COMPUTER APPLICATIONS TO BIOLOGICAL SYSTEMS

Dale L. Jeffries

A Thesis
in
The Department
of
Computer Science

A Thesis
Submitted to the Faculty of Graduate Studies
in Partial Fulfillment of the Requirements for the
Degree of Master of Computer Science

Concordia University
Montreal, Quebec, Canada

September, 1976
ABSTRACT

COMPUTER APPLICATIONS TO BIOLOGICAL SYSTEMS

Dale L. Jeffries

This thesis develops three new techniques for the following:

a) The conversion of any rate dependent system conceptually modelled by Michaelis-Menten notation into simultaneous ordinary differential equations.

b) The computation of values of the observable variables given the initial values of the components and the parameters using a variable order of approximation and variable step size Taylor Series.

c) The iterative determination of extrema of an objective function. By using the weighted least squares function as the objective function the technique was used to estimate the values of parameters consistent with experimentally observed data.

The use of these procedures is illustrated for the modelling and model fitting of biologically important kinetic systems, namely enzyme reaction systems and pharmacokinetic multi-compartment systems.
ACKNOWLEDGMENT

I wish to express my gratitude to Dr. K.V. Leung for his patience in teaching a biologist the mathematics required to write this document, and whose enthusiasm for work was contagious. I want to thank Dr. E.B. Newman for encouraging the question, why, and hence starting me towards becoming a scientist.

A special thank you goes to those whose interest in my personal achievements made this thesis possible: my parents without whom this thesis would not be written till tomorrow and Cornelia without whom it would have stayed a rough draft.
CONTENTS

CHAPTER 1  AN INTRODUCTION TO COMPUTER MODELLING AND MODEL FITTING OF BIOLOGICAL SYSTEMS

CHAPTER 2  A TECHNIQUE FOR THE NUMERICAL INTEGRATION OF SIMULTANEOUS NON-LINEAR DIFFERENTIAL EQUATIONS

2.1 Introduction 8

2.2 The Variable Step Size Variable Order of Approximation of the Taylor Series Solution for the Initial Value Problem 11

2.3 The Computer Programme 14

CHAPTER 3  PARAMETER OPTIMIZATION; THE MODEL FITTING PROCESS

3.1 Introduction 24

3.2 The Finite Element Method for Finding Local Extrema 26

3.3 The Computer Programme 30
CHAPTER 4  SIMULATION OF ENZYME KINETIC SYSTEMS WITH TIME-LAG

4.1 Introduction  33
4.2 Mathematical Description of the Enzyme System and the Kinetic Equations  36
4.3 Illustration of the Use of the Model Solver with an Enzyme System  41

CHAPTER 5 DETERMINATION OF THE PARAMETERS OF AN ENZYME KINETIC SYSTEM

5.1 Introduction  50
5.2 Illustration of the Model Fitter to Find Extrema  51
5.3 Determination of the Values of the Parameters for an Enzyme System  55

CHAPTER 6 PHARMACOKINETICS

6.1 Introduction  58
6.2 Simulation of an Intravenous Dose of a Drug in Man by Compartmental Analysis  61
CHAPTER 7 THE DETERMINATION OF THE TRANSFER RATES FOR MULTI-COMPARTMENT MODELS

7.1 Foreword 67
7.2 A Model for a Three Compartment Model 67
7.3 The Determination of the Transfer Rates for a Three Compartment Model 69

CHAPTER 8. DISCUSSION 74

BIBLIOGRAPHY 80

APPENDIX I THE MODEL SOLVER 90
AI.1 The FORTRAN IV Computer Listings 91

APPENDIX II THE MODEL FITTER 97
AII.1 The FORTRAN IV Computer Listings 98
AII.2 The APL Computer Listings 102

APPENDIX III THE PLOTTER 105
AIII.1 Documentation 106
AIII.2 Computer Listings 107
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Flow Chart of Numerical Integration Program ENZ02.</td>
<td>21</td>
</tr>
<tr>
<td>4.1</td>
<td>Graphical Display of Digital Computer Solution of Reversible-One-Enzyme System with $K_1 = K_2 = K_3 = 1, K_4 = 0$, $X_1(0) = 1, X_2(0) = 0, X_3(0) = 8$, $X_4(0) = 0$.</td>
<td>44</td>
</tr>
<tr>
<td>4.2</td>
<td>Display of 10 Seconds of the Full Time Course of the Reversible Enzyme System with $K_1 = K_2 = K_3 = 1, K_4 = 0$ and the Realistic Relationship where the Enzyme Concentration Exceeds that of the Substrate i.e., $X_1(0) = 8, X_2(0) = 1, X_3(0) = 1$, $X_4(0) = 0$.</td>
<td>45</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>4.3</td>
<td>Graphical Display of 0.2 Seconds Time Course of Computed Concentrations for $X_2$ and $X_4$ of the Very &quot;Stiff&quot; Reversible-One-Enzyme System with $K_1=K_3=10^{-9}$, $K_2=K_4=10^3$, $X_1(0)=10^{-6}$, $X_2(0)=10^{-3}$, $X_3(0)=X_4(0)=0$.</td>
<td>48</td>
</tr>
<tr>
<td>4.4</td>
<td>Graphical Display of $4 \times 10^{-5}$ Seconds Time Course of Computed Concentrations for $X_1$, $X_3$, and $X_4$ of the Very &quot;Stiff&quot; Reversible-One-Enzyme System with $K_1=K_3=10^{-9}$, $K_2=K_4=10^3$, $X_1(0)=10^{-6}$, $X_2(0)=10^{-3}$, $X_3(0)=X_4(0)=0$.</td>
<td>49</td>
</tr>
<tr>
<td>5.1</td>
<td>Contours of Equal Values of the Test Function $F(Z_1, Z_2)$ as in formula (5.2-1).</td>
<td>53</td>
</tr>
<tr>
<td>6.1</td>
<td>Two Compartment Open System with Rapid I.V. Injection of Dose D.</td>
<td>62</td>
</tr>
</tbody>
</table>
6.2 Simulated Display of 5 Hours of the Time Course for the Two Compartment Model of the Distribution of the Antibiotic Amoxicillin after Intravenous Injection with $k_{12} = 1.217$, $K_{21} = 2.523$, $K_{1} = 1.493$ and the Initial Conditions $X_1(0) = 250$, $X_2(0) = 0$, $X_3(0) = 0$. Compare with Figure 6.3.

6.3 Simulated Display of 5 Hours of the Time Course for the Two Compartment Model of the Distribution of the Antibiotic Ampicillin after Intravenous Injection with $k_{12} = 0.53$, $K_{21} = 1.014$ and $K_{1} = 1.413$ and Initial Conditions $X_1(0) = 250$, $X_2(0) = 0$, $X_3(0) = 0$. Compare with Figure 6.2.

7.1 The Three Compartment Model with I.V. Bolus Injection into the Central Compartment.
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.2</td>
<td>Display of the Simulated Three Compartment Model using Equations (7.2-2), the Integrated Form of the Simultaneous Differential Equations (7.2-1).</td>
<td>71</td>
</tr>
<tr>
<td>7.2</td>
<td>Display of the Simulated Three Compartment Model using the Simultaneous Differential Equations (7.2-1) Directly using the Numerical Integration Programme of Chapter 2.</td>
<td>72</td>
</tr>
<tr>
<td>8.1</td>
<td>Recommended Standard Operating Procedure for the Numerical Integration of Simultaneous Differential Equations.</td>
<td>75</td>
</tr>
<tr>
<td>8.2</td>
<td>Recommended Standard Operating Procedure for the Optimization of a Set of Parameters to Best Fit Some Experimental Observations with a Theoretical Model.</td>
<td>77</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Namelist Block Contents for the Control Constants of Numerical Integration Programme ENZ02.</td>
<td>19</td>
</tr>
<tr>
<td>2.2</td>
<td>Important Variables of Programme ENZ02 and Their Description.</td>
<td>20</td>
</tr>
<tr>
<td>3.1</td>
<td>A List of the Computer Routines Used in the Model Fitting Programme.</td>
<td>32</td>
</tr>
<tr>
<td>5.1</td>
<td>Coordinates of Five Points of the Sequence that Converges From (0.2000, 0.4000) to the Maximum Point of ( P(Z_1, Z_2) ).</td>
<td>54</td>
</tr>
<tr>
<td>5.2</td>
<td>Coordinates of Four Points of the Sequence that Converges From (0.4500, 0.4500) to the Minimum Point of ( P(Z_1, Z_2) ).</td>
<td>54</td>
</tr>
<tr>
<td>5.3</td>
<td>Demonstration of the Optimization Programme Using One Thousand Theoretical Data Points From the Model Discussed in Section 4.3 with ( k_1 = k_2 = k_3 = k_4 = 1 ).</td>
<td>57</td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>7.1</td>
<td>Comparison of Results Obtained From Fitting Two Different Functions Representing the Model Shown in Figure 7.1 Consistent with the Experimentally Obtained Data of Appendix IV.</td>
<td>73</td>
</tr>
</tbody>
</table>
CHAPTER 1

AN INTRODUCTION TO COMPUTER MODELLING
AND MODEL FITTING OF BIOLOGICAL SYSTEMS

A fundamental aim of experimentation in the physical sciences is to determine the elementary concepts underlying an observed phenomenon and to establish generalizations that can predict the relation between these and other phenomena. The biological sciences for many years were strictly qualitative. For a quantitative description it is necessary to derive a well defined and rigorous formalization of the system being investigated. Computation requires this formulation to be in the form of a series of algorithms, a mathematical model.

A mathematical model of a biological system consists of variables, some of which are observable, and parameters. The form of the model and value of the parameters determines the value of the variables. Thus once the form of a model has been constructed consistent with contemporary theory, in terms of its parameters, experimental data may then be used
to compute the best statistical estimate of a given set of parameters. Furthermore, within the range of the model, once the parameters are determined it is possible to predict the behavior of the system being modelled for any set of initial conditions.

The computational solution of the mathematical model is often referred to as simulation, but in order to differentiate it from stochastic solutions it shall be referred to here as "modelling". The process of choosing optimal values for the parameters to make experimental data appear "most likely", henceforth will be referred to as model fitting.

The scope of this thesis will be the modelling and model fitting of two aspects of mammalian drug metabolism. The two kinetic systems: the binding of drugs to serum proteins, the multi-compartment model and enzyme reaction systems have been investigated for various reasons leading to certain common goals: to derive general mathematical relations from which it is possible to describe the metabolism and distribution of substances in the body, the mechanisms of this distribution and the design of improved drugs for desired duration and strength.

Most biological processes, such as the above, are sufficiently complex that their time dependent
components cannot be simultaneously observed. Furthermore, it is only in very rare situations that concentrations of intermediate compounds may be observed continuously rather than at a few discrete instants in time. Thus it is necessary to use indirect methods to obtain quantitative information about interactions that occur in metabolic processes (Chance 1969). Representation by a mathematical model is particularly attractive if some predicted values may be compared with experimental measurements.

The mathematical models for each of the systems being considered may be represented by a series of simultaneous non-linear differential equations. The explicit integration of these equations would allow the exact evaluation of the system variables for a given set of parameter values. For these systems, however, the explicit integration expressions which have been presented require a number of simplifying assumptions which limit the applicability of the expressions. The equations for the enzyme systems are even too difficult to explicitly integrate even with simplifying assumptions. To overcome these difficulties many investigators have turned to the use of numerical integration requiring the use of computers. A good summary of the computational
methods and computer utilization for biological systems is given by Garfinkel et al (1970).

Many enzyme catalysed reactions lend themselves to characterizations and simulation by computer methods. Few real enzymes systems have of yet been successfully studied. The metabolic pathway that has been most thoroughly simulated with computers is glycolysis (Garfinkel 1970, Garfinkel 1966, Kerson 1967, Mantle 1969). Several enzymes in the pathway have been studied in detail (Haut 1974, Garfinkel 1966, Mantle 1969) and it is sometimes useful to use computer simulation to differentiate between two possible reaction mechanisms (Gulbinsky 1968).

When dealing with pharmaceutically active compounds, drugs, it is not only the action of the drug at the site of action (usually an enzyme) that determines the intensity or duration of its chemotherapeutic effect, but also, the amount of drug that gets there and stays available. While enzymology allows the investigation of the mode of action of a drug, compartmentation studies allow the study of the amounts of drug and metabolites in the available compartments of the organism such as blood, tissues and excreta. Kinetic models are established to quantify and predict these processes for appropriate dosage regimes. Pharmacokinetics
is the study of the derived constants and quantifiable concepts, such as volumes of distribution and rates of transport between compartments (Wagner 1975 pp. 1-5). These parameters become the focus of attention, rather than the observed data.

While pharmacokinetics allows the investigation of the time course of absorption, distribution and excretion of compounds in an intact organism and enzyme kinetics allows the investigation of interactions and effects of compounds on enzyme catalysed mechanism, the study of the binding of drugs to serum proteins has been reported to allow the pharmacologist to control both the activity and duration of action of a drug by regulating its rate and degree of binding to the serum protein.

The emphasis of this thesis will be on the development and automation of two procedures for the manipulation of kinetic models especially those described above.

The first procedure, based on a variable order, variable step size Taylor series summation, allows the numerical integration of simultaneous differential equations. The development of this procedure is contained in chapter 2 along with a description of its computer implementation. Chapters 4 and 6 illustrate its usage for the
simulation of some enzyme kinetic systems and various multi-compartment models respectively. Although the procedure is not faster or easier to use than the routines already found in the literature, it is however, the only procedure with complete numerical error control.

The second procedure is an iterative finite element method of optimally selecting a set of model parameters to find the critical values of an objective function. When the selected objective function is the sum of squared differences between some experimentally obtained values and corresponding predicted values, the procedure will best fit the parameters to the data in a least square sense. Chapter 3 describes the formulation and the general computer implementation of the procedure. Chapter 5 and 7 are used to illustrate the application of the parameter-fitting procedure to kinetic systems. This procedure differs from other optimizing techniques in that it always converges to the optimal parameter values in the sub-space of the initial estimates. Unfortunately, if the initial estimates are within a sub-space containing local, rather than global optimal values, the optimizer will select the local optimum. The previously reported techniques (see chapter 3 for review) do not necessarily stay
within the initial value sub-space and thus are able to find the global optimum even after poor initial estimates. However, they may still find local minima within other sub-spaces even when their initial estimates place them within the sub-space containing only the global optimum.

The computer routines described in chapters 2 and 3 have been used extensively. The FORTRAN listings for the solving of the differential equations can be found in APPENDIX I, while the listings for the Model Fitter can be found in Appendix II. The APL workspace equivalent to the FORTRAN programme can also be found in APPENDIX II.
CHAPTER 2

A TECHNIQUE FOR THE NUMERICAL INTEGRATION OF
SIMULTANEOUS NON-LINEAR DIFFERENTIAL EQUATIONS

2.1 Introduction

The biological systems presented in later chapters are rate reactions and have been modelled by a series of difference differential equations. A full description of the biological system with respect to the time dependence of the concentrations of individual components could be obtained through explicit integration of the simultaneous differential equations that have been used to describe the system. When the explicit integration is felt to be either too difficult or is filled with too many simplifying assumptions it is possible to employ numerical integration techniques.

Computational techniques to predict the values of the variables for given values for the model's parameters have been described by various authors. Some approaches have been based on the use of Runge-Kutta or point-slope techniques (Chance 1969, Chandler 1972, Bates 1973, Chance 1960, Green 1970,

Methods based on direct summation of the Taylor Series expansion have the attraction that any desired numerical accuracy may be achieved by inclusion of a sufficiently large number of terms of the series. However, inclusion of a large number of terms has the serious disadvantage of requiring the computation of high order derivatives (Froberg 1969).

In choosing an integration formula, there are many factors to be considered: computational accuracy, speed of computation, solution stability, whether it is self-starting and an estimation of the incurred error. These considerations are discussed in detail below.

The accuracy of the simulation of a modelled phenomena is subject to three kinds of errors at each step of the integration. The first kind is called round-off error, and is due to insufficient computer word length, the second kind is truncation error, and is due to the inadequacy of an integration formula to compute the integration step to sufficient
accuracy. For most of the point-slope techniques it is the truncation error which predominates. The third type of error is propagated error, which is due to a combination of the first two types of error propagating from step to step. A numerical integration technique is obviously unstable if the propagation error is unbounded. The simulation of biological systems are usually of long real-time duration and, therefore, contain a very large number of integration steps. Thus, the selected integration technique should be required to pay special attention to propagation, instabilities and oscillations (Hemker 1972).

The numerical solution of a differential equation requires many integration steps; consequentially, the computation time is directly proportional to the number of numerical integration steps required. For any integration step the computer computation time is directly proportional to the time required for solving the integration formula. Therefore, the speed of a simulation is critically dependent on the size of each integration step and efficiency of the particular formula selected to perform the numerical integration. For a desired accuracy there exists a combined error that is optimum in time.

In the following sections of this chapter a
method is described for solution of simultaneous difference differential equations. The method is based on the use of a Taylor Series expansion which selects an optimum step size and order of approximation. It is demonstrated that this combination will give control of truncation and propagation error without significantly increasing computation time compared to other techniques.

The last section of this chapter discusses the computer programme implementing the algorithm. When used with a CDC CYBER 70 computer the FORTRAN IV programme requires less than 22,000 words of 60 bit memory for the treatment of up to forty simultaneous differential equations. The programme with no modifications has been successfully compiled and executed on the XEROX SIGMA 7, the AMDAHL '470-V6 (using IBM FORTRAN-H) and on the mini computer MODCOMP IV.

2.2 The Variable Step Size Variable Order of Approximation of the Taylor Series Solution For the Initial Value Problem

The determination of the full time course of a kinetic system simplifies to the solution of a series of simultaneous ordinary differential equations whose initial values are known. For the purpose of this
chapter the technique for transforming the components of the kinetic equations into the differential equations will be assumed. The technique is provided in exhaustive detail in chapter 4.

Assume there are \( N \) components \( x_n(t) \) whose time course \( t = 0 \) to \( T \) is desired and that \( x_n(0) \) is known.

Let \( h_r \) denote the size of the \( r \) th step. Let \( t_0 = 0, x_{n,r} = x_n(t_r) \), and

\[
\begin{align*}
t_{r+1} &= t_r + h_{r+1} \\ x_{n,r}^{(q)} &= \left\{ \frac{d^q}{dt^q} x_n(t) \right\}_{t=t_r}
\end{align*}
\]

(2.2.1)

(2.2.2)

For each component \( x_n(t) \) the successive values of \( x_{n,r} \) at the grid points \( t_r \) may be computed from the Taylor Series approximation

\[
x_{n,r+1} = x_{n,r} + h_{r+1}^{1} x_{n,r}^{(1)} + \cdots + \frac{(h_{r+1}^{Q})}{Q!} x_{n,r}^{(Q)}
\]

\[
= \sum_{q=0}^{Q} \frac{h_{r+1}^{q}}{q!} x_{n,r}^{(q)}
\]

(2.2.3)

With the assumption that the values of \( x_{n,r} \) and
$x_{n,r}^{(q)}$ may be calculated at the grid points $t_r$. When there is a non-zero time lag such that $(t_r - \text{time lag})$ is not one of the grid points $t_r$ then the values of $x_{n,r}^{(1)}$ (time lag) may be computed using an interpolating polynomial between grid points.

Therefore, all that is required in order to solve the problem is a technique to obtain successive derivatives of the function $x_{n,r}^{(1)}(t)$ from the given first derivative $x_{n,r}^{(1)}$ and the initial value $x_{n,0}$.

Let $X = \text{a vector of } x_1(t), \ldots, x_n(t)$

$Y = \text{a vector of linear multiples of } x_{n,r}^{(1)}(t - \text{time lag})$

$B = \text{a matrix of constants}$

such that

$$\frac{d^{(q+1)} X}{dt^{(q+1)}} = B \begin{bmatrix} \frac{d^{(q)} Y}{dt^{(q)}} \end{bmatrix} \quad (2.2-4)$$

then the required derivatives may be obtained from successive evaluations of equation (2.2-4). This is possible since for all models considered in this paper $d^{(1)} Y/dt^{(1)}$ may be computed using the product rule and known values of $x_n^{(q)}(t)$. 

The computer implementation terminates
computation of the series (2.2-3) when

\[ \text{MAX} \left[ \frac{h^q}{q!} x^{(q)}_{n,r} \right] \leq \varepsilon \]  \hspace{1cm} (2.2-5)

where \( \varepsilon \) is some preassigned tolerance. If equation (2.2-5) has not been satisfied after the computation
of a reasonable number of terms the series may be
re-computed with \( h_r = h_r / 2 \). It should be noted that
the derivatives need not be re-computed to re-sum
the series with the new step size.

Thus during each successive step of the
integration in the approximation of \( x_{i,r+1} \) the last
term used in (2.2-3) has a magnitude not greater
than \( \varepsilon \).

2.3 The Computer Programmes

The computer routines described below were
designed to solve homogenous differential equations
of up to second order. For practical purposes it
became necessary to include two extra modules
specific to each biological model. The first of
these extra modules contained a compiler which
converted a system of rate equations in a notation
easy and familiar to the non-computer scientist into
an efficient mathematical notation. The second
module contains an interactive graphic program for the production of graphs of two dimensions, e.g., component concentration versus time.

2.3.1 Module Definitions

MODULE ONE:

For the conversion of external notation of a model into the internal matrix notation.

The module was first created to convert a modified Michaelis-Menten notation for enzyme systems. Minor changes have made it applicable to the rate equations for multi-compartment drug distribution models.

INPUT:

The input to this module is the desired mechanism in free format with the following sequence.

1. rate constant
2. the reaction in Michaelis-Menten notation with an equal sign replacing the usual arrow
3. the initial concentrations of the components
4. program parameters, the time slice for integration, print intervals, total simulation time, number of allowable reductions of the step size, the maximum number of terms from the Taylor Series to include, and finally the
convergence criteria for each step.

OUTPUT:

The output from this module contains the computer programme diagnostic error messages for the compilation of the biological mechanism and its resulting matrix form. The matrix form and the programme parameters are used as the sole input to module two and as such will described within module two.

EXTERNAL ROUTINES REQUIRED:

Programme INTERP utilizes two subroutines and one function.

1) OUT handles the interactive generation of the desired output

2) TABSET generates a set of state tables for the compiler

3) ERROR handles the correction and re-entry of interactive errors

MODULE TWO:

For the solution by numerical integration of differential equations implemented with both a variable order of approximation Taylor Series Summation.

INPUT:

The module contains only one program requiring input. This program ENZ02, requires the following
elements to be contained in specific blocks, such that ENZ02 may utilize the NAMELIST feature of FTN to obtain the data (Table 2.1).

OUTPUT:

The resulting output from this module is the creation of one file containing the time or total step size in column one followed by the various components at this time.

EXTERNAL ROUTINES:

Upon successful data entry routine ENZ is called by programme MAIN. Routine ENZ performs the majority of the calculations for this module and calls the following routines:

1) Y1    2) MATMUL    3) FACT1

ENZ integrates the given model using step size h and equations (2.2-1 to 2.2-5). Equation (2.2-5) Taylor Series is evaluated for each function but since the right-hand side of equation (2.2-3) requires higher order derivatives than are provided as initial conditions they must be computed using equation (2.2-4). This function describes a recursive algorithm for the computation of the higher derivatives.

The variables used and the description of them are contained in Table 2.2, while the flowchart in
Fig 2-1 describes the flow of the programme and the interaction of the variables.
Table 2.4: Namelist block contents for the control constants of numerical integration programme ENZ02.

<table>
<thead>
<tr>
<th>NAMELIST NAME</th>
<th>VARIABLE</th>
<th>PARAMETER</th>
</tr>
</thead>
<tbody>
<tr>
<td>PARAM,</td>
<td>DT</td>
<td>step size</td>
</tr>
<tr>
<td></td>
<td>MIN</td>
<td>minimum convergence criteria</td>
</tr>
<tr>
<td></td>
<td>MAX</td>
<td>total simulated time</td>
</tr>
<tr>
<td></td>
<td>NT</td>
<td>number of reduction of the step size</td>
</tr>
<tr>
<td></td>
<td>NI</td>
<td>number of terms to use from the Taylor Series maximum</td>
</tr>
<tr>
<td>TRANS-</td>
<td>B</td>
<td>the constant matrix containing the process rates</td>
</tr>
<tr>
<td>VARIB</td>
<td>X</td>
<td>the initial values of the components</td>
</tr>
<tr>
<td></td>
<td>ITN</td>
<td>multi-order derivative flag</td>
</tr>
</tbody>
</table>
Table 2.2: Important variables of programme ENZ02 and their description.

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>HV(I,J)</td>
<td>$\frac{(\frac{DT}{2^J})}{J!}$; the coefficient for the Taylor Series</td>
</tr>
<tr>
<td>NO</td>
<td>Order of differentiation being computed</td>
</tr>
<tr>
<td>N</td>
<td>Number of components</td>
</tr>
<tr>
<td>B</td>
<td>The product of the occurrence matrix and the constant matrix provided as input</td>
</tr>
<tr>
<td>X(I,NO)</td>
<td>The value of the NO th derivative for the I th component</td>
</tr>
<tr>
<td>XX(I)</td>
<td>The derivative of $X^I(t)$ computed using the binomial expansion of the &quot;product&quot; rule</td>
</tr>
<tr>
<td>H</td>
<td>The current value of the step size</td>
</tr>
<tr>
<td>ITN(I)</td>
<td>Flags the order of the I th component</td>
</tr>
</tbody>
</table>
Figure 2.1: Flowchart of Numerical Integration Programme ENZ02.
Figure 2.1 (continued): Flowchart of Numerical Integration Programme ENZ02.
MODULE THREE:

This module consists of two programs. These programs perform interactive two dimensional graphics of a limited nature. The first routine, PLOT, is a line printer plotter generating a graph of the functions over the various steps. It handles up to six graphs per page and ten graphs per run.

The second routine, PLOTIT, uses the TEKTRONIX 4017 software and hardware to create continuous curve graphs of up to ten functions, per display. Since the routines are interactive each display is one run but each session may contain many runs. The routines allow the manipulation of all the graph parameters. The interactively variable parameters are: the scale for the X and Y directions, labels for the axes, maximum and minimum of the displayed graph, plotting symbols, and the number of points to be plotted.

EXTERNAL ROUTINES REQUIRED BY PLOTIT:

1) GRAPH  2) INIT   3) FINIT

All the above are present in the advanced graph package of the TEKTRONIX library.
CHAPTER 3

PARAMETER OPTIMIZATION; THE MODEL FITTING PROCESS

3.1 Introduction

This chapter describes an iterative procedure for the determination of the values of parameters consistent with a set of experimental observations. The parameters are the unknown variables, experimentally unavailable, whose values are required for accurate simulation of events.

Many techniques are available for finding the best fit of some set of model parameters. Most of these procedures are based on the minimization of a function, usually the least-square function. The techniques most often applied to biochemical systems similar to the systems discussed in chapters 5, 7 and 8 are adequately summarized by Garfinkel et al (1970). Experiences in the practical use of optimization for a chain reaction system are described by Chance et al. (1970) and Wagner et al (1971). The specific non-linear optimization procedures have been classified by Swan as

1) direct search methods
2) gradient methods
The various gradient methods are described and compared by Heaps et al. (Heaps, Wells 1965). The gradient methods include: the circle or lattice search methods, the modified random search method and the interpolation method. These methods optimize only one parameter at a time, i.e., all but one of the parameters designated for optimization are fixed, the remaining one being adjusted to best fit the observations. When a certain accuracy is reached with the particular varied constant, it is held fixed and the process is repeated with another, until each in turn has been optimized. It can be expected that optimization with respect to later parameters will perturb the earlier determined parameters from their individual optimum values. Obviously then, it is possible to imagine a situation where convergence will not occur, since the global optimum may be totally divergent from the optimum of the whole set of parameters. Hence, it is desirable to obtain sub-optimum values for the individual values simultaneously which describe a global optimum. Such a method is described below, in the following section.

The finite element method of parameter estimation is as susceptible as the above methods of finding a local rather than a global critical value.
However, consider the K-dimensional space in which the parameters being varied are taken as coordinates. The true global optimum values, local minima, are labelled $P_i$. Now imagine $R_j$ non-intersecting regions each containing one point of $P_0$ to $P_i$. The gradient and other methods when initialized within region $R_i$ do not always converge to the critical value within that region, (Curtis, Chance 1972) whereas the finite element method converges to the maximum or minimum within region $R_i$. This feature is both an asset and a liability since if the initial estimates are poor and thus outside of the region $R_0$ the method cannot converge to $P_0$, but if the estimates are good, i.e., within $R_0$, then it may be confidently assumed that the parameters will converge to the desired global optimum.

3.2 The Finite Element Method For Finding Local Extrema

Let us assume there is an objective function $F$ dependent on a series of parameters $z_1$ to $z_s$ such that

$$F = F(z_1, z_2, \ldots, z_s)$$  \hspace{1cm} (3.2-1)

whose extrema are desired. Suppose that near to its minimum or maximum value the function $F$ may be
approximated by a quadratic expression $\overline{F}(z)$ involving the $z$'s and hence regarded as defining a quadratic function in space of $S$-dimensions. Let $\overline{z} = (z_1, \ldots, z_S)$ be a point that represents an estimate of the optimum $z$ in the region of the global optimum. The equation of $\overline{F}(z)$ may be determined after computation of $\frac{1}{4}(S+1)(S+2)$ values of $F$ at points close to $\overline{z}$ in $S$-dimensional space.

It proves convenient to select the points $z(i,j)$ where $i = 0, 1, 2 \ldots S$ and $j = i+1, i+2, \ldots S$, where the coordinates $z_1(i,j), \ldots, z_S(i,j)$ given by

$$z_r(i,j) = \overline{z}_r - \frac{h}{S+1}$$

$$+ \frac{1}{2} \delta_{i,j}(1-\delta_{i,0})(1-\delta_{j,0})(\delta_{i,r}+\delta_{j,r}) \quad (3.2-2)$$

where $h$ is a small positive number, $\delta_{i,j}$ denotes 1 if $i = j$ and 0 otherwise, and $r = 1, 2, \ldots, S$. The points $z(i,i)$ lie at the corners of an $S$-dimensional tetrahedron of centre $\overline{z}$, and edge length $h$, with $z(0,0)$ at the rectangular apex. The points $z(i,j)$ with $i < j$ are at the midpoints of the edges that join $z(i,i)$ and $z(j,j)$. Let $F(i,j)$ denote the computed value of $F$ at the point $z(i,j)$ so that

$$F(i,j) = F(z(i,j)) \quad (3.2-3)$$
The quadratic interpolation function $\tilde{F}(Z)$ may then be defined in the form

$$F(Z) = \sum_{i=0}^{S} \sum_{j=0}^{S} H(i,j)(\alpha_i, \alpha_j) F(i,j)$$  \hspace{1cm} (3.2-4)

where

$$a_i = \frac{z_i - z^{(0,0)}}{h} \quad \text{for} \quad i = 1, 2, \ldots, S  \hspace{1cm} (3.2-5)$$

$$a_0 = 1 - \sum_{i=1}^{S} a_i  \hspace{1cm} (3.2-6)$$

$$H(i,j)(\alpha_i, \alpha_j) = 2 \left[ 2 \delta_{i,j} \right] a_i a_j - \frac{1}{2} \delta_{i,j}$$

It may be verified that definition of $\tilde{F}(Z)$ in the form (3.2-4) ensures that it takes the values of $F(Z)$ at each of the points $Z^{(i,j)}$.

The stationary value of $\tilde{F}(Z)$ occurs at the point $Z$ for which $\partial F/\partial z_s = 0$ for all $s = 1, 2, \ldots, S$. By virtue of equations (3.2-4) - (3.2-7) the condition may be expressed in the form of the following set of $S + 1$ linear equations

$$\sum_{i=0}^{S} \left[ F(i,s) - F(0,s) \right] a_i = \left[ F(s,s) - F(0,0) \right]  \hspace{1cm} (3.2-8)$$
for $s = 1, 2, \ldots, S$

$$\sum_{i=0}^{S} a_i = 1 \quad (3.2-9)$$

and solved to determine $a_0, a_1, \ldots, a_S$ and hence $z_1, z_2, \ldots, z_S$ by virtue of (3.2-5). In equation (3.2-8) the value of $F^{(i,s)}$ for $i > s$ is defined to be $F^{(s,i)}$.

For an initial point $\bar{Z}$ the above procedure requires computation of the value of $F$ at a set of nearby points $Z^{(i,j)}$, followed by solution of a set of linear equations (3.2-8, 9) to determine the point $Z$ at which $F$ is stationary. This point constitutes a prediction of the point at which $F$ is stationary; it may then be used instead of $\bar{Z}$ as the starting point of a further prediction. The process may be continued to determine a sequence of predicted points

$$Z(0), Z(1), Z(2), \ldots \quad (3.2-10)$$

After determination of a point $Z(k)$ of the sequence (3.2-10) a test may be made to check whether the relations

$$0 \leq z(k)_i - z(k-1)_i^{(0,0)} \leq h \quad (3.2-11)$$

$$0 \leq \sum_{i=1}^{S} z(k)_i - z(k-1)_i^{(0,0)} \leq h \quad (3.2-12)$$
are satisfied for all \( i \) in the range from 1 to \( S \).

If the relations are satisfied then \( Z(k) \) lies within the tetrahedron centered at \( Z(k-1) \), and for determination of \( Z(k+1) \) the value of \( h \) may be replaced by \( h/2 \). It may be shown that if \( Z(k) \) is sufficiently near the stationary point of \( P(Z) \) the accuracy of each successive prediction is proportional to \( h^2 \).

3.3 The Computer Programme

The programme described here was designed to implement the finite element extrema detection algorithm developed in section 3.2.

The computer implementation was written in FORTRAN IV originally for the CDC 6400 using the FTN compiler. It was then converted to execute on the ZEROX SIGMA series computer using the ZEROX FORTRAN compiler, and for demonstration purposes, a version of the programme was written in ZEROX APL+. The source of the latter two versions may be found in appendix II.

Experiments with the APL version of the programme showed that it was the function being analysed for extrema that limited both the accuracy and speed of computer solution. Therefore, when the algorithm was coded into FORTRAN IV it was decided
to emphasize usability and modularity rather than speed. Even with this decision the bottleneck is still in the evaluation of the objective function.

The modularity of the programme allows expansion and modification for the testing and implementation of different applications while requiring only a minimal amount of programming.

The version of the extrema evaluation programme found in appendix II contains seven routines. A list of these routines, their description, and where applicable, the equations from section 3.2 that they numerically evaluate can be found in table 3.1.

To execute the programme the user should refer first to chapters 5 and 8. The user then must write and test the objective function EPSIL making sure that it is a smooth continuous function. If necessary the user can modify the routine main to simplify the input of initial estimates.
Table 3.1: A List of the Computer Routines Used in the Model Fitting Programme.

<table>
<thead>
<tr>
<th>Routine</th>
<th>Description</th>
<th>Equations Evaluated</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAIN</td>
<td>Reads estimates of extrema, controls sequence of calls, prints convergence results.</td>
<td>3.2-1</td>
</tr>
<tr>
<td>RD</td>
<td>Reads experimental observations when Least Squares Fitting is to be used. User supplied.</td>
<td></td>
</tr>
<tr>
<td>EPSIL</td>
<td>Evaluates the objective function in terms of the given parameters. User supplied.</td>
<td>3.2-2</td>
</tr>
<tr>
<td>NODES</td>
<td>Computes the co-ordinates of the nodes surrounding the estimated extrema.</td>
<td></td>
</tr>
<tr>
<td>GAUSS</td>
<td>Uses the evaluations of EPSIL at the nodal co-ordinates to create the polynomial describing a quadratic surface.</td>
<td>3.2-4 3.2-7</td>
</tr>
<tr>
<td>DSCRV</td>
<td>Solves a set of simultaneous equations to evaluate alphas.</td>
<td></td>
</tr>
<tr>
<td>ATOK</td>
<td>Uses $z(0,0)$ and alphas to generate coordinates of an improved estimate of the extrema.</td>
<td>3.2-8,9</td>
</tr>
</tbody>
</table>
CHAPTER 4

SIMULATION OF ENZYME KINETIC SYSTEMS WITH TIME-LAG

4.1 Introduction

The majority of important biological systems are controlled by complicated protein catalysts, called enzymes. The amount of research being carried out around the world to elucidate the mechanism of enzyme action has virtually exploded since 1960. These studies have gone much further than just the definition of the stoichiometry of a reaction, and, in the cases of some enzymes, tracer and stop flow techniques have enabled the characterization of some short lived intermediates. However, in general this characterization is still very difficult.

There is a vast and rapidly growing literature on the kinetics of enzyme systems. The fundamentals of the enzyme systems are most adequately described by K.J. Laidler (1958) or J.M. Reiner (1969) while for reviews refer to the articles of R.A. Alberty (1956) and W.W. Cleland (1967).
A full description of an enzyme system with respect to time dependence could be obtained through the explicit integration of the kinetic equations, the differential equations which describe the evolution in time of the chemical reactions (Garfinkel 1970, Chandler 1972). However, integration by analytical methods allowing the solution in the closed form is generally impossible, even for the simplest examples (Bates and Frieden 1973). Thus, in practice, the system of simultaneous kinetic equations must be integrated by numerical methods requiring the use of a computer (Garfinkel 1970).

The kinetic equations involve various parameters whose interaction determines the time course of the system. Amongst these parameters are both processing rates and time-lags. The processing rate represents an averaged constant describing the rate at which an individual chemical reaction occurs. Time-lags generally are involved whenever some chemical reactions are subjected to certain incubation periods.

In the study of an enzymatic process the following two problems are of particular interest:

i) Given an assumed set of kinetic equations compute the concentration of all the intermediate
compounds for the duration of a reaction for which the processing rates and time-lags are assumed known.

ii) Determine the values of the parameters, the processing rates and time-lags, consistent with a set of experimental observations of the system made at discrete times.

Computer techniques for the above two problems have been described in chapters 2 and 3 respectively.

In the present chapter the kinetic equations of a general enzyme system will first be established. These are a set of differential difference equations with polynomial non-linearities (Sorensen and Schack 1972). Then the simulation of a particular enzyme system will demonstrate the solution of the first problem above and will serve to illustrate the use of the computer routines discussed in chapter 2.

In chapter 5 the solution of the second problem formulated above based upon the method of least squares developed in chapter 3 will be described along with an illustrative example.
4.2 Mathematical Description of the Enzyme System and the Kinetic Equations

Let

\[
\{x\} = \{x_1, x_2, \ldots, x_N\}
\]  

be a set of \( N \) chemical compounds \( \{x_1, x_2, \ldots, x_N\} \) in an enzyme system and let \( x_1(t), x_2(t), \ldots, x_N(t) \) be their respective concentrations at time \( t \). We put

\[
\{x(t)\} = \{x_1(t), x_2(t), \ldots, x_N(t)\}
\]  

The enzyme system is assumed to involve \( M \) chemical processes denoted by \( P_{i,j} \). Process \( P_{i,j} \) is represented symbolically by

\[
P_{i,j}: \{x\}_i, j \xrightarrow{k_{i,j}, \sigma_{i,j}} \{x\}_i, j
\]  

where \( k_{i,j} \) and \( \sigma_{i,j} \) are given non-negative numbers termed respectively the processing rate and the time-lag of the process \( P_{i,j} \) and \( \{x\}_i, j \) and \( \{x\}_i, j \) are two disjoint subsets of \( \{x\} \):

\[
\{x\}_i, j \cap \{x\}_i, j = \{\phi\}
\]  

\( \{\phi\} \) is the empty set. The number of the compounds in \( \{x\}_i, j \) is equal to \( i \). Obviously no chemical compound can be on both sides of \( (4.2-3) \) by virtue
of (4.2-4).

The compounds in \([x]_{i,j}^F\) chemically interact with each other and produce the compounds in \([x]_{i,j}^T\). Let
\[
I = \max\{i\} \tag{4.2-5}
\]

\[ J(i) = \text{the number of the processes which have } \ i \text{ compounds on the left hand side of } (4.2-3) \]

Accordingly,
\[
M = \sum_{i=1}^{I} J(i) \tag{4.2-6}
\]

Further, let
\[
\lambda_{i,j}(t) = \text{the product of the concentrations of the } i \text{ compounds contained in } ([x]_{i,j}^F) \tag{4.2-7}
\]

Hence
\[
\lambda_{i,j}(t) = \prod_{k} x_k(t) \text{ for all } [x]_k \in ([x]_{i,j}^F) \tag{4.2-8}
\]

By the mass action law, the contribution of the process \(p_{i,j}\) to the rate of change of the concentration \(x_n(t), n = 1, 2, \ldots, N,\) is given by
\[ (x_n(t))_{i,j} = \lambda_{i,j;n} k_{i,j} \ell_{i,j}(t-\sigma_{i,j}) \]  

(4.2-9)

where the dot denotes the time derivative of

\[ x_n(t), \quad x_n = \frac{dx_n}{dt}, \quad \text{and} \]

\[ \lambda_{i,j;n} = \left\{ \begin{array}{ll}
+1, & \text{if } [x_n] \in ([x])_{i,j}^r \\
-1, & \text{if } [x_n] \in ([x])_{i,j}^l \\
0, & \text{otherwise} \end{array} \right. \]  

(4.2-10)

The rate of change of concentration \( x_n(t) \) is the sum of the contributions from all the \( M \) processes \( P_{i,j} \):

\[ x_n(t) = \sum_{i=1}^{I} \sum_{j=1}^{J(i)} \lambda_{i,j;n} k_{i,j}(t-\sigma_{i,j}). \]  

(4.2-11)

These equations constitute the kinetic equations under consideration.

We now introduce the following column vectors:

\[ x(t) = [x_1(t), x_2(t), \ldots, x_N(t)]^T \]  

(4.2-12)

\[ L_i(t) = [L_{i,1}(t), L_{i,2}(t), \ldots, L_{i,J(i)}(t)]^T \]

\[ L_i(t-\sigma_1) = [L_{i,1}(t-\sigma_{i,1}), L_{i,2}(t-\sigma_{i,2}), \ldots, L_{i,J(i)}(t-\sigma_{i,J(i)})]^T \]
where $T$ denotes the transpose operation. Then, combining Eqs (4.2-11) and (4.2-12), we obtain the following vector-matrix equation:

$$
 x(t) = \sum_{i=1}^{I} A_i \zeta_i (t-\sigma_i) \quad (4.2-13)
$$

where $A_i$ is a constant matrix of order $N \times J(i)$ whose element at the intersection of the $n$-th row and $j$-th column is given by the relationships:

$$
 (a_{i,j}^n) = \begin{cases} 
 +k_{i,j}, & \text{if } [x_n] \in \{[x]_{i,j}^r \}
 \geq \sigma_{i,j}, \\
 -k_{i,j}, & \text{if } [x_n] \in \{[x]_{i,j}^l \}
 \leq \sigma_{i,j}, \\
 0, & \text{otherwise}
\end{cases} \quad (4.2-14)
$$

for $n = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, J(i)$.

We now put

$$
 \sigma = \max \{\sigma_{i,j} \}, \quad (4.2-15)
$$

$$
 i = 1, 2, \ldots, I \text{ and } j = 1, 2, \ldots, J(i)
$$

Clearly, we can assume that all the concentrations $x_1(t), x_2(t), \ldots, x_N(t)$ are completely known before the chemical reactions start.

Put

$$
 x_{n,0}(t) = x_n(t) \text{ for } -\sigma \leq t \leq 0 \quad (4.2-16)
$$

$$
 (n = 1, 2, \ldots, N).
$$
and call \( x_{1,0}(t), x_{2,0}(t), \ldots, x_{N,0}(t) \) as the initial conditions of the enzyme system, and write

\[
x_0(t) = [x_{1,0}(t), x_{2,0}(t), \ldots, x_{N,0}(t)]^T \quad (4.2-17)
\]

We assume that the functions \( x_{n,0}(t) \)'s are all continuous in the initial interval \(-\sigma \leq t \leq 0\).

The kinetic equations (4.2-11), or, equivalently (4.2-13), are essentially differential difference equations and their right hand sides are polynomials in \( x_1, x_2, \ldots, x_N \). Hence, if \( k_{i,j} \)'s are constants or certain smooth functions of \( t \) and \( x_1, x_2, \ldots, x_N \), then the differential difference system (4.2-13) has a unique continuously differentiable solution \( x(t) \) for \( t > 0 \) satisfying the initial condition

\[
x(t) = x_0(t) \quad \text{for} \quad -\sigma \leq t \leq 0. \quad (4.2-18)
\]

This solution depends on \( k_{i,j} \) continuously. For these results and for further developments related to the initial-value problem (4.2-13), (4.2-18) refer to Bellman (1963), Oguzoreli (1966), and El'sgol'ts (1973), and chapter 2 of this work.

When the time-lags \( \sigma_{i,j} \) are all zero equations (4.2-13) reduce to a system of ordinary differential equations with polynomial right hand sides. This type of differential equations has been investigated.
by Beklemisheva (1971) in a general context. Here we shall consider the kinetic equations directly without invoking the results of Beklemisheva (1971) for the case in which all time-lags are zero.

4.3 Illustration of the Use of the Model Solver with an Enzyme System

The most difficult part of any study with a system is the choice of the model to represent the physical system. This selection is neither purely theoretical nor is it purely experimental. Conceptual models are usually used followed by their translation into mathematical notation. This is how the enzyme models are selected. The Michaelis-Menten (Michaelis 1913) form of the model conceptually shows the interaction of the components consistent with the Laws of Mass Action. The differential equations derived using section 4.2 allows the quantifiable prediction of the components for a given model, given parameter values, and known initial conditions.

The procedure discussed in chapter 2 and section 4.1 will be applied in this section to treat the reversible one enzyme single substrate enzyme system. This system is conceptually very important
as it is the basic element of almost all real enzyme systems. As such it has also been studied extensively by other authors (Chance 1960, Kirby 1969, Garfinkel 1961, Garfinkel 1966). The system involves four compounds subject to four chemical reactions as expressed by the Michaelis-Menten equation

\[
\begin{align*}
\frac{k_1}{k_2} x_1 + x_2 & \xrightarrow{k_3} \frac{k_3}{k_4} x_3 \\
& \xrightarrow{k_4} x_1 + x_4 
\end{align*}
\] (4.3-1)

where \(x_1, x_2, x_3\) and \(x_4\) denote respectively the enzyme, substrate, complex, and product. There are no time lags.

In general the concentrations \(x_n(t)\) satisfy the system of differential equations:

\[
\frac{d}{dt} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix} = \begin{bmatrix} -k_1, 0, k_2 + k_3, -k_4 \\ -k_1, 0, k_2, 0 \\ k_1, 0, -k_2 - k_3, k_4 \\ 0, 0, k_3, -k_4 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix}
\] (4.3-2)

The set of equations (4.3-2) has been solved for various combinations of \(k_1\) to \(k_4\) and for a variety of initial conditions. The values selected by B. Chance (Chance 1960) for the parameters,
\( k_1 = k_2 = k_3 = 1, k_4 = 0 \) and the initial conditions, 
\( x_1(0) = 1, x_2(0) = 8, x_3(0) = x_4(0) = 0 \) were used 
to simulate the system (4.2-1). Using the values 
of the step size and convergence criteria equal to 
0.01 and \( 10^{-6} \) respectively, computation of 
concentrations over a time interval of 15 seconds 
led to the results shown in Figure 4.1. The 
simulation required 10.2 seconds of computer time. 
The system has been simulated with the more 
realistic relationship where the concentration of 
of the enzyme exceeds that of the substrate, e.g., 
\( x_1(0) = 8, x_2(0) = 1 \). This results in Figure 4.2. 
As the ratio of enzyme to substrate concentrations 
is increased, the life-time of the complex, \( x_3 \), 
decreases. This accounts for the experimental 
difficulty in observing transient state inter-
mediaries.

The accuracy of the results may be checked 
from the following physical considerations:
(i) The concentrations of \( x_1(t) \) and \( x_3(t) \) should 
be stationary at the same instant. The computed 
results indicate stationary values at \( t = 0.63499 \) 
and \( t = 0.63498 \) respectively.
(ii) Let \( x_4^* \) denote the maximum value of \( x_4(t) \) and 
let \( t^* \) denote the time interval required for \( x_4(t) \)
Figure 4.1: Graphical display of digital computer solution of reversible one-enzyme system with $k_1=k_2=k_3=1$, $k_4=0$, $x_1(0)=1$, $x_2(0)=8$, $x_3(0) = x_4(0) = 0$. 
Figure 4.2: Display of 15 seconds of the full time course of the reversible enzyme system with $k_1 = k_2 = k_3 = 1$, $k_4 = 0$ and the realistic relationship where the enzyme concentration exceeds that of the substrate i.e., $x_1(0) = 8$, $x_2(0) = 1$, $x_3(0) = x_4(0) = 0$. 
to reduce from $x_4^*$ to $x_4^*/2$. The value of $k_3$ should equal $x_2(0)/(x_4^* t^*)$ which is found to be 1.0017 instead of the given value of $k_3 = 1$.

(iii) At any time $t$ the concentrations should satisfy the conditions

$$x_1(t) + x_3(t) - x_1(0) - x_3(0) = 0$$
$$x_2(t) + x_3(t) + x_4(t) - x_2(0) - x_3(0) - x_4(0) = 0$$

It may be verified that the computed concentrations reduce the left hand sides of these equations to values less than $10^{-10}$.

As can be seen the above system (4.3-1), is adequately simulated in less than real time with a high degree of accuracy independent of the initial conditions.

Unfortunately, rate equations for realistic enzyme reactions are more difficult to integrate due to the large differences in the processing rates (Hemker 1972). Such equations are called "stiff" (Chandler 1972, Hemker 1972, Seinfeld 1970) and their solution is much more demanding.

The system of equations (4.3-2) can be transformed into a very "stiff" system. The system
initialized with

\[ k_1 = k_4 = 10^9, \quad k_2 = k_3 = 10^3 \text{ and} \]

\[ x_1(0) = 10^{-6}, \quad x_2(0) = 10^{-3}, \quad (4.3-3) \]

\[ x_3(0) = x_4(0) = 0 \]

makes (4.3-2) a stiff system. With the initial step size and convergence criteria equal to \(10^{-5}\) and \(10^{-3}\) respectively, the solution over a range of \(t\) from 0 to 0.2 required 362 seconds of CPU time, in contrast to 132 seconds required by use of Kirby's stochastic technique (Kirby 1969) and 1086 seconds required by Kirby's use of the method of E. Chance. The graphical displays of the stiff system can be found in Fig. 4.3 and Fig. 4.4.
Figure 4.3: Graphical display of 0.2 seconds time course of computed concentrations for $x_2$ and $x_4$ of the very "stiff" reversible-one-enzyme system with $k_1 = k_3 = 10^9$, $k_2 = k_4 = 10^3$, $x_1(0) = 10^{-6}$, $x_2(0) = 10^{-3}$, $x_3(0) = x_4(0) = 0$. 
Figure 4.4: Graphical display of $4 \times 10^{-5}$ seconds time course of computed concentrations for $x_1$, $x_3$ and $x_4$ of the very "stiff" reversible-one-enzyme system with $k_1 = k_3 = 10^9$, $k_2 = k_4 = 10^3$, $x_1(0) = 10^{-6}$, $x_2(0) = 10^{-3}$, $x_3(0) = x_4(0) = 0$. 
CHAPTER 5

DETERMINATION OF THE PARAMETERS OF AN
ENZYME KINETIC SYSTEM

5.1 Introduction

In chapter 4 it was assumed that the parameters of an enzyme system had been previously determined. The parameters, processing rates and time-lags, are not always directly measurable with present technology. It is therefore desired to determine the parameters of a kinetic system that are consistent with both the hypothesized model and experimentally observed data using some non-laboratory technique.

In order to demonstrate the speed and accuracy of the extrema finding algorithm of chapter 3, the computer version of the algorithm was used to find the extrema of a function with elliptical ridges (section 4.3). The functions in this family pose many difficulties for gradient following methods since the gradient does not always point to the extrema.

The finite element extrema finding program was
then used to select an optimum set of parameters for an enzyme model in a least squares sense (section 4.3). The model solver of chapter 2 was used to simulate the model with parameters iteratively selected by the model fitter.

Various authors have used conventional techniques in order to optimize a set of parameters for enzyme data, such that these parameters meet a test of "goodness of fit" (Garfinkel 1970, Chandler et al 1972). However, these procedures are limited in their ability to optimize more than a few parameters simultaneously (see chapter 3).

5.2 Illustration of the Model Fitter to Find Extrema

The procedure outlined in the previous section 3.2 may be illustrated by application to determine the stationary values of the function

\[ F(z_1, z_2) = \frac{1}{2} \left\{ 1 - \cos 2\pi \left[ \frac{(2z_1 - 1)^2 + (2z_2 - 1)^2}{2} \right] \right\} \quad (5.2-1) \]

\[ \sqrt{\frac{(z_2 - 3z_1)^2}{1 - \frac{z_2}{3}}} \]

Although this function does not arise in connection with an enzyme reaction it is a suitable one to use to test methods for determination of stationary values within the rectangular region.
\[ 0 \leq z_1 \leq 1, \quad 0 \leq z_2 \leq 1 \]  \hspace{1cm} (5.2-2)

since it contains a sharp curved ridge with a maximum value of 0.9947 at point \((z_1, z_2) = (0.2608, 0.5795)\). Determination of this maximum by use of the gradient, and other hill-climbing techniques, was discussed in a previous paper (Heaps 1965).

The contours of equal values of \(F(z_1, z_2)\) are shown in Fig. 5.1. It may be noted that within the rectangular region the function \(F\) also has a minimum point with a value of zero at \((z_1, z_2) = (0.5000, 0.5000)\).

Depending on selection of the initial point the iterative procedure generates a sequence of points that converge either to the maximum or the minimum value of \(F\). The results tabulated in Table 5.1 indicate that from a starting point of \((0.2000, 0.4000)\) the sequence of predicted points converges to the maximum point after about 20 iterations. Similarly, the results tabulated in Table 5.2 indicate that from a starting point of \((0.4500, 0.4500)\) the sequence converges to the minimum point after 15 iterations.
Figure 5.1: Contours of equal values of the test function $F(z_1, z_2)$ as in formula (5.2-1).
Table 5.1: Coordinates of five points of the sequence that converges from \((0.2000, 0.4000)\) to the maximum point of \(F(z_1, z_2)\).

<table>
<thead>
<tr>
<th>Number of Iterations</th>
<th>(z_1)</th>
<th>(z_2)</th>
<th>(F(z_1, z_2))</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.2000</td>
<td>0.4000</td>
<td>0.83247</td>
</tr>
<tr>
<td>5</td>
<td>0.24729</td>
<td>0.48609</td>
<td>0.99145</td>
</tr>
<tr>
<td>10</td>
<td>0.25389</td>
<td>0.56027</td>
<td>0.99448</td>
</tr>
<tr>
<td>15</td>
<td>0.26058</td>
<td>0.57879</td>
<td>0.99469</td>
</tr>
<tr>
<td>20</td>
<td>0.26076</td>
<td>0.57953</td>
<td>0.99469</td>
</tr>
</tbody>
</table>

Table 5.2: Coordinates of four points of the sequence that converges from \((0.4500, 0.4500)\) to the minimum point of \(F(z_1, z_2)\).

<table>
<thead>
<tr>
<th>Number of Iterations</th>
<th>(z_1)</th>
<th>(z_2)</th>
<th>(F(z_1, z_2))</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.4500</td>
<td>0.4500</td>
<td>0.16604</td>
</tr>
<tr>
<td>5</td>
<td>0.4991</td>
<td>0.4982</td>
<td>0.23181</td>
</tr>
<tr>
<td>10</td>
<td>0.4999</td>
<td>0.4999</td>
<td>(7.506 \times 10^{-6})</td>
</tr>
<tr>
<td>15</td>
<td>0.5000</td>
<td>0.5000</td>
<td>0.0</td>
</tr>
</tbody>
</table>
5.3 Determination of the Values of the Parameters for an Enzyme System

When the objective function $F$ is the result of the weighted sum of square differences between observed data and predicted from the model data the object is to find the set of parameters which minimize $F$ and are therefore the optimum parameters in a least squares sense. The use of the least squares estimate carried with it a number of assumptions which are explained in detail by Daniel and Wood (Daniel 1971).

To demonstrate the use of the parameter optimizer for an enzyme system there are two approaches possible. The first is to experimentally obtain some transient state data from the laboratory and fit some set of parameters to the hypothetical model so as to minimize the weighted sum of squares. However, this would not prove anything regarding the optimizer, since there would be no proof of correctness. The second approach is to select some arbitrary model with some set of parameters, collect some simulation data from this model, then see if when given a set of estimated parameters, perturbed from the known parameters, whether the optimizer converges to the actual parameters. This latter approach has the appeal of allowing observation
of how good the initial estimates must be, and how many parameters the optimizer can handle conveniently. Atkins (Atkins 1974) has also previously made, the conclusion that for theoretical investigations into parameter estimation, large amounts of artificial data of known characteristics must be used rather than small amounts of experimentally obtained data.

The enzyme system selected to demonstrate the model fitting process is the same one-enzyme single substrate model of Chance (Chance 1960) as was used in section 4.3.

A thousand data points were collected over a ten second interval with \( k_1 = k_2 = k_3 = k_4 = 1 \) and time-lags all zero. Various estimates were used as initial guesses to these known values. The results are tabulated in Table 5.3. It can be seen from these results that as long as the initial estimates are reasonably close, the technique will converge with an error less than \((\text{step-size optimizer})^2 \times (\text{truncation error of the model solver})\).
Table 5.3: Demonstration of the optimization programme using one thousand theoretical data points from the model discussed in section 4.3 with $k_1 = k_2 = k_3 = k_4 = 1$.

<table>
<thead>
<tr>
<th>Initial Estimate Perturbed By %</th>
<th>Results Converge to Within %</th>
<th># of Iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.02</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>.02</td>
<td>11</td>
</tr>
<tr>
<td>20</td>
<td>.10</td>
<td>25</td>
</tr>
<tr>
<td>50</td>
<td>Did Not Converge</td>
<td>--</td>
</tr>
</tbody>
</table>
CHAPTER 6

PHARMACOKINETICS

6.1 Introduction

The word "pharmacokinetics" means the application of kinetics to "pharmacon", the Greek word for drugs and poisons. The purpose of pharmacokinetics is to study the time courses of drug and metabolite concentrations in various biological compartments, and to construct suitable models to interpret such data (Wagner 1975). Pharmacokinetics reduces pharmacetic data to sets of parameters which are easier to comprehend and to compare from drug to drug. Furthermore, the parameter values determined may be used to make predictions of results of experiments which would be too costly and time consuming to be carried out.

Compartmentalized analysis is an attempt to lump all the chemical reactions, diffusion processes and variations in drug-protein binding for a particular tissue or organ into an averaged process. Therefore, unlike the enzyme system, pharmacokinetics does not investigate the detailed characteristics
of a system, but rather an abstraction emphasizing the conceptual aspects of the model (Robertson 1962, Wagner 1975).

The simplifications of the biological systems to an easily solved form and the importance of drug distribution has made pharmacokinetic studies one of the most advanced areas of computer applications in the biological sciences. The amount of energy invested in the perfection of the conceptual models and the computer tools developed for their computational solutions is just starting to bear fruit. This chapter and the one following demonstrate the general application of the tools described in chapter 2 and 3.

The multi-compartment system can be represented by using Michaelis-Menten notation (Michaelis 1913) to form a conceptual model. Techniques developed in chapter 4 allow the transformation of this form into a series of simultaneous linear ordinary differential equations. In this way the concentrations $x_n(t)$ represent the amount of drug in compartment $n$ at time $t$, while the parameters $k_i$ represent transfer rates between compartments.
\[
\frac{dx}{dt} = BX \quad (6.1-1)
\]

where: \( X \) is an \( N \times 1 \) matrix of compartment concentrations.

\( B \) is an \( N \times N \) matrix of transfer rates.

The compartmental simplification allows the explicit integration of the differential equations using Laplace transforms for the amount of drug in any compartment of a mammillary model. The inverse Laplace transforms converts the simultaneous differential equations into a series of sums of exponential terms shown as (6.1-2) (Wagner 1975, p.58, Nagashima 1968, Teorell 1974).

\[
c_j = \sum_{i=1}^{N} A_i \exp(-\lambda_i t) \quad (6.1-2)
\]

Variables \( A_i \) and \( \lambda_i \) are combinations of the parameters of the kinetic model. Therefore, simulation of a compartmental model may equivalently be performed using (6.1-2) or the actual differential equations (6.1-1). Parameter estimation from observed data can be obtained by fitting the macro-constants of (6.1-2), and then converting them into
the parameters, the micro-constants. Conversely, the parameters may be fitted directly using the differential equations.

The next section of this chapter will investigate the simulation of multi-compartment systems. Chapter 7 will discuss the determination of the parameters iteratively from both the differential equations (6.1-1) and their integrated form (6.1-2).

6.2 Simulation of an Intravenous Dose of a Drug in Man by Compartmental Analysis

Consider the experiments of Zarowny et al (Zarowny 1974) with the semisynthetic penicillin Amoxicillin, where eight adult men were injected intravenously with 250 mg to characterize the drug's disposition. By graphical methods described by Wagner (Wagner 1975) the model was assumed to be a two compartment open-system kinetic model (see Fig. 6.1).

The differential equations of this model are:

\[
\begin{align*}
\frac{d}{dt} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} &= \begin{bmatrix} -k_{12} - k_{el} & k_{21} & 0 \\ k_{12} & -k_{21} & 0 \\ k_{el} & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} \\
\end{align*}
\]  

(6.2-1)
Figure 6.1: Two compartment open system with rapid I.V. injection of Dose $D$. 
Using the parameters $k_{12}$, $k_{21}$, $kel$ reported for a randomly selected subject and the initial conditions the time course of the drug may be followed. For subject 2 (Table I ref. (Zarowny 1974)) the values tabulated are shown in (6.2-2).

$$k_{12} = 1.217, \quad k_{21} = 2.523, \quad kel = 1.493$$

and $$x_1(0) = 250, \quad x_2(0) = x_3(0) = 0 \quad (6.2-2)$$

The simulation using these values forms Fig. 6.2. The computer time to simulate twenty-five hours was twelve CPU seconds with a step size of 0.01 minutes and a truncation error of $1 \times 10^{-6}$.

Simulation studies of this type are very valuable when it is desired to compare the time course of different drugs of the same family or different dosages of the same drug.

For illustration purposes only, let's compare the time courses of Amoxicillin with Ampicillin, two semisynthetic penicillins. The results for Amoxicillin have been observed in Fig. 6.2. The parameters of Macleod et al (1974) for subject number 2 and the initial concentrations for Ampicillin are

$$k_{12} = 0.53, \quad k_{21} = 1.014, \quad kel = 1.413$$

$$x_1(0) = 250, \quad x_2(0) = x_3(0) = 0 \quad (6.2-3)$$
Figure 6.2: Simulated display of 5 hours of the time course for the two-compartment model of the distribution of the antibiotic Amoxicillin after intravenous injection with $k_{12}=1.217$, $k_{21}=2.523$, $kel=1.493$ and the initial conditions $x_1(0)=250$, $x_2(0)=x_3(0)=0$. Compare with figure 6.3.
The time course for each compartment may be observed in Fig. 6.3. The combination of the figures 6.2 and 6.3, the values of the pharmacokinetic parameters, and the observed pharmaceutical effect provide a clear basis for comparing the drugs. Therefore simulation can be seen to be an aid in pharmaceutics and the model solver described in chapter 2 can be seen to adequately perform these simulations.
Figure 6.3: Simulated display of 5 hours of the time course for the two compartment model of the distribution of the antibiotic Ampicillin after intravenous injection with $k_{12}=0.53$, $k_{21}=1.014$ and $kel=1.413$ and initial conditions $x_1(0)=250$, $x_2(0)=x_3(0)=0$. Compare with figure 6.2.
CHAPTER 7

THE DETERMINATION OF THE TRANSFER RATES FOR
MULTI-COMPARTMENT MODELS

7.1 Foreword

In order to perform the simulations of section 6.2 the parameters of the model must be determined consistent with experimental data. There are a multitude of computer programmes available using various techniques to optimize the parameters $A_i$ and $\lambda_i$ of equation (6.1-2) (Garfinkel 1970). This chapter uses the model fitter of chapter 3 to optimize the transfer rates inherent in the actual differential equations derived from a multi-compartment model, i.e., determine the optimum elements of the $B$ matrix of equations (6.1-1).

7.2 A Model for a Three Compartment Model

Let us assume that the disposition of some drug is hypothesized to be the three compartment model shown in Fig. 7.1. The differential equations of Fig. 7.1 are given by (7.2-1).
Figure 7.1: The three compartment model with I.V. Bolus injection into the central compartment.
\[
\begin{bmatrix}
  x_1(t) \\
  x_2(t) \\
  x_3(t) \\
  x_4(t)
\end{bmatrix} =
\begin{bmatrix}
  -k_{13} - k_{12} & k_{21} & k_{31} & 0 \\
  -k_{12} & -k_{21} & 0 & 0 \\
  k_{13} & 0 & -k_{31} - k_{30} & 0 \\
  0 & 0 & 0 & k_{30}
\end{bmatrix}
\begin{bmatrix}
  x_1(t) \\
  x_2(t) \\
  x_3(t) \\
  x_4(t)
\end{bmatrix}
\]

(7.2-1)

The conversion of this three compartment system into a series of sums of exponentials was performed by Nagashime et al (1968). For the central compartment, \( x_1 \), the sum of exponentials is

\[
\begin{bmatrix}
  x_1(t)
\end{bmatrix} =
\begin{bmatrix}
  \frac{(E_2-a)(E_3-a)}{(b-a)(c-a)} \exp(-at) \\
  \frac{(E_2-b)(E_3-b)}{(a-b)(c-b)} \exp(-bt) \\
  \frac{(E_2-a)(E_3-c)}{(a-c)(b-c)} \exp(-ct)
\end{bmatrix}
\]

(7.2-2)

The terms \( E_2, E_3, a, b, \) and \( c \) are all directly dependent on \( k_{12}, k_{21}, k_{31}, k_{13}, \) and \( k_{30} \). The relationships are shown by Wagner (1975). These allow the evaluation of the \( k \)'s from \( E_2, E_3, a, b, \) and \( c \).

7.3 The Determination of the Transfer Rates for a Three Compartment Model

The determination of the five parameters for
the model in Fig. 7.1 was performed using data in Appendix IV. The equations (7.2-1) and (7.2-2) were fitted individually and the results compared in Table 7.1.

The graphical display of these two results can be seen in Fig. 7.2 and Fig. 7.3. These displays and Table 7.1 show that there is no observable differences between the values assigned to the parameters by the model fitting procedure for equations (7.2-1) and (7.2-2). However, since equation (7.2-2) must be derived for every different model, resulting in re-coding of the computer programmes, it is recommended to use the differential equation model solver from chapter 2 whenever the explicit integration of (7.2-1) is not readily available.
Figure 7.2: Display of the simulated three compartment model using equations 7.2-2, the integrated form of the simultaneous differential equations 7.2-1.
Figure 7.3: Display of the simulated three compartment model using the simultaneous differential equations 7.2-1 directly using the numerical integrations programme of chapter 2.
Table 7.1: Comparison of results obtained from fitting two different functions representing the model shown in figure 7.1 consistent with the experimentally obtained data of appendix IV.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Initial Estimate</th>
<th>Using $A e^{-\lambda t}$</th>
<th>Using Equation 7.2-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_{12}$</td>
<td>$1.02 \times 10^{-2}$</td>
<td>$5.44 \times 10^{-2}$</td>
<td>$5.01 \times 10^{-2}$</td>
</tr>
<tr>
<td>$k_{21}$</td>
<td>$1.22 \times 10^{-2}$</td>
<td>$1.16 \times 10^{-5}$</td>
<td>$6.92 \times 10^{-6}$</td>
</tr>
<tr>
<td>$k_{13}$</td>
<td>$1.06 \times 10^{-1}$</td>
<td>$3.52 \times 10^{-2}$</td>
<td>$3.03 \times 10^{-2}$</td>
</tr>
<tr>
<td>$k_{31}$</td>
<td>$1.01 \times 10^{-1}$</td>
<td>$1.68 \times 10^{-0}$</td>
<td>$1.46 \times 10^{-1}$</td>
</tr>
<tr>
<td>$k_{30}$</td>
<td>$9.03 \times 10^{-3}$</td>
<td>$7.70 \times 10^{-3}$</td>
<td>$8.62 \times 10^{-3}$</td>
</tr>
<tr>
<td>Root Mean Squared Error (RMS)</td>
<td></td>
<td>.096</td>
<td>1.162</td>
</tr>
<tr>
<td>Number of Iterations</td>
<td></td>
<td>65</td>
<td>48</td>
</tr>
</tbody>
</table>
CHAPTER 8

DISCUSSION

Whenever an investigation is made in a new field a discussion of the failures and their causes is as important as the description of the successes. In the modelling and model fitting of biological systems various failures were encountered. These were of two types: systemic errors where the failure was caused by inadequacies of the computer programmes, and methodology errors where the failure was due to an inappropriate choice of model or an inaccurate choice of starting estimates.

The experiences of the author have shown that it was very difficult to identify which of the two categories caused a particular example to fail. This was of course most interesting when difficult examples were being studied and different groups of workers were responsible for the two different aspects. Much time was spent in finger pointing when a failure occurred. It was concluded that a standard operating procedure was required. The chart, figure 8.1, describes the sequence of steps
Figure 8.1: The standard operating procedure used for a computer simulation.
required for routine simulation of biological kinetic models. Figure 8.2 is the sequence followed for optimizing a set of parameters to best fit some experimental observations with some theoretical kinetic model. The charts do not represent the only sequence but serve to give direction for the investigator.

The model solver of chapter 2 was observed to never "fail" to integrate "non-stiff" models. The generation of the correct time course for an oscillating or stiff model required very small step sizes. The simulation was observed to fail due to too large a selected step size or, when a very small step size was used, too much computer time was required for the time period specified. These were frustrating errors, but not fatal to the simulation.

The success of the programme described for model fitting in chapter 3 was observed to be dependent on the difficulty of evaluating the objective function and the proximity of the initial estimates to the global optimum. The model fitter failed most frequently when the estimated parameter values were poor. To reduce the frequency of this type of failure another optimizing technique was employed to improve the starting estimates. The
rigorously define the mathematical and the kinetic model

from the model determine which of the observables are most important

perform an experimental protocol to frequently observe the important observables

develop a priori estimates of parameters consistent with observed data

test method of computing objective function

Figure 8.2: Standard operating procedure for use of model fitting programme.
Figure 8.2 (continued): Standard operating procedure for use of model fitting programme.
Fletcher-Powell algorithm (1963) proved to be most satisfactory. Another, less structured approach was to form a "grid" of initial estimates and then to initiate the optimizer from each point in turn. The grid approach is very similar to trial and error and was only used as a last resort.

When the objective function was difficult to evaluate, additional programmes were developed to improve the computational efficiency of the described method. This usually took the form of specific routines which would replace the general, but time consuming, algorithm. For example, for the enzyme kinetic models the calculation of higher derivatives were directly coded rather than being iteratively computed.

In conclusion then, it can be seen that this thesis has developed techniques which can be successfully used in the study of the underlying concepts of certain biological kinetic systems (chapters 4-7), but that care must be taken in certain difficult cases to obtain consistent results.
BIBLIOGRAPHY


Wagner, M., Erecinska, M. and Pring, M. 


APPENDIX I

The following is the FORTRAN IV listing of the computer programme for the solution of simultaneous differential equations using the Taylor Series expansion with a variable order of approximation with a variable step size developed in section 2.2.
C
DO 40 X=0.0, 6.0, 0.0
50 X(R)=0.0
40 CONTINUE
C
C 10 PASS=0.0
C
C 20 PASS=PASS + X(N)*V(N)*F(N)
C
C 30 PASS=PASS + X(N)*V(N+1)*F(N+1)
C
C 40 PASS=PASS + X(N+1)*V(N)*F(N)
C
C 50 PASS=PASS + X(N+1)*V(N+1)*F(N+1)
C
C 60 END
C
C DETECT IF A PACKET IS TOO OCTAH.
C IF PKT=1, THEN
ENDIF = (PKT=1)
I1 = (P=1)
I2 = (I1 = 1)
N = (N/2)
C PERFORM STEP #74 AND CALL TALIA'S SCRE.
C AT LEAST A FUNCTION PENCIL TO CONFERENCE IN FEBRY'S SITE.
C
C CONTINUE
IF (T=6000) STOP
IF (T=71000) STOP = CASE.
IF (T=2000) STOP = CASE.
X = (X = 1)
N = (N+1)
C IF THE LAST AND THEN OPERE BY MORE THAN 16 IN STOP, THEN
C CONTINUE NEXT TIME OF TABLE SEPER.
C
C STOP = CASE.
X = (X + 1)
C IF THE DERITIVE IS 0, THEN CALL DATA.
C IF THE ROOT HAS BEEN COMPLETED
C COMPUTE THE NORT CERTAINTY OF X(I) = X(I).
C THE FLAG WILL BE TRUE. THIS SIGNFIES THE COMPUTATION OF DATA.
C COMPUTE THE DEERIVATIVE FOR ORDER NO HAS BEEN COMPLETED
000 PRINT READ FROM X(I), I=1,7 TO X(N+1)
000 PRINT READ FROM X(N+2) TO X(N+7)
130 PRINT 150', H=100 LAGE, H=W+16, G=1 5. + I.N=15.90
200 PRINT RETURN
240 PRINT 140', L=NUMBER OF LINES ON TAPE=4 T112
260 PRINT 140', H=NUMBER OF LINES ON TAPE=4 T112
310 IF A<1.1 AND IN. 6E. 2116<IN-1
310 IF T1 ESS
330 DETERMINE INITIAL SIZE FOR NEXT ITERATION
350 XI=X(I+1)
370 DO 120 I=1, N
390 CALL PRINT INT X(I), READ X(N+I), I=1, 110, IN
400 CLICK XTMAIN, LAST TIME=LASTTIME+90+?110
116 LAST TIME=LAST TIME+90+?110
400 CALL PRINT INT X(I), READ X(N+I), I=1, 110, IN
400 IF I.LTIME.LASTTIME, I=1, 110, IN
APPENDIX II

The following is the FORTRAN IV listing of the computer programme for the location of the extrema of a mathematical function. The computer function EPSILON is the value to be minimized.
1.000 DIMENSION EPSILON(11),ALPHA(3),F(3),FP(4,6,6)
2.000 COMMON/DALE/HI(U(6),DATA(100),FDATA(100),HOBS
3.000 HAMELIST/HR(HH)
4.000 HAMELIST/IUPT/F.H,H
5.000 HI=2;HI=2.06;F(1)=1;F(2)=.5
6.000 PRINT 863
7.000 863 FORMAT('INPUT IUPT(F,H,H)')
8.000 F(1)=1400.;F(2)=60.
9.000 PRINT 864
10.000 864 FORMAT('INPUT HI')
11.000 HI(1)=3;HI(2)=8.
12.000 CALL RD
13.000 NR=H+1
14.000 C*** COMPUTE MODAL VALUES FOR INITIAL POINT
15.000 C
16.000 HACCOUNT = 0
17.000 CALL NODES(F,FP,HR,H)
18.000 L=1
19.000 HCOUNT=ACCOUNT+1
20.000 EPSILON(1)=EPSIL(F,H)
21.000 WRITE(2,171)ACCOUNT,F(1),F(2),EPSILON(1)
22.000 X 171 FORMAT(2X,19,3F16.9)
23.000 X 100 FORMAT('IF(H.LT.R-.05)GOTO 2')
24.000 IF(H.LT.R-.05)GOTO 2
25.000 100 FORMAT('X = AT POINT ('E12.6','E12.6') EPSILON = ','E16.8)
26.000 C*** COMPUTE FUNCTION VALUE AT MODAL POINT
27.000 C
28.000 C
29.000 DO 10 J=1,HR
30.000 DO 20 M=1,HR
31.000 IF(H.LT.J)GOTO 30
32.000 DO 40 K=1,H
33.000 F(K)=FP(K,J,H)
34.000 40 CONTINUE
35.000 L=L+1
36.000 EPSILON(L)=EPSIL(F,H)
37.000 PRINT 101,L,H,F(1),F(2)
38.000 101 FORMAT('NODE (''12'',''12'') POINT (''F10.5'',''F10.5'')')
39.000 X 102 FORMAT('GIVE EPSILON OF ''F16.8'')
40.000 X 102 EPSILON(L)
41.000 X 102 EPSILON(L)
42.000 X 102 EPSILON(L)
43.000 CONTINUE
44.000 CONTINUE
45.000 CONTINUE
46.000 C*** USE ROUTINE GAUSS TO SOLVE SIMULTANEOUS EQUATIONS
47.000 C
48.000 CALL GAUSS(H,EPSILON,ALPHA)
49.000 C
50.000 C*** USE Z(0,0) AND ALPHA TO FIND X(1,...,K(V+1)
51.000 C
52.000 CALL AT(1)(ALPHA,FP,H,F,H)
53.000 PRINT 303(F,K),K=1,H
54.000 300 FORMAT('NEW POINT AT ('E16.9','E16.9')/'H=','E16.9')
55.000 CALL NODES(F,FP,HR,H)
56.000 GOTO 1
57.000 2. PRINT 103
58.000 103 FORMAT('CONVERGED')
59.000 PRINT 301,F(K),K=1,H,H,EPSILON(1)
60.000 301 FORMAT('FINAL POINT ('E16.9','E16.9')/'H=','E16.9')
61.000 1 EPSILON AT STATIONARY POINT =''E16.9)
62.000 STOP 66
63.000 END
SUBROUTINE RD
COMMON/DALE/N(6),DATA(100),FDATA(100),NQBS
100 FORMAT(E10.5,1X,E10.3)
1 FORMAT(J*1:GOTO 1
99 NQBS=J-1
100 RETURN
END

FUNCTION EPSIL(F,H)
C *** FUNCTION TO BE MINIMIZED
C
DIMENSION F(3),FP(100)
COMMON/DALE/N(6),DATA(100),FDATA(100),NQBS
DO 205 I=1,NQBS
D=DATA(I)
FP(I)=0.
DO 205 J=1,H
FP(I)=FP(I)+FLOAT(H(1,J))*F(J)*D/(1.+F(J)*D)
FC=0.
DO 102 I=1,NQBS
FL=FL+(FDATA(I)-FP(I))**2
WRITE(1,101) FDATA(I),FP(I)
X 101 FORMAT(2E16.9)
102 CONTINUE
EPSIL=FC
100 RETURN
END

SUBROUTINE NOD(R,F,P,N,R,H)
DIMENSION F(3),FP(4,6)
RO(I,J)=PI*H(I,J)/MAXO(I,J)
C *** COMPUTE POINT FP(I,1,1) EQUV TO ZR(0,0)
QH=1./FLOAT(N)*H
HI=HR-1
C *** COMPUTE (H+1)NODES AS FP(I,J,K) EQUV TO ZR(J,K)
DU 10 J=1,NH
FP(J,1,1)=F(J)-QH
DO 20 I=2,NH
RT=RO(J,1-I)**H
DO 20 J=1,NH
FP(J,IX,I)=FP(J,1,1)+RT
20 CONTINUE
DO 10 J=1,NH
FP(J,IX,IX)=FP(J,1,1)+RT
10 CONTINUE
FP(J,IX,L)=.5*(FP(J,1,1)+FP(J,L,L))
CONTINUE
100 CONTINUE
100 RETURN
END
SUBROUTINE GAUSS(nn,FF,n)
DIMENSION A(3,3),BC(3),B(3),F(4,4),FF(11)

L=1
NR=nn+1

COPY EPSILONS INTO DOUBLE SUBSCRIPT F

DO 40 I=1,NR
DO 45 J=1,NR
IF(J.LT.I)GOTO 45
L=L+1
F(I,J)=FF(L)
45 CONTINUE
DO 40 CONTINUE

FILL COEFFICIENT MATRIX

DO 51 I=1,nn
DO 50 J=1,NR
A(I,J)=F(J,1+I)-F(1,J)
50 CONTINUE
IF(J.GT.I)A(I,J)=F(I+1,J)-F(1,J)
51 CONTINUE

BC(I)=B(I)=[(F(I+1,1+I)-F(1,1))/4.
D(NR)=BC(NR)=1
DO 55 J=1,NR
A(NR,J)=1
55 CONTINUE

SOLVE LINEAR EQUATIONS

CALL DSOLV(A,BC,NR,B,MM)

DO 100 MM=1,NR
BC(MM)=B(MM)

PRINT 129,B(MM),MM
100 FORMAT(1X,E16.9) = ALPHA('13')

CONTINUE

RETURN
END
SUBROUTINE DSOLV(X, B, H, V)
DIMENSION X(3,3), B(3), V(3)
DO 2 K=1,N-1
DO 1 I=K+1,N
CF=X(I,K)
DO 3 J=1,N
X(I,J)=X(K,J)*CF-X(I,J)*X(K,K)
3 CONTINUE
B(I)=B(K)*CF-B(I)*X(K,K)
2 CONTINUE
DO 10 K=N,1,-1
IF(K+1.GT.N)GOTO 22
DO 11 I=K+1,N
B(K)=B(K)-X(K,I)*V(I)
11 CONTINUE
10 CONTINUE
V(K)=B(K)/X(K,K)
CONTINUE
RETURN
END

SUBROUTINE ATOK(ALPHA,F,H)
C *** COMPUTES NEW POINT GIVEN THE RESULTING VALUES OF ALPHA
C *** AND THE Z(Q,0) COORDINATE
DIMENSION ALPHA(3), F(3), FP(4,6,6)
DATA XMIX/1E-10/
S=0
DO 10 J=1,N
IF(ALPHA(J+1).LT.0.OR.ALPHA(J+1).GT.1)GOTO 22
10 CONTINUE
H=N/2
DO 20 J=1,N
S=S-F(J)
20 CONTINUE
F(J)=H*ALPHA(J+1)+FP(J,1,1)
S=(S-S-F(J))**2+S
PRINT 104,FP(J,1,1),J
104 FORMAT('FP = ',E16.8, ' J = ', I3)
CONTINUE
IF(S.LE.XMIN)H=1E-75
RETURN
END
APPENDIX II.2

The following is the XEROX APL workspace for the determination of extrema and is equivalent to the FORTRAN IV programme in appendix II,1.
VLEASS(t)
V R=LEASS, FD

[1] 'ENTER NUMBER OF PARAMETER :
[2] N=1:'
[3] 'ENTER INITIAL VALUE FOR EACH PARAMETER :
[4] U=1:
[5] NN=1:
[6] 'NUMBER OF DATA READ' = N=RDYTA FD; I+1:
[8] +(U=1E-15)/END;EP=1+EP;'POINT ':E=
  EPSILON = EPS[1];EP+PPP EPS FD
[10] END;'CONVERGED ';FINAH POINT ,E'; H = ;U

V

VRDUTA(t)
V R=RDUTA FD; I; A; C; M

[1] I=0;m=L;K;1:OPEN IN FILE FD DCB 1; A+FDA DDA+;0
[2] XX1: C=REDN 1,(I+1),100
[3] XX2: XX2+C(M)+I-1;+(P<C)<M-C;I-1)/XX3
[4] XX3: SAVE 1+I<R,XX1+DDA+DDA,A(2); FDA+FDA,A(1); A+;C

V

VWU(t)
V H=U UD W;P; I; J; L

[1] R=(U*(W)*,+1+P)+((P++/(N+1)),U)+E-U(N+1);L+3+J-Q
[2] SS1:=(J>U)/0;I+J+1
[3] SS2: SS1;+(I+N)/J+2;K[L]+((K;J)+K[I])+(2;I+I+1;L+L+1

V

VEPS(t)
V H=NN EPS PP; S

[1] S+=(I/2);DDA*; FP*Q(RFP);ppNN;1+DDA*; FP
[2] R++/(P(S);R(S); FDA)+2
APPENDIX III

The documentation and FORTRAN IV computer listing for a smooth curve to be implemented with the TEKTRONIX 4010 graphics terminal.
The process of modelling and model solving generates large amounts of numerical data. The proper evaluation of the data requires graphical as well as statistical displays of the final results. To create the graphical displays a computer programme was written to produce smooth curve, multi-function graphs for interactive use with a TEKTRONIX 4010 storage scope. In order to integrate the display programme into a laboratory environment it was implemented as a fully conversational and simple to use computer programme package with the flexibility of modifying display parameters.

The computer programme CALPLOT handles the user/computer interface and then generates the appropriate calls to six routines contained in the TEKTRONIX Advanced Graphics Package II. The user has the ability to interactively modify one, or more of the following display parameters: maximum time, minimum time, maximum concentration, minimum concentration, the number of decimals for the labels and number of labels per axis. The ability to select these parameters means the user may obtain a "Blow-up" of any region of the display.
PROGRAM CALLPLT(INPUT, OUTPUT, TAPE1, TAPE2, TAPE3, TAPE4=OUTPUT)
REAL X(1001), Y(1001, 8)
INTEGER FORM(3)

1 FORMAT(* ENTER INPUT FORMAT*)
READ 2, FORM
2 FORMAT(3A10)
PRINT 3
3 FORMAT(* ENTER NUMBER OF LINES AND Y COL. TO BE READ *)
READ 4, N, M
IF(N.GT.1000, OR. M.GT.8) STOP77.
PRINT 6
6 FORMAT(* UNIT NUMBER FILE IS ON, *)
READ 4, INTAPE
FORM(14)
CALL SUBPLT(N, M, FORM, X, Y, INTAPE)
REMNDE* 1 REMIND 2* REMIND 3
REMNDE 4
IF(INTAPE.EQ.4) GOTO05
STOP1
END

SUBROUTINE SUBPLT(N, M, FORM, X, Y, IN)
REAL X(N), Y(N, M), Y(1:2, 2)
INTEGER IN, FORM(3)
LOGICAL A
DATA XINT, YINT/2-10, XMIN, YMIN/2-0.5, XDEC/2.5, YDEC/7-
NAMELIST /CHANGE/ R, N, XMAX, YMAX, XMIN, YMIN, KDEC, YDEC
Y(1, 1)=1, Y(1, 2)=E48

C +++ SET UP PARAMETER VECTOR
C
Y(1, 1)=M(1:2)=Y(1:2, 2)=1
R(1, 1)=XMIN, R(1, 2)=XMAX
R(1, 1)=YMIN, R(1, 2)=YMAX
R(1, 1)=XINT, R(1, 2)=YINT
R(1, 1)=IN, R(1, 2)=FORM
R(1, 1)=1, R(1, 2)=INTAPE
REMNDE 1 REMIND 2* REMIND 3
REMNDE 4
IF(INTAPE.EQ.4) GOTO05
STOP1
END
V(2,1) = N
V(4,1) = XMAX
V(7,1) = XINT
V(7,2) = YINT
V(4,2) = YMAX
V(3,1) = XMIN
V(3,2) = YMIN
V(12,1) = XDEC
V(12,2) = YDEC
PRINT CHANGE
IF(Y(1,1), EQ., -1, E+240) READ(IN, FORM) (X(I), Y(K, I), I = 1, M), K = 1, N
V(8,1) = V(1,1) = V(8,2) = V(10,1) = V(10,2) = 1

C *** START PLOTS

DO 30 L = 1, M
IF(A) GOTO 35
A = T.
PRINT 10, ((V(I, KK), KK = 1, 2), I = 1, 12).
FORMAT(2F5.0)
REIND 4
CALL GRAPH(X, Y(I, L), V)
GOTO 30

50 V(8,1) = V(8,2) = V(10,1) = V(10,2) = V(1,1) = 0.
REIND 4
CALL GRAPH(X, Y(I, L), V)
CONTINUE
CALL DUMP
CALL BELL
CALL INITT(5, 10)
PRINT 20

FORMAT(1) DO YOU WISH TO ALTER PARAMETERS TRUE/FALSE
READ 20, A

FORMAT(1) DO YOU WISH TO CHANGE DATA FILES TRUE/FALSE
READ 20, A
END
APPENDIX IV

Theoretical data used to demonstrate the model fitting procedure.