INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the original text directly from the copy submitted. Thus, some dissertation copies are in typewriter face, while others may be from a computer printer.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyrighted material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is available as one exposure on a standard 35 mm slide or as a $17'' \times 23''$ black and white photographic print for an additional charge.

Photographs included in the original manuscript have been reproduced xerographically in this copy. 35 mm slides or $6'' \times 9''$ black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Order Number 8726513

Numerical analysis as applied to x-ray scattering curves

Jeffries, Thomas Vincent, D.A.

Middle Tennessee State University, 1987

U·M·I

300 N. Zeeb Rd.
Ann Arbor, MI 48106

NUMERICAL ANALYSIS

AS APPLIED TO

X-RAY SCATTERING CURVES

Thomas V. Jeffries

A Dissertation presented to the
Graduate Faculty of Middle Tennessee State University
in partial fulfillment of the requirements
for the degree Doctor of Arts

August 8, 1987

NUMERICAL ANALYSIS

AS APPLIED TO

X-RAY SCATTERING CURVES

APPROVED:

Graduate Committee:

Martin U. Stewart
_____
Major Professor

Sam R. Wulfberg
_____
Committee Member

_____
Committee Member

Dan W. Scott
_____
Chairman of the Department of Chemistry and Physics

_____
Dean of the Graduate School

ABSTRACT

NUMERICAL ANALYSIS AS APPLIED TO X-RAY SCATTERING CURVES --

Thomas V. Jeffries

Various numerical methods involving polynomials were
employed for both interpolation and linear least-squares
curve fitting of the atomic scattering factors of X-ray
diffraction. This included use of Lagrangian and orthogonal
Legendre polynomials, as well as cubic spline and Stineman
interpolation.

Interpolation is a method that uniquely matches known
data points within small segments of an unknown curve to
approximate points in between. Special emphasis was placed
on establishing both the minimum grid spacing and polynomial
degree that are needed to perform accurate interpolations.
The grid spacing in a region of the x-argument near 1.0 $\overset{\circ}{A}{}^{-1}$
that is commonly employed in standard tabulations of X-ray
scattering factors was found to be too coarse for wholly
accurate interpolation by polynomials of low degree.

Least-squares curve fitting is a procedure that
approximates the entire unknown curve with an analytical
function that matches the known data points as closely as
possible by minimizing the sum of their squared deviations
from the fitted curve. Analytical representations of X-ray
scattering curves are advantageous, since otherwise the
complete scattering table must be stored and the values of

Thomas V. Jeffries

individual scattering factors derived by interpolation. The literature of previous nonlinear least-squares fits of Gaussian expansions and polynomial series to approximate X-ray scattering curves is reviewed exhaustively. Several transformations were tested for linear least-squares curve fitting of X-ray scattering factors. It was concluded that the best approximation is obtained from a combination of Gaussian and polynomial terms.

A special chapter on the historical impact of numerical analysis by computers in modern chemical education is also included.

ACKNOWLEDGEMENTS

I would like to thank Mr. Harlie White, Dr. Franklin Cheatham, and Dr. Robert Street of Campbellsville College for their assistance with computer programming and word processing. Thanks to Dr. Dan D. Scott, Dr. Jim C. Howard, Dr. Roy W. Clark, Dr. Linda A. Wilson, and other members of the faculty at MTSU for their assistance and encouragement. Much credit goes to Gerry D. Thomas and Zena A. Tenenbaum for their many hours of input on this project.

I am deeply indebted to the chairman of my committee, my mentor and friend, Dr. Martin V. Stewart. I wish to express a special thanks for his assistance, patience, and the long hours spent on this project. To the other members of my committee, Dr. Gary P. Wulfsberg and Dr. Jack Arters, I would like to express my thanks for their help.

I also wish to express my appreciation to the Graduate School of Middle Tennessee State University and to the Education Commission of the Southern Baptist Convention for their financial support. Dr. Robert S. Clark, Dean and Academic Vice President of Campbellsville College, has been most encouraging.

I am deeply appreciative of my wife, Wanda, and of my daughters, Camille and Alyssa. Their love, support, and sacrifices have made the duration of this project possible.

TABLE OF CONTENTS

LIST OF TABLES

iv

v

LIST OF FIGURES

vi

LIST OF APPENDICES

vii

CHAPTER 1

Introduction

The first chemical concept presented to beginning students is usually an empirical definition of matter as anything displaying the fundamental quantities of mass and volume, whose ratio is the derived physical property of density for a particular sample. Subsequently, the three states of matter--the gas, liquid, and solid phases--are classified by a similiar empirical approach in terms of the observable qualities of volume and shape (1). A solid is described as a state of matter having a definite volume that, as in the liquid state, is only slightly altered by compression and also having a definite shape that, unlike both gases and liquids, is maintained even in the presence of large elastic forces. Later, students are taught that the term solid is strictly reserved for substances having a crystal structure and, still later, that advanced study of the properties of solid materials is largely a study of the crystalline state of matter (2-7).

The difference between a noncrystalline or amorphous substance and a crystalline one is the same as a comparison between a kindergarten classroom and a military academy. The chemical species, whether they be molecules or ions, of a noncrystalline substance are disordered, being arranged in

1

a chaotic, hit-or-miss fashion; whereas in a crystal, the arrangement forms a symmetrically repeating pattern in three-dimensional space having the lowest potential energy that conforms to the laws of close packing (8). This pattern is extended indefinitely throughout the crystal.

The essential symmetry requirement for crystal packing establishes a space lattice, which is a three-dimensional grid system of points occupying an identical environment within the crystal. There are 14 such Bravais space lattices that, symmetrywise, can be classified into one of seven crystal systems: triclinic, monoclinic, orthorhombic, tetragonal, rhombohedral, hexagonal, and cubic. These crystal systems describe a parallelpiped-shaped volume called the unit cell of the crystal. Repetitious unit translation of this unit cell in three dimensions affords the entire crystal structure because it possesses a complete representation of the symmetrical pattern. Thus, a crystal is described as being made up of parallel planes of atoms, spaced regularly in numerous directions, that are repeated periodically by lattice translations. The total number of intermolecular arrangements capable of producing such patterns of perfect periodicity in three dimensions is limited to 230 unique combinations of symmetry operations called space groups (9-13). They are authoritatively summarized for X-ray crystallographers in an international publication (14).

When a narrow and nearly monochromatic beam of parallel X-rays fall on a crystal, a portion is scattered in certain definite directions with various relative intensities, a phenomenon which is called X-ray diffraction (15-20). The positions of these so-called Laue diffraction maxima are related to the dimensions of the space lattice of the crystal. Their relative intensities are determined by the symmetry of the space group, by the three-dimensional arrangement and thermal vibrations of atoms within the unit cell, and by the different capacity of individual atoms to scatter X-rays. The subsequent development of X-ray crystallography permitted structural determination of the individual repeating unit and, hence, provides an extremely powerful technique for establishing the chemical structure of both organic and inorganic compounds (21-27). Although the effect of single-crystal X-ray diffraction in three dimensions is analogous to the one-dimensional diffraction of visible light by a ruled optical grating (28, 29), the actual process involves absorption of the unidirectional incident radiation by the electrons of atoms and subsequent re-emission in all directions under conditions where coherent Thomson scattering greatly predominates over either incoherent Compton scattering or X-ray fluorescence (30-32).

William Lawrence Bragg, in 1913, showed that the X-ray diffraction pattern from crystals is more easily interpreted if treated mathematically as the result of the interaction

of X-rays with a reflection grating rather than with a diffraction grating (33). To understand this approach, the previously defined concept of the crystallographic space lattice is needed. An infinite number of families of parallel and equidistant geometrical planes may be drawn through the lattice points of a crystal in various periodic ways. Bragg noted, as a theoretical model, that the diffracted X-ray beam acts as if it were being reflected from these sheets of imaginary lattice planes following the ordinary law of mirror reflection: the angle of incidence is equal to the angle of reflection, and both occur in the same plane as the perpendicular to each set of mutually parallel reflecting planes of the crystal lattice.

Bragg theorized, then confirmed by experiment, that such "reflected" X-rays must exhibit mutual interference. Constructive interference can occur only when the incident X-ray beam makes a particular angle, θ, with a successive set of parallel planes of the crystal lattice having interplanar spacing d, for only at this angle are they mutually in phase. For a given family of crystallographic planes, destructive interference predominates at all other angles. Consequently, single-crystal X-ray diffraction patterns appear as a symmetrical distribution of spots on photographic film, one spot for each family of planes. Such diffraction spots are commonly referred to as reflections, the etymology of which is self-evident.

The strict condition for reinforcement is one where differences in path length between waves scattered from successive planes, $2d(\sin \theta)$, are an integral number of wavelengths of the radiation employed. This is summarized by the familiar Bragg equation as $n\lambda = 2d(\sin \theta)$ where n is an integer called the diffraction order, λ is the wavelength of the monochromatic X-rays, d is the perpendicular spacing between lattice planes in the crystal, and the glancing or Bragg angle, θ, is the complement of the angle of incidence for the penetrating X-ray beam satisfying the essential diffraction condition (33).

The contribution of a particular atom to the overall intensity of a diffracted X-ray beam is described by a set of dimensionless quantities called atomic scattering factors, $f(s)$, that are a function of the type of atom and both the Bragg angle and the wavelength of the radiation in the combination of $s = (\sin \theta)/\lambda$. They are necessary to account for the fact that the electrons, which are the scattering matter, are situated in a volume around atomic nuclei rather than being concentrated at discrete points. The amplitude of X-rays scattered in a particular direction by a free, single electron, $A(s)_e$, is well-known (29, 34), and the scattering amplitude due to electrons in an atom, $A(s)_{atom}$, is merely $f(s)$ times that due to an isolated electron under similiar conditions as expressed by the ratio $f(s) = A(s)_{atom}/A(s)_e$, which is sometimes referred to as a

form factor or atomic scattering power. The dependence of f(s) on the atom type is a direct result of their different electron densities. The variation with s is because the constituent waves scattered by each electron partially interfere, which causes a decrease in the amplitude of the resultant wave scattered by the whole atom as a function of the Bragg angle.

Atomic scattering factors are essential parameters employed during the solution and refinement calculations involved in a single-crystal X-ray diffraction analysis. Each atomic species has a unique set of X-ray scattering factors, which collectively are called a scattering curve when the scattering factors are plotted as a function of the crystallographic Bragg angle and are referred to as a scattering table when the numerical values of the individual data points are listed. The evaluation of such coherent X-ray scattering factors, although well-known in practice, is a complicated quantum mechanical calculation (31, 35-38). They are obtained from the theoretical wave functions of an atomic species through numerical integration of the radial distribution of the electron density over the volume of its electron cloud, the atomic center being taken as the origin. James and Brindley (29, 39) were the first to employ the Hartree self-consistent field model to calculate X-ray scattering curves, although some of their early scattering tables were actually obtained through interpolation of the

scattering factors of several different elements and were later shown to be inaccurate (40-42).

X-ray scattering tables are now available from the more refined Hartree-Fock wave functions (36, 43), and the most reliable of these are collected together for all elements and a large number of elemental ions in International Tables (44, 45). The Hartree solution was found unsatisfactory for heavy atoms, but relativistic Hartree-Fock wave functions (46-48) can now be used to replace both the Dirac-Slater approximation (49, 50) and the earlier Thomas-Fermi-Dirac statistical model (51).

These solutions are strictly valid only within three fundamental assumptions inherent to the atomic model that X-ray crystallographers employ by necessity, if not by choice, in their calculations. First, they assume that the radial dependence of electron density can be modeled after an isolated atom or ion that is spherically symmetrical about the nucleus with respect to both its valence-shell and core electrons. Second, they assume that the scattering amplitude of each volume element of the electron cloud during integration is proportional to that of a free, single electron--the previous $A(s)_e$ term--whose absolute scattering amplitude is taken as the unit to normalize the total electron charge to the formal number of electrons of the atomic species. Finally, they assume that the atom is

at rest, with f(s) corresponding to the scattering power of a stationary atom.

The latter two assumptions can be accounted for during refinement of an X-ray crystal-structure solution through use of anomalous dispersion corrections and atomic thermal parameters, respectively, together with various other corrections that depend upon the experimental conditions. The first assumption, that of spherical symmetry for the charge distribution, cannot be so easily corrected and is violated by incomplete electron shells and the aspherical charge density of bound atoms (5, 42, 52, 53). The problem here is that the exact electron-density distribution of an atom depends upon the atom type, its bonding state, and its volume in ways that are unknown and are only grossly approximated by various levels of theory. Fortunately, the anisotropic temperature factors (54, 55) absorb much of this error in the practical use of X-ray scattering factors during the least-squares refinement of a crystal-structure solution. Furthermore, Weiss (35, 46, 82) has shown that the numerical values for the calculated scattering factors are not very sensitive to inaccuracies in the electron densities from which they were derived and that different theories give quite similiar results. He estimates that the theoretical free-atom Hartree-Fock scattering factors are accurate to within about 1%.

Figure 1 shows the X-ray scattering curves of neutral hydrogen, nitrogen, and bromine atoms together with those of the iron(II) and cesium(I) cations. Here, the scattering factors, f(s), are plotted against s = (sin θ)/λ. The shape of these plots is roughly the same for all atoms and ions, with the greatest differences in magnitude occurring at low values of s where the valence electrons chiefly contribute to X-ray diffraction phenomena.

The magnitude of f(s) for discrete atoms and ions at any given s is accurately proportional to their electron density and only roughly proportional to their formal number of electrons, an effect readily noticed when the scattering curves of isoelectronic species are compared (37). The numerical value for f(s) of a neutral atom is equal to its atomic number, Z, in the special case of the unique s = 0 intercept, where the curves reach their maximum. At this point, the resultant diffraction amplitude from the whole atom is proportional to the total number of electrons in the incident beam direction because the X-rays scattered by any particular electron have the same path length and are, therefore, in phase. The value of f(0) is increased or decreased from Z by the magnitude of charge for an ionized atom. As the Bragg angle increases, f(s) rapidly decreases due to the destructive interference of increasing phase differences of X-rays diffracted from different parts of the atom. Thus, a set of X-ray scattering factors for a given
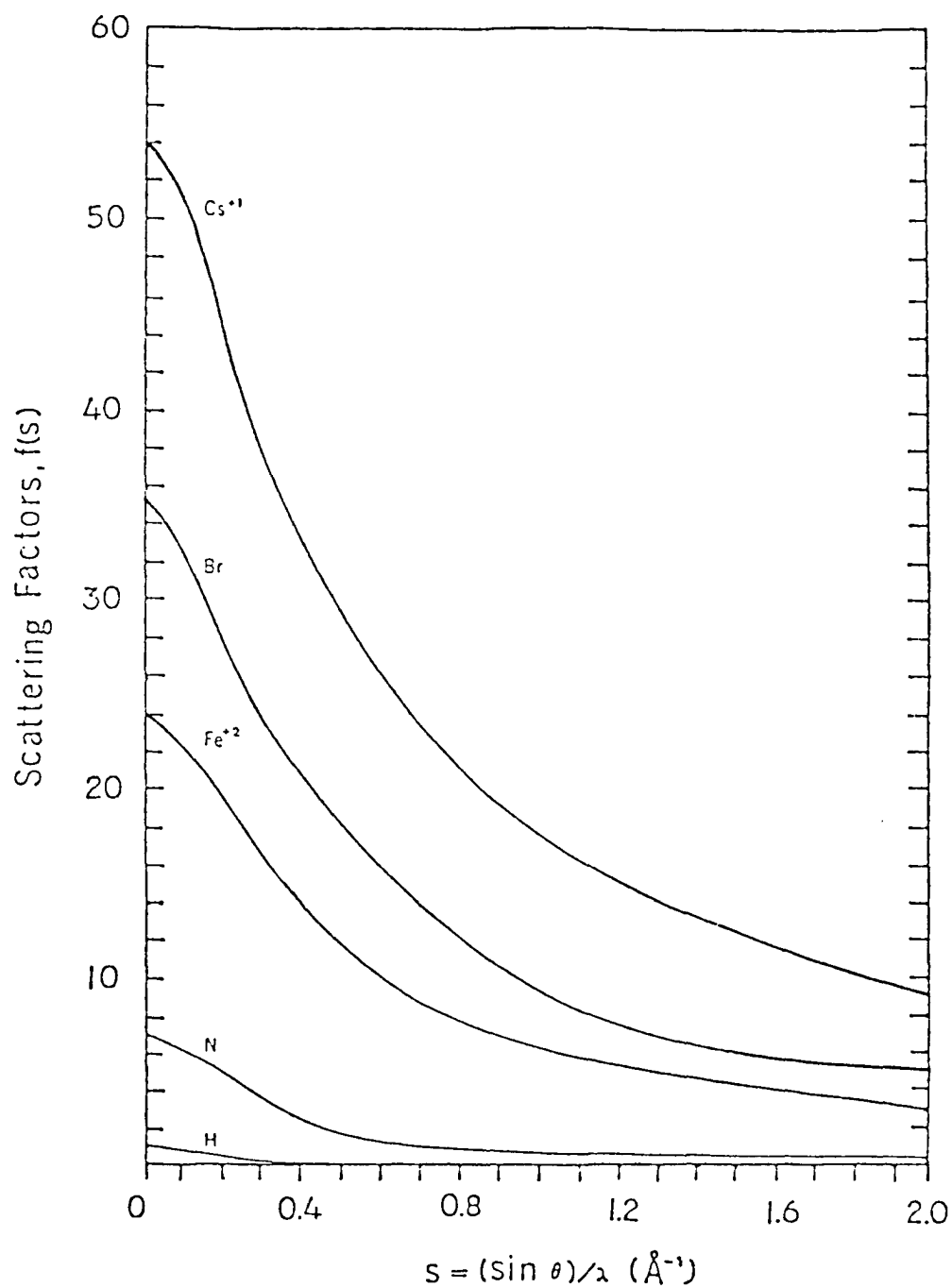
Figure 1

Atomic Scattering Curves for
H, N, Fe(II), Br, and Cs(I)

atom or ion represent the effective number of electrons at any diffraction angle.

The result is an S-shaped curve with no points of inflection that rapidly decreases as s increases, then asymptotes to a known function (36, 46, 56). This shape is similiar to plots resulting from a combination of two curves, one which shows an exponential rate of decrease dominating at low values of s and the other which shows an exponential rate of increase dominating at high scattering angles. The intermediate values of f(s) are nearly linear with s and lie between two regions of opposite curvature and rapidly changing slope.

The storage of entire scattering tables for all atoms and ions at sufficiently small intervals of s to insure completely accurate numerical interpolation of f(s) during crystal-structure solution is, of course, possible for a high-speed computer with a large memory (57-61). The limited memories of early computers, however, encouraged a search for mathematical functions that could serve as accurate, analytical representations for individual X-ray scattering curves in condensed form (62-64). Although the steady improvement of main-frame computers has overcome such difficulties, the possibility of doing crystallographic calculations on personal computers (65) and the present needs of protein crystallographers (25, 66, 67) has returned the problem of storage to its former significance. One

serious disadvantage of an analytical approximation, however, is that it is expressed as a set of coefficients for a function. Any typographical errors in the publication of these values will cause the entire X-ray scattering curve, not just a single scattering factor, to be erroneous. This worry is not mere paranoia, for decimal errors in two different coefficients of separately fitted scattering curves have already been published (68).

It is the purpose of the present dissertation to apply the numerical-analysis techniques of both interpolation and curve fitting to X-ray scattering curves. Much theory and many computer programs are available for fitting curves to data (69-80); however, the proper technique comes from a careful and deliberate choice. The hasty selection of an inappropriate procedure that merely "connects the dots" can be misleading and may ultimately misrepresent the actual curve.

Several possible functions were considered in the early literature for use as an approximation of X-ray scattering curves; however, the one most frequently employed is the nonlinear least-squares fitting of a Gaussian expansion (69) having the form:

$$f(x) = c + \sum_{i=1}^{k} a_i [\exp(-b_i x^2)] \qquad \text{Equation 1.1}$$

where k is the number of Gaussian terms and c is a constant term. The scattering curve is represented when the set of a_i, b_i, and c coefficients are obtained from least-squares fitting of the scattering factors, f(x), taken as a variable dependent upon an expression, x, which is related to the crystallographic Bragg angle. Such a Gaussian function is already contained as a subprogram in standard software for other computations in X-ray crystallography, and it has the advantage of very rapid convergence. It also has a simple Fourier transform (81) that allows calculation of electron density using parameters proportional to the coefficients of the fitted X-ray scattering curve (35, 82-85).

Costain (86) and Booth (87) were among the first to suggest and apply a symmetric two-parameter (k = 1, c = Ø) Gaussian function to approximate the distribution of atomic electron densities, which has found more recent application in estimating the separate contribution of core electrons to X-ray scattering factors (88-9Ø). Increasingly better fits are obtained when additional Gaussian terms are added to form an expansion, but each new term requires that the values of two additional parameters (i.e., the a_i and b_i coefficients) be evaluated. Witte and Wölfel (91) used the four-parameter (k = 2, c = Ø) expansion to approximate the X-ray scattering factors of NaCl. Vand, Eiland, and Pepinsky (62) employed it to systematically calculate each set of four coefficients needed to fit early tabulations of

X-ray scattering factors. Successive cycles of nonlinear least-squares approximation of f(s) showed errors rarely exceeding $\pm 1\%$ over the limited $0.00 \leq s \leq 0.65$ range of copper radiation, but poor fits occurred for the much larger $0.00 \leq s \leq 1.40$ range accessible by the short wavelength of molybdenum radiation. These calculations were facilitated for the computational machines of the 1950's by imposing a least-squares fitting condition that $a_1 + a_2 = N$, where N is the number of electrons in the neutral atom (N = Z) or ion (N = Z $\pm$ charge).

They suggested that improved fits over larger ranges of s would be obtained by inclusion of a constant term and the condition $a_1 + a_2 + c = N$, but also noted that subsequent least-squares calculation of the five coefficients would cease to be straightforward and would involve a laborious successive approximation procedure. The inclusion of the constant, c, rather than a third Gaussian term is not a mathematical requirement, but was based solely on the fact that this is the simplest possible extention of the equation for the limited computing facilities then available. It has, however, an important effect on high values of s, for the expanded Gaussian tends to asymptote to the numerical magnitude of the constant term as the limit. It should also be noted that these functions and their coefficients are only numerical fits and have no theoretical basis or significance.

The five-parameter (k = 2) Gaussian expansion, with some computational modifications, was applied by Forsyth and Wells (92) and later by Moore (93) using three cycles of least-squares reiteration to fit the best scattering factors of most atoms and many ions through the publication of Volume 3 of International Tables (45). The use of a five-parameter (k = 2) expansion for the X-ray scattering factors of O^{2-} (94) has been compared critically with the results obtained from a nine-parameter (k = 4) expansion (95), the latter also being employed to fit the atomic scattering curve of bonded hydrogen in H_2 (36, 44, 96) and the molecular scattering factors of H_2O (97). The Gaussian expansion of Equation 1.1 has also been applied as an analytic approximation for incoherent X-ray scattering factors (97, 98), for the scattering factors of electron diffraction (47, 99, 100), and for the scattering factors of the core and valence electrons independently as three- and four-term Gaussian expansions (52, 88, 101). Representation of scattering factors with one, two, or three Gaussian terms finds special application in protein crystallography (25), a device that allows ready inclusion of isotropic temperature factors (66, 67).

Current practice is reflected by International Tables, Volume 4 (44), which employs a nine-parameter (k = 4) Gaussian expansion. Use of the reciprocal space variable as $x = s = (\sin \theta)/\lambda$ allows the Gaussian coefficients of the

least-squares fit to be independent of the wavelength of the radiation and is, therefore, preferred over the original use of $x = \sin \theta$ that was employed in the early literature. Least-squares fitting is now carried out as the minimum sum of squares for nonlinear functions by cycles of successive approximation (i.e., reiteration), which is no longer a tedious operation with a modern computer. Thus, the original $c + \Sigma a_i = N$ condition has been relaxed resulting in an improved fit over a more useful range of s. A weighting factor having the form $w(s) = \exp[-(s - \bar{s})^2]$ is generally employed during the least-squares refinement to insure best agreement of the fitted curve with the tabulated data near the "middle" of most X-ray scattering curves and corresponds to a Bragg angle where most X-ray data are measured for crystal-structure determination. The value of $\bar{s}$ is often taken as 0.5 and 1.0 for copper and molybdenum radiation, respectively.

Table 1 presents the average errors reported in the literature for least-squares fitting of the indicated number of X-ray scattering curves over the range $0.00 \leq s \leq s_{max}$ for the x-argument as a Gaussian expansion of Equation 1.1 having various numbers of terms. The average error is given as E, which is the standard deviation of the fit expressed as a percentage of the number of electrons associated with individual atoms or ions as defined by Equation 2.3 in Chapter 2. The results obtained by these different workers

Table 1

Summary of Gaussian Expansion Fitting Errors

Number of Parameters in Gaussian Expansion	Number of $f(s)$ Curves Averaged	s_{max}, Å^{-1}	$w(s)^a$	Mean E, %	Ref.
2	4	0.34	C	± 0.8	103
2	2	1.50	unit	± 4.6	35
3	2	1.50	unit	± 1.1	35
4	10	0.34	C	± 0.1	103
4	29	0.70	unit	± 0.52	62
5	258	0.70	A	± 0.102	93
5	100	0.65	A	± 0.166	102
5	94	1.30	A	± 0.576	92
5	210	1.30	B	± 0.429	93
5	100	1.40	B	± 0.459	102
9	36	1.20	A	± 0.063	97
9	208	2.00	A	± 0.015	49
9	54	2.00	unit	± 0.083	47

[a]Weighting schemes are denoted as: A for $w(s) = \exp[-(s - 0.5)^2]$, B for $w(s) = \exp[-(s - 1.0)^2]$, C for $w(s) = s^2$, and unit for $w(s) = 1$.

are not strictly comparable because the details of their
computational methods--such as the weighting scheme, the
data grid spacing, the method of least-squares refinement,
and the accuracy of the original sets of scattering
factors--are not exactly the same. Nevertheless, the
general trends of E decreasing as more Gaussian terms are
added and E increasing as the range of s_{max} is extended
can be seen clearly.

The crystallographic literature involving polynomial
approximation, when compared with the similar use of the
Gaussian expansion, is relatively sparse, diffuse, and
obscure. The first application of a polynomial series to
represent X-ray scattering factors was by Mayer (64, 104),
but was rejected by him as being "not significantly faster
then table-searching and...appreciably less convenient
unless the problem of storage room is critical." This was
reinvestigated by Freeman and Smith (63), who applied a
nested form (105, 106) of Equation 2.1 (Chapter 2, p 28) for
a nonlinear least-squares fitting of the X-ray scattering
tables for 37 atoms and ions. An average fitting error of
E = $\pm 0.17\%$ was obtained for the $0.00 \leq s \leq 0.70$ range using
a sixth-degree polynomial series and, generally, 0.05
intervals of s. A glance at Table 1 shows this to
correspond approximately to the typical performance of a
five-parameter Gaussian expansion. They reported that no
improvement was obtained when polynomials of higher degree

were employed and found that inclusion of **s** > 0.7 caused the function to diverge wildly.

Gillis (107) claimed pessimistically that the error of a nonlinear least-squares fitting of a ninth-degree polynomial is less than 1% for $0.00 \leq s \leq 0.8$ and less than 5% for $0.0 \leq s \leq 1.3$ when applied to the carbon scattering table. Silverman and Simonsen (108) attempted to extend this range by fitting $\ln[f(s)]$ against s as a more nearly linear function and obtained $E = \pm 0.33\%$ for $0.00 \leq s \leq 0.70$ and $E = \pm 0.90\%$ for $0.00 \leq s \leq 1.30$ using a sixth-degree polynomial for eleven fitted curves. Onken and Fisher (109) employed this transformation in a nonlinear least-squares fitting of a sixth-degree polynomial for all X-ray scattering tables listed in Volume 3 of International Tables (45). Silverman and Simonsen (108) obtained exact fits for $0.00 \leq s \leq 1.30$ when scattering curves were subdivided into four to six successive ranges containing six tabulated points and each was separately fitted by a fifth-degree polynomial, but now four to six sets of coefficients would be needed to cover this range and the least-squares problem begins to approach one of interpolation. Polynomials have been used to fit electron scattering curves (110, 111) and have also been applied in both Lorentzian (47, 65) and orthogonal (112) forms for X-ray scattering factors over the limited range of copper radiation.

Polynomials have always played a central role in mathematical approximation (69-80). They can wiggle in a number of satisfactory ways to fit a variety of curves. No polynomial, however, ever has a true asymptote. Thus, a polynomial series should be grossly inefficient in fitting $f(s)$ for both high and low values of s. A Gaussian expansion (69, 70, 80, 113), on the other hand, is essentially an asymptotic series and should improve in accuracy as s increases.

It is obvious from this review of the literature that X-ray crystallographers have generally accepted expanded Gaussian functions for approximation of X-ray scattering curves. Experts in the field of numerical analysis, however, do not rely on such exponential fitting methods and even caution against their use. An exponential series of this type can be extremely ill-conditioned because it fits data as differences of very large numbers. Consequently, the use of exponentials tend to limit significant figures as an inherent feature of the function in even the best cases and produce instability in the worst. Another disadvantage of least-squares fitting of an expanded Gaussian function is that it really is not applicable to memory-limited personal computers (65, 114), being a nonlinear summation of exponentials.

For these reasons, the best approximation of an X-ray scattering curve would probably employ a polynomial times a

suitable Gaussian term so that the best features of both functions could be exploited and the worst minimized. This idea is not original, but is standard practice in the field of numerical analysis for dealing with problems that are exponential in character (115). The recent incorporation of Jacobi polynomials (52) into the previous expanded Gaussian approximation (101) for the generalized valence-shell X-ray scattering factors of R. F. Stewart (54, 83, 84) is an example of this approach. Indeed, its widespread use should delegate the methods both cited and applied herein as belonging to the classical period of X-ray scattering factor approximation.

The present dissertation examines the behavior of polynomials in fitting these curves. There is more concern here with understanding this behavior than in artificially obtaining "good fits." Consequently, polynomial series will be applied in unadulterated form as linear least-squares fits using unit weights. No pretense is made here to have solved this problem, but the basis for its solution has been defined in the practical terms of numerical analysis.

CHAPTER 2

Computational Methods

The results of our application of various methods of numerical analysis (69-80) to the distinctly different problems of both interpolating and curve fitting the atomic scattering factors of X-ray diffraction are discussed in Chapter 3. This was accomplished by several computer programs referred to as Curve Fitter, Curve Fitter-PC, Polft, and Lagi. Curve Fitter and Curve Fitter-PC are multi-purpose programs with the capabilities of polynomial, cubic spline, and Stineman interpolation as well as curve fitting by linear least squares. Polft is a specialized program that only performs a linear least-squares fit with a classical Legendre orthogonal polynomial series. Lagi is also a specialized program capable of interpolation by the Lagrange polynomial method. These methods of numerical analysis are all based on operations involving polynomial functions, the details of which are discussed below.

The general algebraic polynomial (116) is given as

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_i x^n$$

where $f(x) = y$ is the dependent variable, the x-argument is the independent variable, the n is some nonnegative

22

integer called the degree or order of the polynomial, and the set of a_i (where $i = 0, 1, ..., n$) are $n + 1$ real constants called the coefficients of x^i in the polynomial series. The general product $a_i x^i$ is called the _ith-order term of the polynomial, and the $a_0 x^0 = a_0$ coefficient of the zero-order term is sometimes referred to as the constant term. A condensed form of this equation using summation notation is shown in Equation 2.1. Solution for the unique polynomial amounts to finding the values of the coefficients from a given set of known x,y data pairs, which are called the bivariate standards.

The solution has many applications, one of which is to produce a smooth curve so that values of unknown points may be calculated. The existence of polynomials that accomplish this is assured by the Weierstrass Approximation Theorem (73, 117, 118), which states that any continuous function can be uniformly approximated in a finite interval by a unique polynomial of sufficiently high degree; however, this should not be taken to mean that any particular method using polynomials will be universally valid because some functions are simply not "polynomial-like" throughout their entire domain and have special properties that polynomials are incapable of handling effectively unless a prohibitively high degree is employed. The practical realization of Weierstrass' Theorem of 1885 for the general case awaited the development of the high-speed computer and has been

realized only within the past thirty years (119-123). Such

numerical approximation may be applied to either the fitting

of curves of empirical data whose functional values contain

random errors or, as in our case, the approximation of

complicated mathematical functions of a real variable with

much simpler functions like polynomials that are numerically

equivalent within some arbitrary accuracy. This is

accomplished, in turn, by either of two conceptually

different approaches called curve fitting and interpolation.

In the case of interpolation, a smooth curve is

constructed from a polynomial of degree n that exactly

matches n + 1 successive data points spanning each

interpolated point within small segments of the real, but

unknown, curve. An exact fitting of invarient standards is

inherent to all interpolation processes. When this is

achieved, the model may be used to estimate additional

points so that a curve.can be plotted. The form of the

interpolation polynomial may appear to be altered, as with

Lagrange interpolation; however, these apparent differences

are for computational purposes only, and they all can be

manipulated back to the same unique polynomial of degree n

that yields their common and exact solution for n + 1

standards. These various forms of the unique interpolating

polynomial may be applied differently, however, as with the

specialized Stineman (124, 125) or with piecewise fitting of

the first and second derivitives of a third-order polynomial

as in interpolation with cubic splines (78, 117, 120, 122).
Whatever method is selected, interpolation is best suited
for fairly precise data because it provides no smoothing for
any random errors in the exactly fitted standards.

Alternately, the concept of invarient standards is
abandoned, and the polynomial may be used to approximate the
entire curve as an analytical function by fitting the data
as closely as possible through a least-squares technique
(i.e., the so-called L_2 or Euclidean norm) where the sum of
the squared deviations between the standards and the fitted
curve is minimized (71, 126). Such least-squares fitting
routines, also called regression analysis or merely curve
fitting, assume that major errors occur only in the y values
and that the errors in the x values are negligible. This
approximation, while not related by any coherent or physical
theory, is numerically equivalent to the real curve and is
also "smoothed" of any random errors of the original
standards; therefore, it may be used to find f(x) for any
value of x, including unknown values between the data
points. Unlike the Chebyshev or other more advanced norms,
least squares requires only K = n + 1 data points to solve
the single system of simultaneous equations having degree n.
The polynomial is not unique because its solution depends
upon both n and K, often becoming ill-conditioned (127) when
n + 1 approaches K. In practice, an overdetermination of
data (i.e., K >> n + 1) is usually required to find a

solution that minimizes the L_2 norm of the least-squares fit.

In the lowest nonconstant case of first-order or linear polynomials, it is assumed that the unknown points between closest standards during interpolation, or of the entire data set during curve fitting, can be described by the familiar equation of a straight line where a_0 is the y-intercept, a_1 is the slope, and all other coefficients of the polynomial are zero. The polynomial series may be expanded to higher and higher degree; however, this does not, necessarily, yield a more accurate approximation because higher-order polynomials exhibit what is called "induced stability" and tend to introduce oscillations into the fitted curve (72, 77, 128, 129) that afford a gross misrepresentation of the overall shape of the real curve. This occurs even during the interpolation process, where the fitted curve is required to pass through all the standards. Here, agreement between the standards of the unknown function and the interpolating polynomial is obtained only at a cost of very large deviations within intervals between the points. In practice, numerical experiments will often suggest the appropriate polynomial degree as a compromise between the smoothness of the overall curve and the closeness to the standards, generally being the lowest degree that fits the data within the desired accuracy and is not wasteful of computational effort.

In the present dissertation, all bivariate curve fitting by a linear least-squares orthogonal polynomial method (70, 73, 74, 79, 120, 128, 130) was accomplished using the program Polft running under the CP-6 operating system on the Honeywell DPS 8/49 mainframe computer. Polft is a public domain program written in FORTRAN and is part of the Time-Sharing-System Programs Library (131, 132) that was supplied by Honeywell.

The program allows the user to specify the lowest degree, n, of the algebraic polynomial (Equation 2.1) to be fitted, then successively fits polynomials of ascending degree in the range of a minimum $n = 0$ to a maximum $n = 11$. It is limited to 100 data points and employs unit weights. Values of the independent variable, x, should be spaced at constant intervals and have identical random errors, each individually equal to the common standard deviation of the population, for exact orthogonality of the polynomial fit to be attained (77, 132, 133). As a practical test of this orthogonality condition, least-squares fitting of higher and higher order polynomials must result in nearly unchanged coefficients for the lower-order terms (77). In a test of Polft, values of $y = 8 + 2x^4$ were fitted for all integral x values from 1 to 12. An exact fit was obtained for a polynomial order of $n = 4$, and neither the coefficients nor the error estimates were altered on going to higher orders.

Output includes a tabular presentation of the x and y standards, x_{std} and y_{std}; the calculated values of y for the fit, y_{calc}, and their absolute and relative deviations, dev_y and % dev_y, respectively; the index of determination, r^2, whose square root is the correlation coefficient, r; the standard error of estimate for the y standards, s_y; and the n + 1 polynomial coefficients. These are displayed on the monitor at each stage, and the user has the choice of going on to the next highest degree of fit or entering new read data. The definition of absolute error given by Polft in Equation 2.2 is <u>not</u> followed in either the text or the tables of Chapter 3. Instead, the equivalent expression with reversed algebraic sign was used. We merely note that both definitions are found in the literature of numerical analysis. Other statistical conventions (134) employed in these calculations follow:

$$Y_{calc} = \sum_{i=0}^{n} a_i (x_{std})^n \qquad \text{Equation 2.1}$$

$$dev_y = y_{std} - y_{calc} \qquad \text{Equation 2.2}$$

$$\% \; dev_y = 100[(dev_y)/(y_{calc})]$$

$$s_y = \sqrt{\sum_{i=1}^{K} (dev_{y,i})^2/(K - n - 1)}$$

$$\bar{y} = \sum_{i=1}^{K} y_i/K$$

$$SSTO = \sum_{i=1}^{K} (y_{std,i} - \bar{y}_{std})^2$$

$$SSE = \sum_{i=1}^{K} (y_{std,i} - \bar{y}_{calc})^2$$

$$SSR = \sum_{i=1}^{K} (y_{calc,i} - \bar{y}_{std})^2 = SSTO - SSE$$

$$r^2 = SSR/SSTO = (SSTO - SSE)/SSTO = 1 - SSE/SSTO$$

where $\bar{y}$ is the simple mean, K is the number of standards (i.e., the number of independent variables), n is the degree of the polynomial, K - n - 1 is the number of degrees of freedom (because the K data calculates n + 1 coefficients), SSTO is the "total sum of squares," SSE is the "error sum of squares," and SSR is the "regression sum of squares." This treatment contains the standard statistical assumptions that y_{std} is a normally distributed and mutually uncorrelated variable and that the residuals of the least-squares fit are independent of each other, of x_{std}, and of the parameters, being normally distributed with zero mean and constant variance (76, 134-141). These calculations were performed with a floating decimal place and double precision accuracy of 18 digits internally for the fractional part, but only eight digits (i.e., seven decimal places) are displayed in the calculated coefficients when expressed in scientific notation and only seven digits in the other calculated results.

The following equation has been selected by previous workers in this area as the standard least-squares fitting error for an X-ray scattering curve (47, 49, 62, 63, 92, 93, 97, 102, 112). It is the standard deviation of the fit expressed as a percentage of the X-ray scattering factor, f(s), at zero Bragg angle, which is a maximum on any particular scattering curve and is equal to the number of electrons associated with the atom or ion.

$$E = 100/f(0) \sqrt{\sum_{i=1}^{K} w_i (dev_{y,i})^2 / \sum_{i=1}^{K} w_i}$$

Equation 2.3

Here, w_i is the weight employed for an individual data point in the least-squares fit. In our case, weights of unity were used, so $w_i = 1$ and $\Sigma w_i = K$, the number of tabulated values fitted by the function. It should be remembered that this is only an estimate of the average error. For example, the fact that E is found to be smaller than unity does not mean that deviations greater than 1% cannot occur over parts of the curve. Consequently, a very small value of E should not be taken too seriously without scanning the individual deviations to confirm that regions of significant errors do not occur in isolated segments of the fitted curve. The fit should also be tested graphically for polynomial oscillation by using the least-squares coefficients to calculate points at much smaller intervals of the x-argument than were the original standards (109).

The mainframe computer of the Middle Tennessee State University Computer Center is located in the basement of the Cope Administration Building. Interactive access from remote terminals is provided by the Time-Sharing System

(TSS), which is an integrated and multidimensional information processing system operating under the common command language of the General Comprehensive Operating Supervisor (142). Although TSS jobs may be sent to the batch job stream, the conversational mode of entering data on a Hazeltine Esprit II terminal in the Student Computer Laboratory of Kirksey Old Main and obtaining the output from a Honeywell Model B300-22 printer located in the same room was preferred.

X-ray scattering factors, $f(s) = f(x) = y$, were calculated for any value of the Bragg angle, $s = (\sin \theta)/\lambda = x$, from the least-squares polynomial coefficients obtained from Polft, or either version of Curve Fitter, using the program Tenth Degree, which was written in BASIC by Tom Jeffries (Appendix I, June 1984) for the IBM PC and compatible microcomputers (143, 144) and performed with a floating decimal place in single precision (9 digit) accuracy. This program prompts the user for the coefficients of any given algebraic polynomial (Equation 2.1) up to the tenth degree. If the degree of the fitted polynomial is lower than $n = 10$, zeros are entered for the coefficients of higher-order terms (i.e., $i > n$). The program allows the continuation with the same coefficients by entering the number 1 when the user is prompted for a change. If a change is desired, any other single number digit is entered. The numerical value of the independent variable, x, must correspond to the transposed

form, such as $x = s^2$ or $x = \exp(-s)$, that was employed in the original least-squares fit. The calculated result, the entered x value, and the coefficients are displayed on the monitor, and a hardcopy may be obtained from the printer.

Lagrange interpolation (70, 73-75, 80, 129, 145) was performed initially with a Perkin-Elmer 3600 Data Station (16K bytes of random access memory) using the program Lagi, which was written in BASIC by Tom Jeffries (July 1983) and later modified for IBM PC and compatible microcomputers (Appendix II, October 1986). This program was validated by checking it against a set of test data (146). Lagrangian interpolation was originally developed to allow solution by hand calculation (147), but it is still used today with most of the computational work delegated to the computer. It is expressed explicitly in terms of the ordinates, rather than in finite or divided differences, and may be applied to both evenly spaced and unevenly spaced data. It should be noted that a "K-point Lagrange" requires K = n + 1 data arranged as symmetrically as possible about the point to be interpolated and corresponds to an interpolating polynomial of degree n.

The program Lagi requires that a set of standards be entered on line 230 of the program. When the data statement is entered, the program prompts the user for the number of standards and the value to be interpolated. The user then has the choice of continuing interpolation using the same

set of standards or writing a new data statement on line 230. Only 249 characters are allowed as a maximum line length on the Perkin-Elmer Data Station. This includes the number "230" (3 characters), the word "data" (4 characters), one space following the line number and another following the line statement (1 + 1 = 2 characters), the 2K decimal points and 2(4)K digits of the K standards, and the 2K - 1 commas separating the data entries. For example, a total of K = 20 bivariate data having four significant figures may be entered as calculated from

$$3 + 4 + 2 + 2K + 8K + (2K - 1) \leq 249$$

These calculations were performed with a floating decimal point at single precision (6 digit) accuracy, but double precision accuracy (16 digit) may be chosen as an option. Lagi is not as convenient a program to use as the previous ones described. A hardcopy listing the standards and interpolated values is only obtainable on the IBM PC version, and the Lagrange coefficients employed in the interpolations are not included as output in any form. Furthermore, the data statement of the program must be rewritten each time the standards are changed.

Finally, the program Curve Fitter was written in 1980 for Interactive Microware, Inc. by Paul K. Warme (125). Curve Fitter has many options available including data

scaling (to produce a compact plot), N-point data averaging (to reduce the number of data), and N-point data smoothing (to average noise); interpolation of y values by polynomial (degree 1 to 6), cubic spline, or Stineman Methods (124); and linear least-squares fitting by any of three functions: an ordinary (i.e., nonorthogonal) algebraic polynomial function (degree 1 to 6) solved directly by Gaussian elimination (148), a geometric function (i.e., $y = a(x)^b$ by fitting the logarithmic transformation), or a one-term exponential function (i.e., $y = a[exp(bx)]$ also by fitting the logarithmic transformation). The program is relatively easy to use, being versatile in user interaction and selection of methods.

Data may be entered from the keyboard or input from either a previously stored disc file or from an optional analog to digital converter for instrumental interfacing. Pairs of either evenly-spaced bivariate data or y values with constant x interval can be entered. After all the standards have been input, the program allows the abscissa and ordinate axes to be redefined so that inverse interpolation (i.e., interchange of x and y data) is possible; however, only linear interpolation of the x values is then applied. The raw data at this point are in the natural units that were originally measured, but these can be scaled through use of a multiplying scale factor, offset by a constant amount, or converted to either common or

natural logarithms if desired. The standards can be listed and plotted using 320 by 200 point graphics while the program automatically calculates suitable numeric labels along the scaled axes. Any one of the interpolation or least-squares methods may then be applied to calculate a curve superimposed upon the standards.

After all calculations have been made and the curve plotted, the user must decide if the fit is satisfactory and, also, whether the fitted curve is to be erased. One may choose not to erase the curve so that it can be visually compared with other curves obtained from different fitting methods or with polynomials of higher degree. When an acceptable fit has been achieved, unknown values of other points may be calculated. Statistical measures of the accuracy of the least-squares fit include the coefficient of determination, the correlation coefficient (i.e., the square root of the coefficient of determination), and the standard error of estimates. The observed and calculated values, together with the percentage deviation, are also listed. All working files may be saved on the same disc for later addition or deletion of individual data points. A hardcopy of the plotted curve can be obtained using the compatible program Scientific Plotter, also written in 1980 by Paul K. Warme (149), which accepts the formats of the standards file from Curve Fitter as input and also accepts data from the keyboard as well as data calculated by a user-defined

subroutine. Multiple sets of data can be plotted on the
same graph using any one of 20 different plotting symbols
and, as an option, superimposed variable-length error bars.

Early use of Curve Fitter at Middle Tennessee State
University was with a 64K Apple II+ microcomputer having
Applesoft ROM, floating decimal point, and single precision
(9 digit) accuracy; however, most of our final results that
are reported in Chapter 3 were calculated on a Kaypro-PC
microcomputer at Campbellsville College, Campbellsville,
Kentucky using the program Curve Fitter-PC, which is version
of Curve Fitter written for the IBM PC. The Kaypro-PC has
an 8088-2 chip that runs on eight megahertz instead of the
4.7 MHz chip of the IBM PC. This provides a much faster
calculation time, about 1/4 the time required for the IBM PC
and about 1/6 that required for the Apple II+.

The minimal requirements of Curve Fitter-PC is a 128K
IBM PC or compatible microcomputer, a color/graphics board,
a color monitor, one disc drive, and at least PC DOS 1.1 or
greater. It contained the 1984 improvements of Corl and
Warme (150), including function-key plot and graph-labeling
capabilities; however, the actual calculation routines in
single precision remained unchanged. The advanced IBM
BASIC, BASICA, is contained, in part, on the ROM of the IBM
PC and is required for the program to automatically execute
itself. This had to be changed to Microsoft BASIC (143) in
order for it to be totally compatible with the Kaypro-PC.

CHAPTER 3

Results and Discussion

In Chapter 1, the crystallographic literature was reviewed for previous applications of numerical analysis to the atomic scattering curves of X-ray diffraction. In the present chapter, the results of our own investigations in this area using the computational methods detailed in Chapter 2 are discussed. There are many tables in this chapter that must be referred to in several different orders. Consequently, all tables have been placed at the end of the chapter as a convenience to the reader.

The present study is limited to numerical methods that perform an analytical operation with polynomial functions. Other approximations that are based upon trigonometric and exponential functions enjoy wide use as computational algorithms; however, none of these have managed to displace polynomials from their central role in numerical analysis for computer evaluation of mathematical functions (105). Whatever approximating function is finally selected, it can be applied in either of two conceptually different ways (69-80). First, it can be employed in an interpolation process that matches the data exactly and estimates additional, unknown values of the real curve within intervals between these standards. Alternately, it can be

38

used to curve fit the data by a least-squares method where the sum of the squared deviations between the data and the fitted curve of the approximating function is made to be a minimum at the data points.

Least-squares fits are often applied to data where random errors occur in the measured values, the curve passing as close as possible to the known points within the limitations of the fitting function. The purpose here, however, would be to approximate the precise values of an X-ray scattering table in the condensed form of a single, analytical function whose least-squares coefficients accurately reproduce the continuous scattering curve and, hence, the set of individual scattering factors. Thus, the complex function from physical theory that originally calculated the X-ray scattering curve is replaced with a numerically equivalent, but mathematically much simpler, polynomial function in a process which is called numerical approximation or, sometimes, analytic substitution.

We have followed, for both interpolation and curve fitting, the common practice of taking $s = (\sin \emptyset)/\lambda = x$ as an exact, independent variable and the value of the atomic scattering factor, $f(s) = f(x) = y$, as a real function dependent only on s. While the absolute value of $f(s)$ will contain determinate errors inherent to their original quantum mechanical evaluation, the relative values within any particular X-ray scattering curve must be internally

self-consistent because they were all calculated from the same wave function. Consequently, the atomic scattering tables contain no indeterminate or random errors and require no smoothing for numerical analysis. We have also assumed that all data of a given X-ray scattering table are of equal importance in determining its scattering curve and have used unit weights in all our calculations. The Criterion of Fit that defines the successful interpolation of a scattering factor or a good least-squares fit of the scattering table is taken as $dev_{f(s)} = f(s)_{calc} - f(s)_{table} \leq \pm 0.002$ because the X-ray scattering factors are generally calculated with three decimal places. The choice of this criterion over ± 0.001 avoids trivial complications due to roundoff errors in the last significant figure for differences between our fitted values.

X-ray scattering factors for the light atoms hydrogen (1 e$^-$) and nitrogen (7 e$^-$), the intermediate iron(II) cation (24 e$^-$), and the heavy bromine atom (35 e$^-$) and cesium(I) cation (54 e$^-$) were selected for our interpolation study from International Tables for X-ray Crystallography (44) and are plotted in Chapter 1 as Figure 1. This selection was made for the somewhat arbitrary reasons that the species are chemically significant and represent a range of atomic numbers, ionization states, and different positions on the periodic table. The values of f(s) are tabulated at s intervals of 0.01 up to 0.20 Å^{-1}, 0.02 up to 0.50 Å^{-1}, 0.05

up to 0.70 Å^{-1}, and 0.10 thereafter. The scattering tables contained 51 f(s) values ranging $0.00 \leq s \leq 1.50$ for atomic hydrogen and 56 f(s) values ranging $0.00 \leq s \leq 2.00$ for each of the other atoms or ions. It was noted by the editors of International Tables that "these intervals are such that linear interpolation is reasonably accurate (44)."

Except for hydrogen, these X-ray scattering factors were calculated by Doyle and Turner (47, 48) using the relativistic Hartree-Fock atomic wave functions of Coulthard (151) that included the potential of a point, rather than a finite, nucleus (i.e., neglect of magnetic term) and also ignored the effects of electron correlation within the atom (i.e., neglect of off-diagonal Lagrange parameters). Doyle and Turner (47) noted that their results are tabulated to three decimal places "as an aid to interpolation" even though the absolute accuracy of the third (and probably also the second) decimal place is questionable. These tables have been since revised to include high-angle data (46); however, the numerical values for the ranges studied herein have remained unchanged. For hydrogen, the non-relativistic Hartree-Fock calculations of Cromer and Mann (152), using the wave functions of Mann (153), were employed. The atomic model in all cases was the usual isolated, symmetrical sphere of uniform electron density in its ground state and at rest.

The act of sampling a continuous function implies a potential loss of information. Thus, our first concern was to find the minimum interval of s that is required for an accurate interpolation of f(s). This is sometimes called the step size or grid spacing. The values of f(s) for the X-ray scattering factors described above (47) were obtained by numerical integration over the range $0.00 \leq s \leq 6.00$ by repeated use of an equal-interval formula based on seven successive points for up to 180 of the 201 values of the radial electron density provided by the wave function (151). It would seem, perhaps naively, that $\Delta s_{min} = 6.00/180 = 0.033$ should be the "natural" grid spacing of this data set.

The literature establishing the intervals employed in current tabulations of X-ray scattering factors (44, 45) is sparse, nonrigorous, and appears based upon conclusions drawn during the first applications of digital computers in X-ray crystallography (20, 29, 63, 87, 112, 154, 155) using early and often inaccurate scattering tables. Simple table searching without interpolation would result in a maximum, absolute error in s of one-half the grid spacing selected (64). Storage of the f(s) values at smaller and smaller intervals of s would, of course, increasingly improve the accuracy of the table search, but at a corresponding increase in the computer-storage requirements as well as the cost of the calculation in time and money.

Computer memory requirements can be reduced with equivalent accuracy by linear interpolation of scattering factors for each atomic species when tabulated as $f(s^2)$, a more nearly linear function. This has been applied at equally spaced intervals of s^2 (156), but is much more powerfully employed using successive intervals graded in groups so as to give roughly constant increments in the intervals of the corresponding values of the scattering factors (157). The latter method is in use today in many computer programs for crystallographic calculations, and variations of this procedure involve the Fourier Transforms of the X-ray scattering curves (158) as well as semi-log and log-log cubic spline interpolation (36, 46). It is usually also possible to roughly fit two straight lines through a plot of $\log[f(s)]$ against s^2, one through the high-angle region and one through the low-angle region (62). The size of the grid spacing has also been shown to be critical in the related calculation of structure factors during single-crystal X-ray structure determinations (66, 67, 103, 158, 159).

Initial interpolation studies of the atomic scattering factors of bromine and nitrogen were performed by cubic spline, Stineman, and polynomial ($2 \leq n \leq 6$) methods using the program Curve Fitter. The first degree interpolating polynomial ($n = 1$) is, of course, merely a straight line and was excluded. Our methodology was to enter all $f(s)$ values

on the Curve Fitter disc at selected table intervals of s

over the $0.00 \leq s \leq 2.00$ range of the x-argument, then

remove each individual value one at a time and use the

remaining data to interpolate it. The deleted value was

then returned to the standards file, and new ones were

selected for successive deletion and interpolation.

As shown in Tables 2 and 5, interpolation of the

nitrogen and bromine atomic scattering factors using only

those values of s spaced at 0.10 were within ± 0.002 for

values at s above ca. 1.0 for third and higher polynomial

orders; however, interpolation at smaller values of s are

seen to be progressively inaccurate. The effect of using

interpolating polynomials of higher degree is shown in

Figure 2, where the domains of the scattering table that

accurately interpolate within the Criterion of Fit are

plotted. The curves for nitrogen and bromine are roughly

equivalent and show that no polynomial of any order can be

made to interpolate X-ray scattering factors within ± 0.002

when they are below an s of about 0.6 or 0.7 and spaced at

0.10 intervals. Figure 3 shows that, when the grid density

was increased to include also every s unit of 0.05 up to s

= 0.70, the interpolation of nitrogen scattering factors

(Table 3) was somewhat improved, but the domain within the

Criterion of Fit for bromine (Table 6) was virtually

unchanged. The full range of data, excepting f(s) values

at s = 1.0, could be interpolated within ± 0.002 at fourth
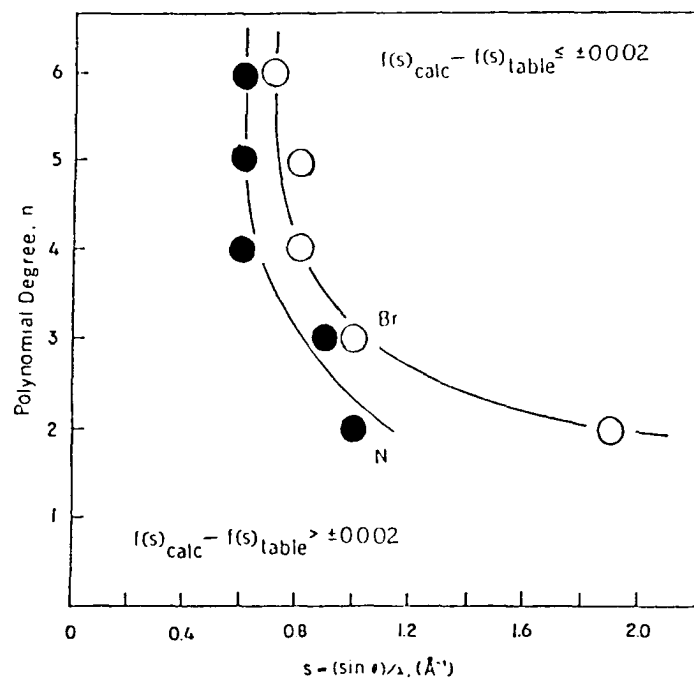
Figure 2

Domains Within Criterion of Fit
for s = 0.10 Intervals
During Interpolation
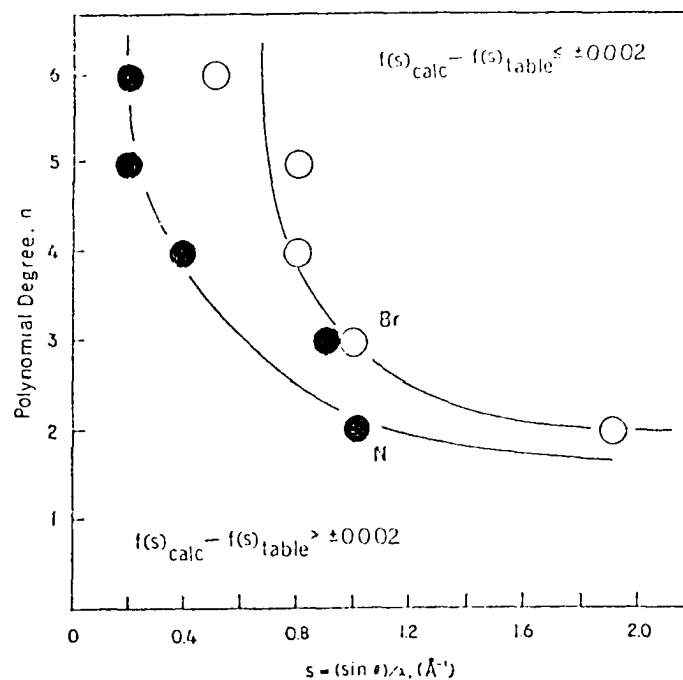
Figure 3

Domains Within Criterion of Fit
for s = 0.10 and Available 0.05
Intervals During Interpolation

and higher polynomial orders when intervals of 0.02 were added below s $\leq$ 0.50 (Tables 4 and 7).

The interpolations were repeated using the cubic spline, Stineman, third- and fourth-degree polynomials, and the four-point Lagrange applied to s intervals of 0.02 up to 0.50 Å^{-1}, 0.05 up to 0.70 Å^{-1}, and 0.10 Å^{-1} thereafter. In addition to nitrogen (Table 9) and bromine (Table 11), hydrogen (Table 8), iron(II) (Table 10), and cesium(I) (Table 12) were also included in the final calculations. These various interpolation methods all share the common feature of being expressed explicitly in terms of the ordinates rather than in terms of their differences or divided differences. As a result, they all may be applied to both evenly spaced and unevenly spaced data. The four-point Lagrange is mathematically identical to the third-degree interpolating polynomial, and small ±0.001 differences between these two provide a measure of the roundoff error involved in the present calculations. The four coefficients of this third-degree interpolating polynomial are calculated from n + 1 = 4 successive standards spanning each point to be fitted. In the case of cubic spline interpolation, each set of four successive points are employed in a piecewise fashion to calculate the three coefficients of the continuous first and second derivitives of a cubic polynomial for each pair of adjacent standards. As can be seen in Tables 2-12, the cubic spline

consistently out-performed the third-order interpolating polynomial for all but the most extreme value of s. It was also more efficient in computer time. The Stineman method, however, gave inconsistent results to show that this specialized interpolation technique is inappropriate for X-ray scattering curves. Its general performance was somewhere between that of interpolating polynomials having second and third order.

It was found that the light elements hydrogen and nitrogen could be accurately interpolated within ± 0.002 for all values of s using third- and higher-order polynomials or even the Stineman. The iron(II), bromine, and cesium(I) interpolated scattering factors had a small patch of relatively small error in the intermediate range of s from $0.8-1.1$, $0.8-1.0$, and $0.7-1.3$, respectively, when a third-order polynomial was applied and still showed a "delinquent" point at s = 1.0 with an error of ± 0.002 to ± 0.004 at the fourth degree. This latter error, as seen for bromine in Table 7, did not disappear until an interpolating polynomial of the sixth degree was employed. It is clear that the x-grid spacing is too coarse in this region for reliable interpolation; however, inclusion of finer intervals over the range of about $0.50 \leq s \leq 1.50$ would then allow completely accurate interpolation using even a third-degree polynomial. This might amount to extending the 0.05 intervals to s = 1.50, but might also require that an

even finer grid spacing be employed for this especially difficult region to fit.

The calculations up to this point have dealt with interpolation only. Analytical representations of X-ray scattering curves are advantageous in computer calculations since otherwise the complete scattering table must be stored for interpolation. Our attempts at linear least-squares curve fitting involved an orthogonal Legendre polynomial as contained in the program Polft. Polft was employed to curve fit the atomic scattering factors from International Tables (44) for all neutral atoms from hydrogen (1 e⁻) through neon (10 e⁻) at all polynomial orders from zero (n = 0) through eleven (n = 11) as shown in Tables 13-16. The data should be evenly spaced in order to maintain the orthogonality of the polynomial. The standards here, however, were unevenly spaced. Thus, the results should actually be similiar to least-squares fits with ordinary polynomials.

The linear least-squares fitting errors from the orthogonal polynomial are approximately a factor of ten larger than those obtained by nonlinear least-squares fitting of equivalent Gaussian expansions or polynomial series as documented by the citations in Chapter 1. For example, the values presented in Table 15 have an average error of $E = \pm 0.60\%$ for the linear least-squares fitting of an eighth-degree polynomial, which has the same number of coefficients as the nine-parameter Gaussian expansion whose

nonlinear fitting error is summarized in Table 1 of Chapter 1. It is clear, then, that the solution for the best set of polynomial coefficients presents a nonlinear minimization problem.

Still, it is interesting to consider how close a linear least-squares fitting technique could actually approach the best approximation of an X-ray scattering curve. For example, the fitting errors shown in Tables 13-16 decrease as the polynomial order increases from sixth to ninth degree suggesting a beneficial use of high-degree polynomials. On the other hand, merely increasing the polynomial order to higher and higher degree does not, necessarily, produce a better approximation of the overall X-ray scattering curve, because convergence to the standards may result in divergence from the points on the real curve in between due to polynomial oscillation. Such oscillations grow without bound as the degree of the polynomial series is increased.

Another approach to this problem is transformation. When polynomial approximation alone does not appear to be appropriate, one or both of the variables can be transformed to make the function more amenable to linear least-squares fitting. There are no strict rules concerning this option, and the suitable transformation must be found by the trial and error results of numerical experiment.

The X-ray scattering factors due to a single hydrogen atom covalently bound in a hydrogen molecule was selected as

a simple test case. The X-ray scattering table for bonded

hydrogen has been calculated by Stewart, Davidson, and

Simpson (96) using a modified form of the Kolos-Roothaan

wave function. These calculations were facilitated by the

symmetry of the hydrogen molecule. Both sides of the

diatomic molecule give the same one-electron density

function that, when normalized to unity, is half the value

of the H_2 molecular scattering factor. A spherical shape

was assumed (i.e., a spherical approximation) for both

hydrogen atoms individually within the molecule, and the

electron density of each atom was allowed to float away from

its central proton until an error function was minimized. A

minimum error of $\pm 0.11\%$ was obtained when the center of the

spherically symmetrical electron density of each atom was

optimally positioned 0.070 Å off the protons and into the

covalent bond. The resulting X-ray scattering curve of

bonded hydrogen is compared with the X-ray scattering curve

obtained from the isolated atom in Figure 4.

An analytical approximation of the X-ray scattering

curve of bonded hydrogen using the coefficients obtained by

nonlinear least-squares fitting of a nine-parameter Gaussian

expansion have been reported (44). The scattering table has

been extrapolated to higher Bragg angles and interpolated to

finer grid spacing (36). For our study, we employed the 46

original values of f(s) for the range $0.0000 \leq s \leq 1.7176$

reported by Stewart, Davidson, and Simpson (96). Our
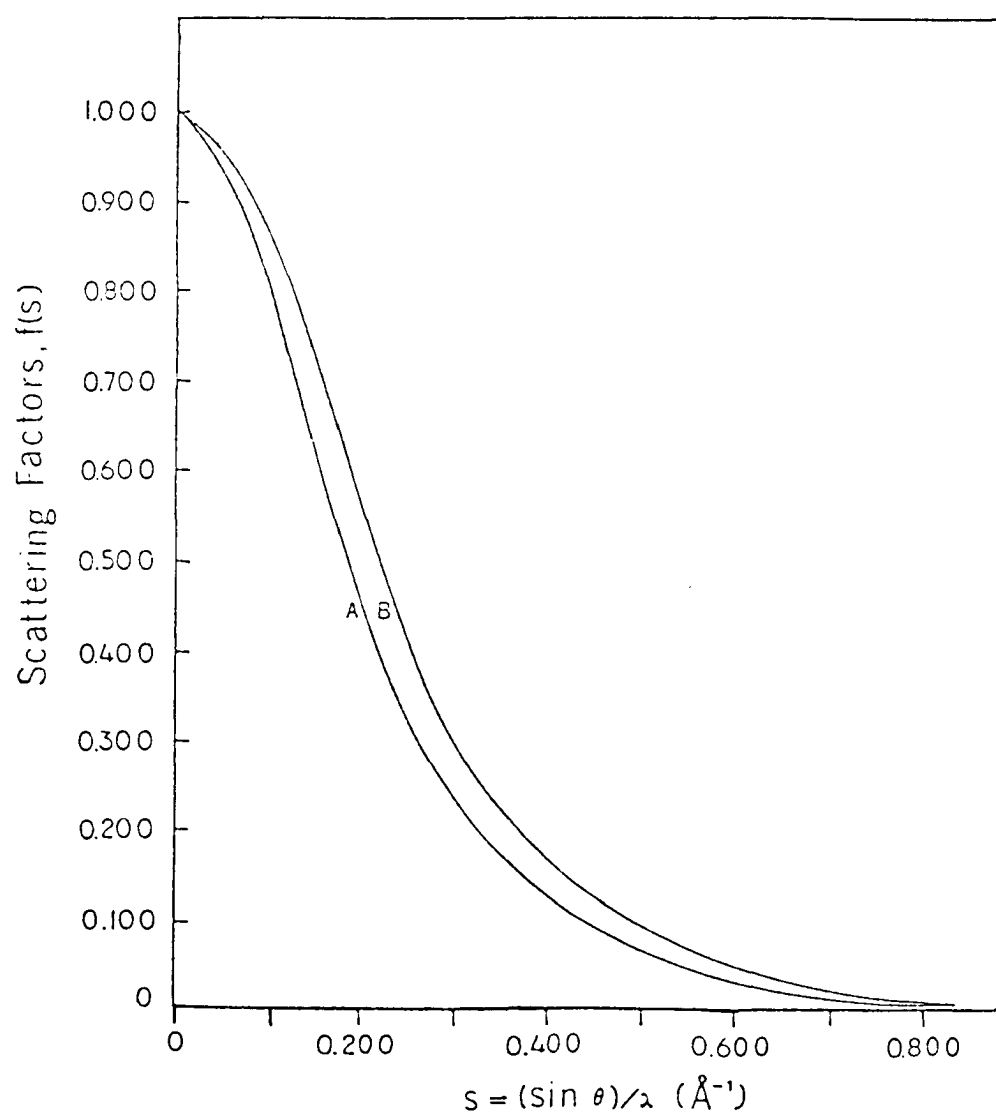
Figure 4

Atomic Scattering Curves for Hydrogen:
Free Atom (Curve A) and Bonded (Curve B)

least-squares fits were repeated using 41 values selected at evenly spaced intervals of s = 0.0429 in an attempt to maintain the orthogonality of the polynomial.

The results from x = s, x = exp$(-s^2)$, and x = exp(-s) are shown in Tables 17-19, respectively. Small improvement in linear least-squares fitting did result from using evenly spaced data; however, the polynomial coefficients varied with the degree of the polynomial to show that orthogonality was not attained. Clearly, the x = exp(-s) transformation of Table 19 produced the best fit. That this inclusion of a Gaussian function into a polynomial series also affords a better <u>linear</u> least-squares fit than does the <u>nonlinear</u> least-squares fitting of a Gaussian expansion is highly significant and implies that its "natural" shape more closely approximates that of an X-ray scattering curve prior to any iterations. This comparison is made between the linear fit of our eighth-degree orthogonal polynomial (E = 0.063%) with the corresponding nine-parameter Gaussian expansion (E = 0.072%) fitted by nonlinear least-squares (44).

Table 2

Scattering Factor Differences (X 10^3) Interpolated for
Nitrogen at 0.10 Intervals of s = (sin Θ)/λ

s, $\overset{\circ}{A}^{-1}$	Cubic Spline	Stineman	POLYNOMIALS				
			n = 2	n = 3	n = 4	n = 5	n = 6
0.00	–	–	–	–	–	–	–
0.10	-336	-398	357	206	-87	-12	28
0.20	146	-136	355	137	44	0	-9
0.30	4	136	82	65	-44	-4	6
0.40	4	50	48	4	-36	-4	-6
0.50	-7	27	56	-10	-9	-5	5
0.60	-3	11	36	-8	1	0	3
0.70	-2	2	20	-6	2	0	0
0.80	-1	1	9	-3	2	0	0
0.90	0	1	3	-1	-1	0	0
1.00	0	0	-1	-1	0	0	0
1.10	0	0	0	0	1	0	0
1.20	0	0	0	0	0	0	0
1.30	0	0	1	0	0	0	0
1.40	0	0	1	0	0	0	0
1.50	0	0	0	0	0	0	0
1.60	0	0	0	0	0	0	0
1.70	0	0	0	0	0	0	0
1.80	0	0	0	0	0	0	0
1.90	1	2	0	0	-1	-1	-1
2.00	–	–	–	–	–	–	–
Mean Error	26	40	51	23	12	1	3

Table 3

Scattering Factor Differences (X 10^3) Interpolated for
Nitrogen from Available Data at 0.05 and
0.10 Intervals of s = (sin θ)/λ

s, Å^{-1}	Cubic Spline	Stineman	POLYNOMIALS				
			n = 2	n = 3	n = 4	n = 5	n = 6
0.00	–	–	–	–	–	–	–
0.10	20	-36	58	0	-6	-4	-4
0.20	5	3	4	1	0	1	-1
0.30	1	6	1	3	3	-1	-1
0.40	1	4	-8	-1	-1	-1	0
0.50	-1	2	-7	-1	0	0	0
0.60	-1	1	-4	-1	0	0	0
0.70	-1	-1	-7	0	0	0	0
0.80	-1	1	9	-3	2	0	0
0.90	0	1	3	-1	-1	0	0
1.00	0	0	-1	1	0	0	0
1.10	0	0	0	0	1	0	0
1.20	0	0	0	0	0	0	0
1.30	0	0	1	0	0	0	0
1.40	0	0	1	0	0	0	0
1.50	0	0	0	0	0	0	0
1.60	0	0	0	0	0	0	0
1.70	0	0	0	0	0	0	0
1.80	0	0	0	0	0	0	0
1.90	1	2	0	0	-1	-1	-1
2.00	–	–	–	–	–	–	–
Mean Error	2	3	6	1	1	<1	<1

Table 4

Scattering Factor Differences $(X\ 10^3)$ Interpolated
for Nitrogen from Available Data at 0.02, 0.05,
and 0.10 Intervals of $s = (\sin\theta)/\lambda$

s, Å⁻¹	Cubic Spline	Stineman	POLYNOMIALS				
			n = 2	n = 3	n = 4	n = 5	n = 6
0.00	–	–	–	–	–	–	–
0.10	0	-1	4	0	0	0	0
0.20	0	1	2	0	0	0	0
0.30	0	0	0	0	0	0	0
0.40	0	0	-1	-1	0	-1	0
0.50	0	0	-7	0	0	0	0
0.60	0	1	-4	-1	0	0	0
0.70	-1	1	-7	-1	0	0	0
0.80	-1	1	9	-1	0	0	0
0.90	0	1	3	-1	-1	0	0
1.00	-1	-1	-1	-1	0	0	0
1.10	0	0	0	0	1	0	0
1.20	0	0	0	0	0	0	0
1.30	0	0	1	0	0	0	0
1.40	0	1	1	1	0	0	0
1.50	0	0	0	0	0	0	0
1.60	0	0	0	0	0	0	0
1.70	0	0	0	0	0	0	0
1.80	0	0	0	0	0	0	0
1.90	1	2	0	0	-1	-1	-1
2.00	–	–	–	–	–	–	–
Mean Error	<1	1	2	<1	<1	<1	<1

Table 5

Scattering Factor Differences (X 10^3) Interpolated for
Bromine at 0.10 Intervals of s = (sin θ)/λ

s, Å^{-1}	Cubic Spline	Stineman	POLYNOMIALS n = 2	n = 3	n = 4	n = 5	n = 6
0.00	−	−	−	−	−	−	−
0.10	−949	−1076	−987	−715	−454	−229	−58
0.20	486	404	986	477	227	92	19
0.30	−54	121	34	99	−226	−70	−11
0.40	−13	7	−164	−49	−89	−49	12
0.50	−18	−27	−66	−35	1	−4	26
0.60	−1	−6	3	−8	16	4	5
0.70	−2	8	19	5	8	3	1
0.80	4	11	10	7	2	2	−1
0.90	1	9	−4	3	−2	0	−1
1.00	−1	−8	−2	−1	−2	−2	−1
1.10	0	8	−17	0	−1	0	0
1.20	0	6	−17	1	0	0	0
1.30	−1	4	−9	−1	0	0	0
1.40	−1	3	−13	−2	0	0	0
1.50	0	1	0	−2	1	0	2
1.60	0	1	−7	−1	0	0	0
1.70	0	1	−5	−1	0	0	0
1.80	0	0	−3	−1	0	0	0
1.90	0	3	−2	1	0	0	0
2.00	−	−	−	−	−	−	−
Mean Error	81	90	124	74	54	24	7

Table 6

Scattering Factor Differences (X 10^3) Interpolated
for Bromine from Available Data at 0.05 and
0.10 Intervals of s = (sin θ)/λ

| s, Å⁻¹ | Cubic Spline | Stineman | POLYNOMIALS | | | | |
			n = 2	n = 3	n = 4	n = 5	n = 6
0.00	–	–	–	–	–	–	–
0.10	49	-120	206	16	-28	-56	-109
0.20	-56	28	49	-32	-57	-72	-66
0.30	-73	-38	-119	-64	-107	-77	-100
0.40	-10	1	-19	-4	-10	-5	-19
0.50	-2	-2	-5	-2	1	0	0
0.60	-3	0	2	-1	0	0	-1
0.70	-2	8	19	5	8	3	1
0.80	4	11	10	7	2	2	-1
0.90	1	9	-4	3	-2	0	-1
1.00	-1	-8	-2	-1	-2	-2	-1
1.10	0	8	-17	0	-1	0	0
1.20	0	6	-17	1	0	0	0
1.30	-1	4	-9	-1	0	0	0
1.40	-1	3	-13	-2	0	0	0
1.50	0	1	0	-2	1	0	2
1.60	0	1	-7	-1	0	0	0
1.70	0	1	-5	-1	0	0	0
1.80	0	0	-3	-1	0	0	0
1.90	0	3	-2	1	0	0	0
2.00	–	–	–	–	–	–	–
Mean Error	11	13	27	8	12	11	16

Table 7

Scattering Factor Differences (X 10^3) Interpolated
for Bromine from Available Data at 0.02, 0.05,
and 0.10 Intervals of s = (sin θ)/λ

s, Å⁻¹	Cubic Spline	Stineman	POLYNOMIALS				
			n = 2	n = 3	n = 4	n = 5	n = 6
0.00	–	–	–	–	–	–	–
0.10	0	0	14	0	0	0	0
0.20	0	1	3	0	-1	-1	-1
0.30	0	0	-2	0	-1	0	0
0.40	0	0	0	0	0	0	0
0.50	0	0	0	0	0	0	0
0.60	0	0	2	-1	0	0	0
0.70	0	2	4	1	0	0	0
0.80	3	1	7	5	1	1	0
0.90	1	9	-4	4	-2	-1	0
1.00	1	10	-5	3	-4	-3	-2
1.10	0	8	-17	0	-1	0	0
1.20	0	6	-17	-1	0	0	0
1.30	-1	4	-15	-1	0	0	0
1.40	-1	3	-13	-1	0	0	0
1.50	-1	1	-10	-1	0	0	0
1.60	0	1	-7	-1	0	0	0
1.70	0	1	-5	-1	0	0	0
1.80	0	0	-3	-1	0	0	0
1.90	0	3	-2	1	0	0	0
2.00	–	–	–	–	–	–	–
Mean Error	<1	3	6	1	1	<1	<1

Table 8

Scattering Factor Differences Interpolated for Hydrogen
from the Available Data at 0.02, 0.05, and 0.10
Intervals of s = (sin θ)/λ

s, $\AA^{-1}$	$f(s)$	Cubic Spline	Stineman	-POLYNOMIALS- n = 3	n = 4	4-Point Lagrange
0.00	1.000	-	-	-	-	-
0.10	0.811	-0.001	-0.001	-0.001	0.000	0.000
0.20	0.481	-0.001	0.000	0.000	-0.001	0.000
0.30	0.251	0.000	0.000	0.000	0.000	0.000
0.40	0.13	0.000	0.000	0.000	0.000	0.000
0.50	0.071	-0.001	-0.001	-0.001	0.000	-0.001
0.60	0.040	0.000	0.000	0.000	-0.001	0.000
0.70	0.024	0.000	0.000	0.000	0.000	0.000
0.80	0.015	0.000	0.000	0.000	0.000	0.000
0.90	0.010	0.000	0.000	0.000	0.000	0.000
1.00	0.007	0.000	0.000	0.000	0.000	0.000
1.10	0.005	-0.001	0.000	0.000	0.000	-0.001
1.20	0.003	0.001	0.001	0.001	0.001	0.001
1.30	0.003	-0.001	-0.001	-0.001	0.000	-0.001
1.40	0.002	0.000	0.000	0.001	0.000	0.000
1.50	0.001	-	-	-	-	-
1.60	none	-	-	-	-	-
1.70	none	-	-	-	-	-
1.80	none	-	-	-	-	-
1.90	none	-	-	-	-	-
2.00	none	-	-	-	-	-

Table 9

Scattering Factor Differences Interpolated for Nitrogen
from the Available Data at 0.02, 0.05, and 0.10
Intervals of s = (sin θ)/λ

s, $Å^{-1}$	f(s)	Cubic Spline	Stineman	-POLYNOMIALS- n = 3	n = 4	4-Point Lagrange
0.00	7.000	-	-	-	-	-
0.10	6.180	0.000	-0.001	0.000	0.000	0.000
0.20	4.563	0.000	0.001	0.000	0.000	0.000
0.30	3.219	0.000	0.000	0.000	0.000	0.000
0.40	2.393	0.000	0.000	-0.001	0.000	-0.001
0.50	1.942	0.000	0.000	0.000	0.000	0.000
0.60	1.697	0.000	0.001	-0.001	0.000	-0.001
0.70	1.551	-0.001	0.001	-0.001	0.000	-0.001
0.80	1.445	-0.001	0.001	-0.001	0.000	-0.002
0.90	1.353	0.000	0.001	-0.001	-0.001	-0.001
1.00	1.265	-0.001	-0.001	-0.001	0.000	-0.001
1.10	1.177	0.000	0.000	0.000	0.001	0.000
1.20	1.090	0.000	0.000	0.000	0.000	0.000
1.30	1.004	0.000	0.000	0.000	0.000	0.000
1.40	0.921	0.000	0.001	0.001	0.000	0.001
1.50	0.843	0.000	0.000	0.000	0.000	0.000
1.60	0.769	0.000	0.000	0.000	0.000	0.000
1.70	0.700	0.000	0.000	0.000	0.000	0.000
1.80	0.636	0.000	0.000	0.000	0.000	0.000
1.90	0.578	0.001	0.002	0.000	-0.001	0.000
2.00	0.525	-	-	-	-	-

Table 10

Scattering Factor Differences Interpolated for Iron(II)
from the Available Data at 0.02, 0.05, and 0.10
Intervals of s = (sin θ)/λ

s, Å^{-1}	f(s)	Cubic Spline	Stineman	-POLYNOMIALS- n = 3	n = 4	4-Point Lagrange
0.00	24.000	–	–	–	–	–
0.10	22.889	0.000	-0.001	0.000	0.000	0.000
0.20	20.140	-0.001	-0.002	-0.001	-0.001	-0.001
0.30	16.871	0.000	0.000	0.000	0.000	0.000
0.40	13.881	0.001	0.001	0.001	0.001	0.001
0.50	11.494	0.000	0.001	0.000	0.000	0.000
0.60	9.737	0.000	0.003	0.001	0.000	0.000
0.70	8.501	0.000	0.006	-0.001	0.000	-0.001
0.80	7.640	-0.001	0.007	-0.003	0.000	-0.003
0.90	7.023	-0.003	0.007	-0.005	-0.001	-0.004
1.00	6.546	-0.001	0.005	-0.004	0.002	-0.005
1.10	6.144	-0.002	0.002	-0.004	0.000	-0.004
1.20	5.775	-0.001	0.001	-0.002	0.001	-0.002
1.30	5.419	-0.001	-0.001	-0.001	0.000	-0.001
1.40	5.068	0.000	0.000	0.000	0.000	-0.001
1.50	4.722	0.000	0.000	0.000	0.000	0.000
1.60	4.384	0.000	0.000	0.000	0.000	0.000
1.70	4.058	0.001	0.001	0.000	0.000	0.000
1.80	3.749	-0.001	0.000	0.000	0.000	0.000
1.90	3.459	0.004	0.012	0.001	0.001	-0.001
2.00	3.192	–	–	–	–	–

Table 11

Scattering Factor Differences Interpolated for Bromine
from the Available Data at 0.02, 0.05, and 0.10
Intervals of s = (sin θ)/λ

s, Å^{-1}	f(s)	Cubic Spline	Stineman	-POLYNOMIALS- n = 3	n = 4	4-Point Lagrange
0.00	35.000	–	–	–	–	–
0.10	32.450	0.000	0.000	0.000	0.000	0.000
0.20	27.749	0.000	0.001	0.000	-0.001	0.000
0.30	23.857	0.000	0.000	0.000	0.000	0.000
0.40	20.874	0.000	0.000	0.000	0.000	0.000
0.50	18.307	0.000	0.000	0.000	0.000	0.000
0.60	15.958	0.000	0.000	-0.001	0.000	-0.001
0.70	13.837	0.000	0.002	0.001	0.000	0.001
0.80	12.001	0.003	0.001	0.005	0.001	0.005
0.90	10.480	0.001	0.009	0.004	-0.002	0.004
1.00	9.262	0.001	0.010	0.003	-0.004	0.002
1.10	8.312	0.000	0.008	0.000	-0.001	0.000
1.20	7.580	0.000	0.006	-0.001	0.000	-0.001
1.30	7.016	-0.001	0.004	-0.001	0.000	-0.001
1.40	6.574	-0.001	0.003	-0.001	0.000	-0.002
1.50	6.216	-0.001	0.001	-0.001	0.000	-0.001
1.60	5.913	0.000	0.001	-0.001	0.000	-0.001
1.70	5.645	0.000	0.001	-0.001	0.000	-0.001
1.80	5.398	0.000	0.000	-0.001	0.000	-0.001
1.90	5.162	0.000	0.003	0.001	0.000	-0.002
2.00	4.932	–	–	–	–	–

Table 12

Scattering Factor Differences Interpolated for Cesium(I)
from the Available Data at 0.02, 0.05, and 0.10
Intervals of s = (sin θ)/λ

s, Å^{-1}	f(s)	Cubic Spline	Stineman	-POLYNOMIALS- n = 3	n = 4	4-Point Lagrange
0.00	54.000	-	-	-	-	-
0.10	50.635	-0.001	-0.005	-0.001	0.000	-0.001
0.20	43.897	0.001	0.003	0.002	0.000	0.002
0.30	37.893	0.000	0.001	0.001	0.000	0.000
0.40	33.240	0.000	0.000	0.000	0.000	0.000
0.50	29.385	0.000	0.000	0.000	0.001	0.000
0.60	26.074	0.001	0.001	0.001	0.001	0.001
0.70	23.303	0.002	0.006	0.004	0.001	0.003
0.80	21.071	0.003	0.007	0.004	-0.002	0.005
0.90	19.309	0.000	0.016	0.001	-0.002	0.002
1.00	17.900	-0.001	0.009	-0.004	-0.003	-0.004
1.10	16.721	-0.002	0.005	-0.005	0.000	-0.005
1.20	15.676	-0.002	0.003	-0.005	0.000	-0.005
1.30	14.701	-0.002	-0.002	-0.004	0.000	-0.005
1.40	13.760	-0.001	-0.002	-0.002	0.001	-0.002
1.50	12.844	0.000	-0.001	-0.001	0.001	-0.001
1.60	11.956	-0.001	0.000	-0.001	0.000	-0.001
1.70	11.104	0.001	0.002	0.001	0.001	0.001
1.80	10.302	-0.002	0.001	0.000	0.000	0.000
1.90	9.559	0.012	0.033	0.000	0.001	0.002
2.00	8.882	-	-	-	-	-

Table 13

Sixth Degree Fit by Polft of Atomic Scattering
Factors for Each Neutral Atom from Hydrogen
(Z = 1) to Neon (Z = 10)

Element	Atomic Number Z	E, %	Standard Error of Estimate	Index of Determination
H	1	1.743	0.019	0.9976
He	2	1.032	0.022	0.9992
Li	3	1.614	0.052	0.9969
Be	4	2.123	0.091	0.9944
B	5	2.392	0.128	0.9933
C	6	2.213	0.142	0.9946
N	7	1.824	0.136	0.9966
O	8	1.392	0.119	0.9981
F	9	1.013	0.098	0.9990
Ne	10	0.714	0.076	0.9995

Table 14

Seventh Degree Fit by Polft of Atomic Scattering
Factors for Each Neutral Atom from Hydrogen
(Z = 1) to Neon (Z = 10)

Element	Atomic Number Z	E, %	Standard Error of Estimate	Index of Determination
H	1	0.737	0.008	0.9996
He	2	0.336	0.007	0.9999
Li	3	1.357	0.044	0.9978
Be	4	2.062	0.089	0.9948
B	5	1.892	0.102	0.9958
C	6	1.401	0.091	0.9979
N	7	0.920	0.070	0.9991
O	8	0.559	0.048	0.9997
F	9	0.322	0.031	0.9999
Ne	10	0.185	0.020	0.9999

Table 15

Eighth Degree Fit by Polft of Atomic Scattering
Factors for Each Neutral Atom from Hydrogen
(Z = 1) to Neon (Z = 10)

Element	Atomic Number Z	E, %	Standard Error of Estimate	Index of Determination
H	1	0.232	0.003	0.9999
He	2	0.143	0.003	0.9999
Li	3	1.354	0.044	0.9978
Be	4	1.678	0.073	0.9965
B	5	1.187	0.065	0.9984
C	6	0.679	0.044	0.9995
N	7	0.334	0.026	0.9999
O	8	0.166	0.015	0.9999
F	9	0.123	0.012	0.9999
Ne	10	0.120	0.013	0.9999

Table 16

Ninth Degree Fit by Polft of Atomic Scattering
Factors for Each Neutral Atom from Hydrogen
(Z = 1) to Neon (Z = 10)

Element	Atomic Number Z	E, %	Standard Error of Estimate	Index of Determination
H	1	0.185	0.002	0.9999
He	2	0.142	0.003	0.9999
Li	3	1.270	0.042	0.9981
Be	4	1.107	0.049	0.9985
B	5	0.568	0.031	0.9996
C	6	0.240	0.016	0.9999
N	7	0.139	0.011	0.9999
O	8	0.134	0.012	0.9999
F	9	0.123	0.012	0.9999
Ne	10	0.103	0.011	0.9999

Table 17

Orthogonal Polynomial Fit of f(s) versus s
for Bonded Hydrogen

	---- Full Data Set ---- K = 46 Data		-- Evenly Spaced Data -- K = 41 Data	
Order	Standard Error of Estimate for f(s)	Index of Determination	Standard Error of Estimate for f(s)	Index of Determination
1	0.2059	0.6503447	0.1907	0.5782925
2	0.0921	0.9316453	0.0913	0.9059052
3	0.0350	0.9903303	0.0342	0.9871070
4	0.0313	0.9924806	0.0284	0.9913723
5	0.0272	0.9944525	0.0257	0.9931082
6	0.0168	0.9979269	0.0165	0.9972487
7	0.0075	0.9995971	0.0075	0.9994444
8	0.0024	0.9999618	0.0024	0.9999473
9	0.001	0.9999845	0.0013	0.9999843
10	0.0014	0.9999865	0.0013	0.9999854
11	0.0010	0.9999937	0.0009	0.9999928

Table 18

Orthogonal Polynomial Fit of f(s) versus $\exp(-s^2)$
for Bonded Hydrogen

| | ---- Full Data Set ----
K = 46 Data | | -- Evenly Spaced Data --
K = 41 Data | |
Order	Standard Error of Estimate for f(s)	Index of Determination	Standard Error of Estimate for f(s)	Index of Determination
1	0.2117	0.6303771	0.1939	0.5641825
2	0.1262	0.8717222	0.1203	0.8363969
3	0.0742	0.9567019	0.0713	0.9440936
4	0.0430	0.9857708	0.0408	0.9821542
5	0.0246	0.9954810	0.0229	0.9945454
6	0.0139	0.9985967	0.0128	0.9983344
7	0.0077	0.9995808	0.0072	0.9994970
8	0.0042	0.9998781	0.0040	0.9998499
9	0.0023	0.9999652	0.0022	0.9999564
10	0.0012	0.9999901	0.0012	0.9999877
11	0.0007	0.9999971	0.0006	0.9999964

Table 19

Orthogonal Polynomial Fit of f(s) versus exp(-s)
for Bonded Hydrogen

	---- Full Data Set ---- K = 46 Data		-- Evenly Spaced Data -- K = 41 Data	
Order	Standard Error of Estimate for f(s)	Index of Determination	Standard Error of Estimate for f(s)	Index of Determination
1	0.1479	0.8195917	0.1420	0.7661639
2	0.0386	0.9880030	0.0373	0.9842798
3	0.0370	0.9892208	0.0339	0.9873841
4	0.0216	0.9964183	0.0210	0.9952695
5	0.0058	0.9997510	0.0058	0.9996440
6	0.0025	0.9999553	0.0021	0.9999538
7	0.0019	0.9999749	0.0017	0.9999711
8	0.0007	0.9999965	0.0007	0.9999956
9	0.0002	0.9999997	0.0002	0.9999996
10	0.0002	0.9999998	0.0001	0.9999999
11	0.0001	0.9999999	0.0001	0.9999999

CHAPTER 4

Educational Implications

It is important for a dissertation of this type to be relevant to the teaching of undergraduates. The purpose of the present chapter is to document how numerical analysis with computers applies to the undergraduate chemistry curriculum. A historical approach has been chosen to show how trends in science education have led to the current, ubiquitous use of computers as educational tools. Analog computers have been employed in engineering schools for several decades; however, the emphasis here is on technical developments in computer science and related mathematics since the early 1960's, and the topic is limited to more modern digital computers.

In 1939, John V. Atanasoff, a mathematics professor at Iowa State University, along with his assistant, Clifford Berry, designed and built the first electronic digital computer. The Atanasoff-Berry-Computer or "ABC" was given birth out of a need for rapid mathematical calculations by some twenty masters and doctoral candidates and their professor, Dr. Atanasoff. This invention set the stage for a revolution in the fields of business, science, technology, and education (143, 144, 160-162) and has become a subject of popular interest (163, 164).

71

The first UNIVersal Automatic Computer, UNIVAC I, was delivered to the United States Bureau of Census for use in the 1950 census. This event marked the first time that a stored-program computer dedicated to business was employed. Previous applications had been for scientific, military, or engineering data processing.

During the period 1958-1964, a myriad of machines by IBM, RCA, Sperry-Rand, and others, sometimes referred to as the "second-generation" of computers, were introduced into the market. A shortcoming of these second-generation computers was their mutually incompatible programming, being roughly classified as either scientific or business. For example, FORmula TRANslation or FORTRAN was developed in 1947 as a problem-oriented language for mathematical and scientific calculations, and COmmon Business-Oriented Language or COBOL was released for business applications in 1960.

The significance of these unprecedented developments was appreciated by the Mathematical Association of America, MAA, whose main objective is to promote an interest in mathematics, especially at the college level, and it began to up-date college libraries for the forthcoming computer revolution. In 1960, a series of grants from the National Science Foundation, NSF, enabled the MAA and its specially appointed Committee on the Undergraduate Programs in Mathematics, CUPM, to initiate a sustained effort of

providing critical information concerning the content of undergraduate mathematics textbooks (165).

During the post-Sputnik era, science educators of all areas were evaluating their individual subject matter and methods also, and the National Science Foundation was providing grants to improve the teaching of these subjects at the high school as well as the college level. In the summer of 1960, the NSF began funding of the Chemical Education Materials Study, CHEM Study, with Nobel Laureate Glenn T. Seaborg as its head. The first through third drafts of the CHEM Study textbook (166, 167) were tested in several high schools and junior colleges from 1960-1965. Both the Physical Science Study Committee (168), PSSC Physics, and the Earth Science Curriculum Project (169), ESCP Earth Science, released their textbooks at about the same time. These "new approaches" placed special emphasis on the relationship of unifying principles to the laboratory work associated with each science.

From 1963-1968, I worked with other science educators in the development and testing of laboratory modules for these courses. Our emphasis was upon the traditional values of the scientific method, and we sought to develop simple, inexpensive, and meaningful experiences for high school and first-year college students. These lab experiments were designed so that much data was collected, but data analysis was conducted by graphical methods because computers were

then unavailable for classroom use. Still, the widespread acceptance of this more analytical approach to the student laboratory was, in my opinion, an important factor in the readiness with which the academic community adopted the microcomputer upon its later arrival.

In Europe, the computer revolution was having an effect on the teaching of mathematics in British high schools and universities. In 1965, the Institute of Mathematics and its Applications sponsored a symposium entitled "The State of the Art in Numerical Analysis" held at Birmingham, England, at which specialists gave introductory lectures for novice users. These lectures were later edited by Joan Walsh, also one of the contributors, and published as a general survey of recent developments in numerical analysis, a chapter of which was devoted entirely to the increasing importance of numerical work in teaching at all levels (170). One of the speakers at this conference expressed the opinion that "desk-calculating machines" had now become so commonplace in British schools that some numerical analysis was actually being taught, but not in an organized or even recognizable manner. Accordingly, a study group of the Committee for Mathematics in Education and Industry presented a syllabus for pre-university students suggesting that numerical analysis be included as an integral part of the mathematics curriculum and be applied with calculating machines whenever possible. Several other contributors proposed syllabi for

undergraduate courses in computing and numerical methods, some in pure mathematics and some for science and engineering students. It was recommended that computer programming be taught at the university level as a college major, since prospective employers would have neither the time nor the inclination to teach it on the job. As late as 1975, however, twenty-one British chemical educators, in a discussion of curricular changes and new techniques in teaching chemistry, made an assessment of the state of chemical education in the British Isles without a single reference to computers (171). Thus, a large gap must have existed between classroom lectures and the state of the art in computational chemistry.

The third generation of computers, those using chips rather than transistors or diodes, was ushered into the world in April of 1964 with the announcement of the System/360 computer by IBM. The "360" notation indicates that this computer system can be used for all types of processing, encircling both business and science. Familiar terms like "batch-processing" and "time-sharing" came into existence during this period. In 1964, John Kemeny developed Beginners All-purpose Symbolic Instructional Code, BASIC, so that students at Dartmouth College could use the time-sharing terminals. The software libraries of most university computer centers, including that of Middle Tennessee State University, contain the Dartmouth

Time-Sharing-System Programs Library that are written in BASIC and encompass a whole series of mathematical, statistical, and business procedures designed mostly for small sets of data, as well as for both demonstrations and tutorials (142).

This new emphasis in science and mathematics courses, coupled with the simultaneous growth of a computer industry, brought on a deluge of books in many different areas of applied mathematics, which caused CUPM's Basic Library List of 1965 to become outdated (165). In 1974, the NSF provided funds for an updated list, which was published in 1976 and increased from 300 titles to approximately 700 separate titles (172). Many of the books on the original list were either out of print or outdated and were replaced. New headings entitled "numerical analysis," "computation," and "mathematical programming" were created to include prominent new books (69, 70, 79, 118, 126, 127, 173) on approximation, computer programming, interpolation, and the application of these areas to many disciplines including the physical sciences.

The gradual introduction of the computer into research and education closely parallels the development and marketing of an inexpensive, user-friendly microcomputer. Early use of the computer during the 1950's and 1960's was limited because of the initial cost of the hardware and unfamiliarity on the part of potential users; however, the

period 1964-1974 saw the birth of a software industry and the emergence of the microcomputer.

In 1974, Commodore startled the world by presenting the Personal Electronic Transactor, PET, for about six hundred dollars including the chip for BASIC. Two years later, Steve Wozniak and Steven Jobs, working for Hewlett-Packard and Atari, respectively, assembled the original Apple I computer in Wozniak's parents' garage, and Apple Computer, Inc. delivered its first Apple II computer in May of 1977. By the end of the year, Apple, Commodore, and Radio Shack had a complete line of computers on the market. In 1981, IBM introduced the IBM PC, which is regarded by many as the best general usage computer. This was followed by several IBM "clones," including the Kaypro-PC discussed in Chapter 2. It was now apparent that the microcomputer was having a great and lasting impact on homes, schools, business, and research.

The need for routine treatment of data, especially in the physical chemistry laboratory (129, 174), has been recognized for many years; however, the introduction of digital computers to undergraduate chemistry majors would seem to require some knowledge of either FORTRAN compiler language (175-178) for mainframe computers or BASIC Interpreter language (129, 143, 144, 174) for microcomputers as a prerequisite. The most common example of curve fitting in laboratory courses is drawing the best straight line

through data points when a linear relationship between two variables is suspected, and least-squares fitting applied as a simple linear regression is an obvious choice (139, 177, 179, 180). The method of least squares may be extended by multiple regression to data having more than two variables (76, 181) and also to functional relationships between two or more variables that are not linear. In the latter case, a linear regression can sometimes be used to find the coefficients of a polynomial, logarithmic, or other linear transformation of the original equation; however, an <u>iterative</u> least-squares method may be more appropriate if the equation is nonlinear in the coefficients and cannot be recast into a linear form (129, 174, 176, 182). Owing to the frequent occurrence of nonlinear equations in chemistry, much emphasis has been placed on their least-square fitting in the recent literature of chemical education (183-189).

The Clausius-Clapeyron equation provides an interesting example involving a nonlinear least-squares analysis of data. Physical chemistry textbooks and laboratory manuals usually include an experiment in which the vapor pressure of a pure liquid is measured as a function of temperature. From this measurement, the heat of vaporization, ΔH_{vap}, is calculated. In cases where the heat of vaporization does not change with temperature, the Clausius-Clapeyron equation may be written in a linear form:

$$\log P = A - 0.4342 \, \Delta H_{vap}/RT$$

This equation is linear with respect to $\log P$ and $1/T$, where A is the y-intercept, $-0.4342 \, \Delta H_{vap}/R$ is the slope of the line, R is the ideal gas constant, P is the vapor pressure in atmospheres, and T is the absolute temperature (190).

It has been suggested by Pollnow (191, 192) that the observed values of the vapor pressure, rather than $\log P$, should give a more accurate least-squares fit against the reciprocal of the temperature since P is the quantity that is measured directly; however, this form of the equation is not linear with respect to the y-intercept and the slope. Pollnow suggests that a Taylor series be used to approximate the y-intercept and the slope, which are then refined using an _iterative_ least-squares technique until neither of the coefficients show any significant change. In a recent supplementary textbook designed to apply microcomputer methods in the physical chemistry laboratory, Noggle (129) proposes that a four-point Lagrange or other interpolation method involving cubic polynomials would be simpler, faster, and more accurate when applied to the Clausius-Clapeyron equation. Interpolation should give good results with fairly precise data, but spurious interpolated values would be expected when the data contains large, random errors. In the latter case, the iterative least-squares fitting method is preferred.

Many other examples of the curve fitting of empirical data may be found in the literature of chemical education (193-195), including applications for specific brands of computers (185, 196-198). The impact of the microcomputer on all areas of chemical education is the subject of a recent review (199). Computers have affected science and education in many obvious ways; however, the effects are often hidden. The Chemical Abstracts Service data base, CAS ONLINE, and other similar information retrieval systems are good examples where little computer knowledge is needed, and many modern instruments require no knowledge of programming, even though they are completely computer operated. Other instruments, such as the Varian 3600 gas chromatograph and the Perkin-Elmer 983 infrared spectrophotometer, require only a bare minimum of programming ability in a simple language, usually BASIC or some version of BASIC.

Computer technology has become so commonplace that chemical educators have been forced to produce a new generation of students who are not only chemists, but are also literate in the area of computers. Many schools are now requiring computer programming as a prerequisite to physical chemistry, usually in the junior or senior year of a chemistry major. This ranges from a one-hour laboratory briefing on FORTRAN at the University of Missouri (200) to a full course entitled "Computer Applications in Chemistry" at Penn State University (201).

The course at Penn State University, which has been offered for over ten years, is divided into three major topic areas: an introduction, which uses 15% of the course time; numerical topics, 50% of the course; and nonnumerical topics, 35% of the time.

The course begins with several books and magazine articles being assigned as optional reading. These deal mainly with the impact of the computer on society, the integral organization of the computer system, and the specification of hardware capability. Discussion of these topics is followed by an overview of scientific computer usage to include number crunching, time-limited problems, computer optimization, storage and retrieval of information, experiment management and control, intelligent problem solving, and modeling and simulation. The remaining topics in the introduction involve the basics of writing computer programs and statistical analysis.

The numerical methods section of the course consumes the largest block of time. Its purpose is to cover as many topics as possible, in just enough detail, to enable the student to pursue areas of interest on his own. The first block of material deals with algorithms and their goodness of fit. The requirements of iterative methods are also discussed: how to make a good first guess, how to improve on previous guesses, and when to stop the iteration after sufficient accuracy is reached. This is followed by a

discussion of the various kinds of approximating functions available. The main point here is that applied numerical analysis can be based upon many different functional forms. A homework assignment is given that develops a potential energy surface for atom-transfer reactions. The largest amount of time is devoted to curve fitting, starting with a simple linear fit without weighting and proceeding to one where the points are weighted. This introduction to curve fitting is followed by higher order polynomial fits using matrix notation. A second homework problem involves the fitting of kinetic data from an enzyme reaction to derive its Michaelis-Menten constants. Also covered are iterative least-squares fitting, interpolation methods including spline fitting, Fourier transform, integration, differential equations, and the literature of research applications.

Nonnumerical topics are covered in the concluding section of the course. These include information handling capabilities, molecular mechanics and modeling, and artificial intelligence. The course is constantly updated, and reading lists are provided and revised periodically since no single textbook is considered adequate. About 25-30 students in the fields of chemistry, biochemistry, and computer science enroll each semester, and their response has been positive. In fact, the course is often recommended to others by students who have taken it.

The previous discussion concerns a specific course offered at Penn State University; however, other similar courses are taught at Purdue University (202, 203), Trinity University (204), Loyola College (205), Creighton University (206), Case Western Reserve University (132), and elsewhere (123, 207, 208). Numerous other references, besides those previously cited, are available on such subjects as data handling, statistics, mathematical and computational methods, and computer and electronic calculator programming, all of which would be relevant for a chemistry course of this type (114, 123, 209-218).

CHAPTER 5

Conclusions

A review of the literature involving previous use of analytical approximations for describing the shapes of atomic scattering curves of X-ray diffraction was given in Chapter 1. This was to document that a finite set of X-ray scattering factors taken from the scattering curve of any particular atomic species are, indeed, related by the coefficients of either a Gaussian expansion or a polynomial series, and we have explored the less studied application of polynomials for both curve fitting and interpolating such data in Chapter 3. The following conclusions of this study will be of general interest to practicing X-ray crystallographers and may contribute to an improvement in methods for the refinement of X-ray crystal-structure determinations.

1. Interpolation of X-ray scattering factors equally spaced at 0.10 intervals of s is accurate within ± 0.002 only for values of $f(s)$ above 0.6 or 0.7 in the x-argument. Inclusion of 0.05 intervals in the x-argument range of $0.00 \leq s \leq 0.70$ affords little improvement in the interpolation of X-ray scattering factors.

2. The grid spacing in the x-argument range of $0.50 \leq s \leq 1.50$ commonly employed in standard tabulations of X-ray scattering factors is too coarse for interpolation of $f(s)$ within ± 0.002 near $s = 1.00$ unless an interpolating polynomial of sixth or higher degree is employed.

84

3. A third-degree polynomial would accurately
   interpolate values of f(s) if a grid spacing of
   0.05 (or 0.02) intervals of s were extended to
   about 1.50 in the x-argument.

4. Cubic spline interpolation of X-ray scattering
   factors is generally superior to interpolating
   polynomials of the third degree.

5. Interpolation of X-ray scattering factors by
   the Stineman method is generally inferior to
   interpolating polynomials of the third degree.

6. As expected, the interpolating polynomial of
   third degree and the four-point Lagrange
   perform identically within roundoff error.

7. An orthogonal polynomial series provides a slightly
   better linear least-squares fit than does a non-
   orthogonal one, but the fitted X-ray scattering
   curve is still too nonlinear in the coefficients
   to permit true orthogonality. A nonlinear norm
   is required for the accurate analytical approx-
   imation of an X-ray scattering curve with either
   a Gaussian expansion or a polynomial series alone.

8. The best analytical approximation of an X-ray
   scattering curve involves a combination of
   Gaussian and polynomial terms that maximizes
   the best features of both and minimizes the
   worst. One is suggested that gives an accurate
   linear least-squares fit for the X-ray scattering
   factors of bonded hydrogen.

9. Numerical analysis has the potential for providing
   a unique methodology to uncover mathematical rela-
   tionships both within and between sets of X-ray
   scattering factors that could be applied as
   practical criteria to test their individual
   accuracy. It would be especially important if
   these criteria were absolute, being based upon
   self-evident principles that presuppose no assumpt-
   ions and are not subject to empirical uncertainty.
   Such criteria, if simple and absolute, could serve
   as reliable and efficient guides to check the
   accuracy of existing tables of scattering factors,
   to test new ones for improvement, and to provide
   similiar comparisons between the different types
   of wave functions from which the X-ray scattering
   factors are ultimately derived.

The final conclusion needs some further elaboration. The a priori calculation of X-ray scattering factors is an application of advanced mathematics to complex physical theory, but the values obtained from different levels of theory and mathematical approximation are not subject to any straightforward testing of their absolute accuracy. Nevertheless, indirect evidence supports the growing awareness that improper treatment of valence electrons by the scattering factors can introduce significant systematic error into an otherwise highly refined crystal-structure determination. This is certainly true for ionic crystal structures, where the current practice of using scattering factors calculated from the wave functions of isolated ions is seriously flawed because this completely ignores the influence of counter ions on the distribution of electron density (219-221). In the present dissertation, we have sought to lay a basis for our continuing work (223) in this area by establishing some of the boundaries within which these various methods of numerical analysis may be applied with appropriateness and validity.

The current project made use of a Honeywell DPS 8/49 mainframe computer, a Perkin-Elmer 3600 Data Station, and the Apple II+, the IBM PC, and the Kaypro-PC microcomputer. Several simple programs were written in BASIC, and other commercial and public domain programs were employed. The appropriateness of these computer programs for use by

students was demonstrated by the fact that virtually all of the various calculations of this project have involved the supervised participation of undergraduates at both Middle Tennessee State University and Campbellsville College. In fact, most of our calculations during the summer of 1986 were actually conducted by a high school student having no previous computer experience. One of the undergraduates, a college senior, presented a group paper at a regional student meeting (222). The same slides and much of the content of the student talk were equally appropriate for presentations given at professional meetings (223-225). This is because the nature of crystals and the electronic structure of atomic species are not only sophisticated topics at the frontiers of chemical research, but are also basic subjects pursued with interest and understanding by even young science students at elementary levels (226). Most areas of the physical sciences are like that, being understandable at different levels while always bearing a relationship to a few fundamental concepts.

The application of this research project to teaching chemistry was addressed in the preceeding chapter, devoted entirely to the relevance of numerical analysis as a topic in the undergraduate chemistry curriculum. The recent evolution of the low cost microcomputer has been one of the most significant innovations for chemical education since the introduction of the Bunsen burner; and it has required

that chemical educators now produce a new breed of graduates that are computer literate. The main areas of computer use in education today include not only those of instruction and demonstrations, but also the collection and analysis of student laboratory data as well. Data-handling programs, including curve fitting, allow a quicker and more detailed analysis of undergraduate laboratory assignments by both students and professors. Other software permits the chemistry department with a small budget to have simulated instrumentation.

I have recently purchased and employed such programs for a new course in instrumental analysis offered at Campbellsville College. In the Spring of 1984, I designed and wrote the syllabus (Appendix III) for this course, Chemistry 350 entitled "Instrumental Analysis," as an upper division course for both the minor and major in chemistry. The course uses Skoog's Principles of Instrumental Analysis as the textbook (227) together with the supplementary Explorations in Chemistry by Allendoerfer (228) including its accompanying disc, whose interactive programs contain fourteen different types of chemistry problems. The student also has the option of printing a set of problems, along with the answers, for homework. In one experiment, data taken from a Perkin-Elmer UV-vis spectrophotometer is fitted using Curve Fitter-PC (150). Permanent plots are then made and incorporated into the laboratory write-up.

Another topic of Chemistry 350 is NMR spectroscopy. In the small college, it is unrealistic to purchase expensive instrumentation; however, in January of 1986, I purchased the program NMR Simulator (229), which was written by Paul F. Schatz, to acquaint students with the use of the Varian EM-360A nuclear magnetic resonance spectrometer without my actually owning such an instrument. The simulated computer controls and adjustments closely mimic those on the actual instrument, as does the appearance of the spectrum obtained from these settings. The program allows me to present samples as unknowns, randomly vary the sample concentration, or vary the amount of "background noise" to enhance the realism of the simulation. The spectrum can be printed along with the selected instrument settings on any one of several printers compatible with the program. The program is versatile in that new data discs can be purchased at a low cost once the initial program package is bought. The response of students has been favorable to this course, as well as to the associated tutorials, the study aids, and the simulated instrumentation.

My involvement in the research area of the present dissertation has given me, a chemical educator at a small liberal arts college, a new outlook on the future of chemical education. Instrumentation once too expensive, interactive study supplements for both the gifted and the poorly prepared student once unavailable, and data analysis

techniques once too time consuming are now plentiful in the form of computer programs. Consequently, my participation in this project has provided a key that unlocks the door to these opportunities.

APPENDIX I

Tenth Degree, a Program in BASIC for the Calculation of f(x) = y at any x from the Coefficients of an Algebraic Polynomial up to the Tenth Degree

```
100 REM THIS PROGRAM WILL CALC Y FROM THE COEF OF A 10TH
105 REM DEGREE POLYNOMIAL
110 PRINT "WHAT IS THE VALUE OF X?"
120 INPUT "X"; X
122 INPUT "PLACE A VALUE FOR COM TO COMPARE"; C1
130 LPRINT "THE VALUE OF SIN THETA/LAMBDA TO BE CALCULATED"
131 LPRINT "IS = "; X
133 IF C1 = 1 THEN 360
140 INPUT "A"; A
150 LPRINT "A = " ; A
160 INPUT "B"; B
170 LPRINT "B = "; B
180 INPUT "C"; C
190 LPRINT "C = "; C
200 INPUT "D"; D
210 LPRINT "D = "; D
220 INPUT "E"; E
230 LPRINT "E = "; E
240 INPUT "F"; F
250 LPRINT "F = "; F
260 INPUT "G"; G
270 LPRINT "G = "; G
280 INPUT "H"; H
290 LPRINT "H = "; H
300 INPUT "I"; I
310 LPRINT "I = "; I
320 INPUT "J"; J
330 LPRINT "J = "; J
340 INPUT "K"; K
350 LPRINT "K = "; K
360 LET Q = A + B*X + C*X^2 + D*X^3
370 LET R = E*X^4 + F*X^5 + G*X^6
380 LET Z = H*X^7 + I*X^8 + J*X^9 + K*X^10
390 LET Y = Q + R + Z
400 LPRINT "Y = "; Y
405 GOTO 110
410 END
```

APPENDIX II

Lagi, a Program in BASIC for Lagrange Interpolation[a]

```
10 REM PROGRAM TO CALCULATE THE LAGRANGE POLYNOMIAL
20 DIM XT(25), FT(25)
50 INPUT "ENTER THE DEGREE OF THE POLYNOMIAL"; N
51 LPRINT "THE DEGREE OF THE POLYNOMIAL IS"; N
60 N1 = N
70 LPRINT "      XT(I)          FT(I)      "
80 FOR I = 1 TO N1
90 READ XT(I), FT(I)
100 NEXT I
120 INPUT "ENTER A VALUE FOR X"; X
130 FX = 0
140 FOR I = 1 TO N1
150 P = 1
155 FOR J = 1 TO N1
160 IF J<>I THEN P=P*((X-XT(J))/(XT(I)-XT(J)))
180 NEXT J
190 FX = FX + P*FT(I)
200 NEXT I
210 LPRINT "X IS  "; X; TAB(20); "F(X) = "; FX
211 LPRINT "  "
212 FOR I = 1 TO N1
213 LPRINT XT(I); "      "; FT(I)
214 NEXT I
220 GOTO 120
230 DATA 60,156,50,135,80,204
231 FOR I = 1 TO N1
232 LPRINT XT(I); "     "; FT(I)
233 NEXT I
240 END
```

---

[a]Sample data from Daniels (146) has been included in line 230 in the form $x_1, y_1, x_2, y_2, x_3, y_3$.

APPENDIX III

Syllabus for Chemistry 350, Instrumental Analysis, at
Campbellsville College, Campbellsville, Kentucky

CAMPBELLSVILLE COLLEGE

Natural Science Division

Course Syllabus

I.   TITLE: Chemistry 350, Instrumental Analysis,
     Three Semester Hours Credit

II. DESCRIPTION: Instrumental Analysis is a study of the
     modern chemistry laboratory involving sophisticated
     instrumentation and includes spectroscopy, gas
     chromatography, atomic absorption, and other methods.

III. TEXT: Skoog, D. A. Principles of Instrumental
     Analysis, 3rd ed.; Saunders College Publishing:
     Philadelphia, PA, 1985.

IV. PURPOSE: The purpose of this course is to provide the
     minor and major in chemistry with an introduction to
     instrumental methods.

V.   COURSE OUTLINE: Chemistry 350, Instrumental Analysis

     1. Introduction to Chemical Instrumentation
         A. Electroanalytical Methods
         B. Spectroscopic Methods
         C. Chromatographic Methods

     2. Gas Chromatography
         A. Theoretical Plate Calculation
             1. Temperature Effects
             2. Flow Rate Effects
             3. Qualitative Analysis
             4. Quantitative Analysis
         B. Computer-Simulated Instruction
             (When Available)

     3. Infrared Spectroscopy
         A. Theory
         B. Quantitative Analysis--The determination of
             the percentage composition of xylenes

            C.   Qualitative Analysis
            D.   Sample Preparation

       4.   UV-Vis Spectroscopy
            A.   Theory
            B.   Basic Instrumentation
            C.   Using CURVE FITTER to construct a calibration curve

       5.   Miscellaneous Methods
            A.   Atomic Absorption
            B.   Nuclear Magnetic Resonance
            C.   Polarography
            D.   Others as time permits

VI.   COMPETENCIES:

    1.   Students will demonstrate a knowledge of the theory of modern instrumentation.

    2.   The student will demonstrate the use of modern instrumentation.

    3.   The student will learn the techniques of good precision and accuracy in their measurements.

    4.   Students will demonstrate interpretation of infrared, UV-vis, and NMR spectra and the analysis of a chromatogram.

VII.   METHODS: The lecture method will be used, as well as hands-on laboratory experiences in instrumental methods. Computer simulation will be used extensively.

VIII.   EVALUATION: Four regular exams will be given on the theory and methods. The final examination will be comprehensive. Laboratory analysis and write-ups will count one-third of the final grade.

IX.   READINGS AND SUPPLEMENTARY MATERIALS:

Silverstein, R. M.; Bassler, G. C.; Morrill, T. C. Spectrometric Identification of Organic Compounds, 3rd ed.; John Wiley: New York, 1974.

Williams, D. H.; Fleming, I. Spectroscopic Methods in Organic Chemistry, 2nd ed.; McGraw-Hill: London, 1973.

Shapiro, R. H.; Depuy, C. H. Exercises in Organic
Spectroscopy, 2nd ed.; Holt, Rinehart and Winston:
New York, 1977.

Yost, R. W.; Ettre, L. S.; Conlon, R. D. Practical
Liquid Chromatography: An Introduction;
Perkin-Elmer Corporation, Chromatography Division:
Norwalk, CT, 1980.

Allendoerfer, R. D. Explorations in Chemistry;
Prentice-Hall: Englewood Cliffs, NJ, 1985.

Schatz, P. F. NMR Simulator for the IBM PC; City
Software: Milwaukee, WI, 1985.

LITERATURE CITED

1. Partington, J. R. An Advanced Treatise on Physical Chemistry, Vol. 3: The Properties of Solids; Longmans, Green: London, 1952.

2. Clarke, L. J. Surface Crystallography: An Introduction to Low Energy Electron Diffraction; Wiley-Interscience: Chichester, England, 1985.

3. West, A. R. Solid State Chemistry and Its Applications; John Wiley: Chichester, England, 1984.

4. Hazen, R. M.; Finger, L. W. Comparative Crystal Chemistry: Temperature, Pressure, Composition, and the Variation of Crystal Structure; Wiley-Interscience: Chichester, England, 1982.

5. Structure and Bonding in Crystals; O'Keeffe, M.; Navrotsky, A., Eds.; Academic Press: New York, 1981; 2 vols.

6. Treatise on Solid State Chemistry; Hannay, N. B., Ed.; Plenum Press: New York, 1976; 7 vols.

7. Bondi, A. Physical Properties of Molecular Crystals, Liquids, and Glasses; John Wiley: New York, 1968.

8. (a) Kitaigorodsky, A. I. Molecular Crystals and Molecules; Academic Press: New York, 1973.
(b) Kitaigorodskiy, A. I. Order and Disorder in the World of Atoms; Springer-Verlag: New York, 1967.
(c) Kitaigorodskii, A. I. Organic Chemical Crystallography; Consultants Bureau: New York, 1961.

9. Engel, P. Geometric Crystallography: An Axiomatic Introduction to Crystallography; D. Reidel: Dordrecht, Holland, 1986.

10. Smith, J. V. Geometrical and Structural Crystallography; John Wiley: New York, 1982.

11. Verma, A. R.; Srivastava, O. N. Crystallography for Solid State Physics; John Wiley: New York, 1982.

12. Vainshtein, B. K. Modern Crystallography I: Symmetry of Crystals, Methods of Structural Crystallography; Springer-Verlag: Berlin, 1981.

96

13. (a) Buerger, M. J. Elementary Crystallography: An Introduction to the Fundamental Geometrical Features of Crystals; Wiley: New York, 1963. (b) Phillips, F. C. An Introduction to Crystallography, 3rd ed.; John Wiley: New York, 1963.

14. (a) International Tables for Crystallography, Vol A: Space-Group Symmetry; Hahn, T., Ed.; D. Reidel: Dordrecht, Holland, 1987. (b) International Tables for X-ray Crystallography, Vol. 1: Symmetry Groups; Henry, N. F. M.; Lonsdale, K., Eds.; Kynoch Press: Birmingham, England, 1965.

15. McKie, D.; McKie, C. Essentials of Crystallography; Blackwell Scientific: Oxford, United Kingdom, 1986.

16. Glusker, J. P.; Trueblood, K. N. Crystal Structure Analysis: A Primer, 2nd ed.; Oxford University Press: New York, 1985.

17. (a) Luger, P. Modern X-ray Analysis on Single Crystals; Walter de Gruyter: Berlin, 1980. (b) Glasser, L. S. D. Crystallography and its Applications; Van Nostrand-Reinhold: Workingham, Berkshire, England, 1977.

18. (a) Buerger, M. J. Contemporary Crystallography; McGraw-Hill: New York, 1970. (b) Buerger, M. J. Crystal-Structure Analysis; John Wiley: New York, 1960.

19. Stout, G. H.; Jensen, L. H. X-ray Structure Determination: A Practical Guide; Macmillan: New York, 1968.

20. Lipson, H.; Cochran, W. The Crystalline State, Vol. III: The Determination of Crystal Structures; Cornell University Press: Ithaca, NY, 1966.

21. (a) Ladd, M. F. C. Structure and Bonding in Solid-State Chemistry; John Wiley: Chichester, England, 1986. (b) Ladd, M. F. C.; Palmer, R. A. Structure Determination by X-ray Crystallography; Plenum Press: New York, 1977.

22. Vainshtein, B. K.; Fridkin, V. M.; Indenblom, V. L. Modern Crystallography II: Structure of Crystals; Springer-Verlag: Berlin, 1982.

23. Dunitz, J. D. X-ray Analysis and the Structure of Organic Molecules; Cornell University Press: Ithaca, NY, 1979.

24. Tollenaere, J. P.; Moereels, H.; Raymaekers, L. A. Janssen Research Foundation Series, Vol. 1: Atlas of the Three-Dimensional Structure of Drugs; Elsevier North-Holland: New York, 1979.

25. (a) Sherwood, D. Crystals, X-rays, and Proteins; John Wiley: New York, 1976. (b) Blundell, T. L.; Johnson, L. N. Protein Crystallography; Academic Press: New York, 1976.

26. (a) Adams, D. M. Inorganic Solids: An Introduction to Concepts in Solid-State Structural Chemistry; John Wiley: London, 1974. (b) Krebs, H.; Walter, P. H. L. Fundamentals of Inorganic Crystal Chemistry; McGraw-Hill: London, 1968.

27. (a) Mills, H. H.; Speakman, J. C. Prog. Stereochem. 1969, 4, 273-298. (b) Speakman, J. C. ibid. 1958, 2, 1-38. (c) Robertson, J. M. Organic Crystals and Molecules; Cornell University Press: Ithaca, NY, 1953.

28. (a) Jenkins, F. A.; White, H. E. Fundamentals of Optics, 4th ed.; McGraw-Hill: New York, 1976. (b) Cowley, J. M. Diffraction Physics; Elsevier North-Holland: Amsterdam, 1975.

29. James, R. W. The Crystalline State, Vol. II: The Optical Principles of the Diffraction of X-rays; Cornell University Press: Ithaca, NY, 1965. Reprinted by Ox Bow Press: Woodbridge, CT.

30. Woolfson, M. M. An Introduction to X-ray Crystallography; Cambridge University Press: London, 1970.

31. Guinier, A. X-ray Diffraction: In Crystals, Imperfect Crystals, and Amorphous Bodies; W. H. Freeman: San Francisco, 1963.

32. Warren, B. E. X-ray Diffraction; Addison-Wesley: Reading, MA, 1969.

33. (a) Bragg, L. The Crystalline State, Vol. I: A General Survey; Cornell University Press: Ithaca, NY, 1949. (b) Bragg, W. L. The Development of X-ray Analysis; Phillips, D. C.; Lipson, H., Eds.; G. Bell: London, 1975.

34. Compton, A. H.; Allison, S. K. X-rays in Theory and Experiment, 2nd ed.; Van Nostrand: Princeton, NJ, 1954.

35. Azaroff, L. V.; Kaplow, R.; Kato, N.; Weiss, R. J.; Wilson, A. J. C.; Young, R. A. X-ray Diffraction; McGraw-Hill: New York, 1974; Chapter 1.

36. Hubbell, J. H.; Veigele, W. J.; Briggs, E. A.; Brown, R. T.; Cromer, D. T.; Howerton, R. J. J. Phys. Chem. Ref. Data 1975, 4, 471-538.

37. Azaroff, L. V. Elements of X-ray Crystallography; McGraw-Hill: New York, 1968.

38. (a) Benesch, R.; Malli, G. Int. J. Quantum Chem. 1968, 2, 371-389. (b) Bonham, R. A. NATO Adv. Study Inst. Ser., Ser. B 1980, B48, 743-755.

39. (a) James, R. W.; Brindley, G. W. Phil. Mag., Ser. 7 1931, 12, 81-112. (b) James, R. W.; Brindley, G. W. ibid. 1931, 12, 729.

40. (a) Cochran, W. Acta Crystallogr. 1953, 6, 812. (b) Bacon, G. E. Acta Crystallogr. 1952, 5, 492-499.

41. (a) Hoerni, J. A.; Ibers, J. A. Acta Crystallogr. 1954, 7, 744-746. (b) Viervoll, H.; Ögrim, O. Acta Crystallogr. 1949, 2, 277-279.

42. (a) McWeeny, R. Acta Crystallogr. 1951, 4, 513-519. (b) McWeeny, R. ibid. 1952, 5, 463-468. (c) McWeeny, R. ibid. 1953, 6, 631-637. (d) McWeeny, R. ibid. 1954, 7, 180-186.

43. Levine, I. N. Physical Chemistry, 2nd ed.; McGraw-Hill: New York, 1983.

44. International Tables for X-ray Crystallography, Vol. 4: Revised and Supplementary Tables; Ibers, J. A.; Hamilton, W. A., Eds.; Kynoch Press: Birmingham, England, 1974.

45. International Tables for X-ray Crystallography, Vol. 3: Physical and Chemical Tables; MacGillavry, C. H.; Rieck, G. D., Eds.; Kynoch Press: Birmingham, England, 1968.

46. Hubbell, J. H.; Øverbø, I. J. Phys. Chem. Ref. Data 1979, 8, 69-105.

47. Doyle, P. A.; Turner, P. S. Acta Crystallogr., Sect. A: Found. Crystallogr. 1968, 24, 390-397.

48. Doyle, P. A.; Turner, P. S. Acta Crystallogr. 1967, 22, 153.

49. (a) Cromer, D. T.; Waber, J. T. Acta Crystallogr. 1965, 18, 104-109. (b) Cromer, D. T. Acta Crystallogr. 1965, 19, 224-227.

50. Lie, G. C. J. Phys. Soc. Jpn. 1977, 42, 1327-1333.

51. (a) Yusaf, M. S.; Lawes, G. P.; March, N. H. Int. J. Quantum Chem. 1980, 17, 833-844. (b) Enstrom, J. E. Phys. Rev. A 1973, 7, 979-983. (c) Bandopadhyay, P. K. J. Comput. Phys. 1973, 12, 280-284.

52. (a) Van der Wal, R. J.; Stewart, R. F. Acta Crystallogr., Sect. A: Found. Crystallogr. 1984, 40, 587-593. (b) Van der Wal, R. J.; Stewart, R. F. ibid. 1983, 39, 422-424.

53. (a) Deutsch, M.; Hart, M. Phys. Rev. B: Condens. Matter 1985, 31, 3846-3858. (b) Takama, T.; Kobayashi, K.; Hyugaji, M.; Nittono, O.; Sato, S. Jpn. J. Appl. Phys., Part I 1984, 23, 11-14. (c) Hansen, N. K.; Coppens, P. Acta Crystallogr., Sect. A: Found. Crystallogr. 1978, 34, 909-921.

54. Stewart, R. F. Acta Crystallogr., Sect. A: Found. Crystallogr. 1973, 29, 602-605.

55. (a) Sakurai, T. Rep. Comput. Cent., Univ. Tokyo 1970, 3(1-2), 19-39. (b) Coppens, P. Acta Crystallogr., Sect. A: Found. Crystallogr. 1969, 25, 180-186.

56. (a) Goscinski, O.; Lindner, P. J. Chem. Phys. 1970, 52, 2539-2540. (b) Smith, V. H. Chem. Phys. Lett. 1970, 7, 226-228.

57. Methods and Applications in Crystallographic Computing; Hall, S. R.; Ashida, T., Eds.; Clarendon Press: Oxford, 1984.

58. Computational Crystallography; Sayre, D., Ed.; Clarendon Press: Oxford, 1982. See especially: Stewart, J. M., pp 497-505.

59. (a) Prince, E. Mathematical Techniques in Crystallography and Materials Science; Springer-Verlag: New York, 1982. (b) Ramaseshan, S.; Richardson, M. F.; Wilson, A. J. C. Crystallographic Statistics; Indian Academy of Sciences: Bangalore, 1982.

60. (a) Crystallographic Computing Techniques; Ahmed, F. R.; Huml, K.; Sedlacek, B., Eds.; Munksgaard: Copenhagen, 1976. (b) Crystallographic Computing; Ahmed, F. R., Ed.; Munksgaard: Copenhagen, 1970.

61. Rollett, J. S. Computing Methods in Crystallography; Pergamon Press: Oxford, 1965.

62. Vand, V.; Eiland, P. F.; Pepinsky, R. Acta Crystallogr. 1957, 10, 303-306.

63. Freeman, H. C.; Smith, J. E. W. L. Acta Crystallogr. 1958. 11, 819-822.

64. Sparks, R. A.; Prosen, R. J.; Kruse, F. H.; Trueblood, K. N. Acta Crystallogr. 1956, 9, 350-358.

65. Wright, A. C. Clays Clay Miner. 1973, 21, 489-490.

66. Isaacs, N. In ref. 57, pp 193-205.

67. Isaacs, N. In ref. 58, pp 381-397.

68. Marsh, R. E.; Slagle, K. J. Acta Crystallogr., Sect A: Found. Crystallogr. 1983, 39, 173.

69. Hamming, R. W. Numerical Methods for Scientists and Engineers, 2nd ed.; McGraw-Hill: New York, 1973.

70. Hildebrand, F. B. Introduction to Numerical Analysis, 2nd ed.; McGraw-Hill: New York, 1974.

71. Forsythe, G. E.; Malcolm, M. A.; Moler, C. B. Computer Methods for Mathematical Computations; Prentice-Hall: Englewood Cliffs, NJ, 1977.

72. Rice, J. R. Numerical Methods, Software, and Analysis; McGraw-Hill: New York, 1983.

73. Davis, P. J. Interpolation and Approximation; Blaisdell: New York, 1963.

74. Phillips, G. M.; Taylor, P. J. Theory and Applications of Numerical Analysis; Academic Press: London, 1973.

75. Dodes, I. A. Numerical Analysis for Computer Science; Elsevier North-Holland: New York, 1978.

76. Green, P. E. Analyzing Multivariate Data; The Dryden Press: Hinsdale, IL, 1978

77. Draper, N. R.; Smith, H. Applied Regression Analysis; John Wiley: New York, 1966.

78. Daniels, R. W. An Introduction to Numerical Methods and Optimization Techniques; Elsevier North-Holland: New York, 1978.

79. Kelly, L. G. Handbook of Numerical Methods and Applications; Addison-Wesley: Reading, MA, 1967.

80. Acton, F. S. Numerical Methods that Work; Harper and Row: New York, 1970.

81. (a) Ramachandran, G. N.; Srinivasan, R. Fourier Methods in Crystallography; Wiley-Interscience: New York, 1970; Chapter 1. (b) Lipson, H.; Taylor, C. A. Fourier Transform and X-ray Diffraction; Bell: London, 1958

82. Weiss, R. J. X-ray Determination of Electron Distributions; Elsevier North-Holland: Amsterdam, 1966.

83. Stewart, R. F.; Spackman, M. A. In ref. 5; Vol. 1.

84. Spackman, M. A.; Stewart, R. F. In ref. 57, pp 302-320.

85. Marumo, F. In ref. 57, pp 321-332.

86. Costain, W. Ph.D. Thesis, University of Birmingham, England, 1941.

87. (a) Booth, A. D. Proc. Roy. Soc. (London) A 1946, 188, 77-92. (b) Booth, A. D. ibid. 1947, 190, 482-489. (c) Clews, C. J. B.; Cochran, W. Acta Crystallogr. 1949, 2, 46-57. (d) Cruickshank, D. W. J. ibid. 1949, 2, 65-82.

88. De Ranter, C. J.; Van Dijck, M. Z. Kristallogr., Kristallgeometrie, Kristallphys., Kristallchem. 1971, 134, 97-106.

89. (a) Linkoaho, M. V. H. Acta Crystallogr., Sect. A: Found. Crystallogr. 1980, 36, 155-156. (b) Linkoaho, M. V. H. ibid. 1979, 35, 140-142. (c) Patomäki, L. K.; Linkoaho, M. V. ibid. 1969, 25, 304-305.

90. Linkoaho, M. V. Phys. Scr. 1972, 5, 271-272.

91. (a) Witte, H.; Wölfel, E. Z. physik. Chem. 1955, 3, 296-329. (b) Wagner, B.; Witte, H.; Wölfel, E. ibid., 273-295.

92. Forsyth, J. B.; Wells, M. Acta Crystallogr. 1959, 12, 412-415.

93. Moore, F. H. Acta Crystallogr. 1963, 16, 1169-1175.

94. Tokonami, M. Acta Crystallogr. 1965, 19, 486.

95. Hovestreydt, E. Acta Crystallogr., Sect. A: Found. Crystallogr. 1983, 39, 268-269.

96. Stewart, R. F.; Davidson, E. R.; Simpson, W. T. J. Chem. Phys. 1965, 42, 3175-3187.

97. (a) Hajdu, F. Acta Crystallogr., Sect. A: Found. Crystallogr. 1972, 28, 250-252. (b) Hajdu, F. ibid. 1971, 27, 73-74.

98. Smith, V. H.; Thakkar, A. J.; Chapman, D. C. Acta Crystallogr., Sect. A: Found. Crystallogr. 1975, 31, 391-392.

99. Jiang, J.; Li, F. Wuli Xuebao 1984, 33(6), 845-849; Chem. Abstr. 1984, 101, 220293s.

100. Smith, G. H.; Burge, R. E. Acta Crystallogr. 1962, 15, 182-186.

101. (a) Yáñez, M.; Stewart, R. F. Acta Crystallogr., Sect. A: Found. Crystallogr. 1978, 34, 648-651. (b) Yáñez, M.; Stewart, R.F.; Pople, J.A. ibid 1978, 34, 641-648.

102. Lee, J. D.; Pakes, H. W. Acta Crystallogr., Sect. A: Found. Crystallogr. 1969, 25, 712-713.

103. (a) Agarwal, R. C. Acta Crystallogr., Sect. A: Found. Crystallogr. 1978, 34, 791-809. (b) Isaacs, N, W.; Agarwal, R. C. ibid. 1978, 34, 782-791.

104. Mayer, S. W. Ph.D. Thesis, University of California, Los Angles, 1953.

105. Fike, C. T. Computer Evaluation of Mathematical Functions; Prentice-Hall: Englewood Cliffs, NJ, 1968.

106. Hart, J. F.; Cheney, E. W.; Lawson, C. L.; Maehly, H. J.; Mesztenyi, C. K.; Rice, J. R.; Thacher, H. G.; Witzgall, C. Computer Approximations; John Wiley: New York, 1968.

107. Gillis, J. Acta Crystallogr. 1958, 11, 833-834.

108. Silverman, J. N.; Simonsen, S. H. Acta Crystallogr. 1960, 13, 50-54.

109. Onken, H.; Fisher, K. F. Z. Kristallogr. 1968, 127, 188-199.

110. Pilyankevich, A. N.; Vereshchaka, V. M. Zavod. Lab. 1983, 49, 26-29. Chem. Abstr. 1983, 99, 218917v.

111. Shibata, S. Bull. Chem. Soc. Jpn. 1966, 39, 1612-1613.

112. Ladd, M. F. C. Z. Kristallogr. 1966, 123, 373-381.

113. (a) Gallant, A. R. Am. Statistician 1975, 29, 73-81. (b) Clenshaw, C. W. In ref. 122, Chapter 2. (c) Donnelly, J. D. P. In ref. 120, Chapter 15.

114. Lark, P. D.; Craven, B. R.; Bosworth, R. C. L. The Handling of Chemical Data; Pergamon Press: New York, 1968; pp 251-253.

115. Ref. 69, p 626.

116. McCoy, N. H. Introduction to Modern Algebra, 3rd ed.; Allyn and Bacon: Boston, 1975; Chapter 9.

117. (a) Linz, P. Theoretical Numerical Analysis: An Introduction to Advanced Techniques; Wiley-Interscience: New York, 1979; Chapter 2. (b) See also: ref. 120.

118. Wendroff, B. Theoretical Numerical Analysis; Academic Press: New York, 1966; Chapter 2.

119. Approximation of Functions; Garabedian, H. L., Ed.; Elsevier: Amsterdam, 1965.

120. Methods of Numerical Approximation; Handscomb, D. C., Ed.; Pergamon Press: Oxford, 1966.

121. Numerical Analysis: An Introduction; Walsh, J., Ed.; Thompson: Washington, DC, 1967.

122. Numerical Approximation to Functions and Data; Hayes, J. G., Ed.; Athlone: London, 1970.

123. Personal Computers in Chemistry; Lykos, P., Ed.; Wiley-Interscience: New York, 1981.

124. Stineman, R. W. Creative Computing 1980, 6(7), 54-57.

125. Warme, P. K. Curve Fitter; Interactive Microware: State College, PA, 1980.

126. Forsythe, G. E.; Moler, C. B. Computer Solution of Linear Algebraic Systems; Prentice-Hall: Englewood Cliffs, NJ, 1967.

127. Ortega, J. M. Numerical Analysis: A Second Course; Academic Press: New York, 1972.

128. Hayes, J. G. In ref. 122, Chapter 5.

129. Noggle, J. H. Physical Chemistry on a Microcomputer; Little, Brown: Boston, 1985.

130. Kuo, S. S. Computer Application of Numerical Methods; Addison-Wesley: Reading, MA, 1972.

131. Series 600/6000 Time-Sharing Application Library Guide, Volume II -- Statistics; Honeywell No. DA44; Honeywell Information Systems, Inc., 1973; Vol. 2, p ST-127.

132. Bevington, P. R. Data Reduction and Error Analysis for the Physical Sciences; McGraw-Hill: New York, 1969; Chapter 8.

133. Ref. 136, pp 334-341. See also: ref. 231.

134. Neter, J.; Wasserman, W.; Whitmore, G. A. Applied Statistics; Allyn and Bacon: Boston, 1978.

135. Shorter, J. Correlation Analysis of Organic Reactivity; Research Studies Press: New York, 1982; Chapters 2 and 7.

136. Burr, I. W. Applied Statistical Methods; Academic Press: New York, 1974.

137. Thorndike, R. M. Correlational Procedures for Research; Gardner Press: New York, 1978.

138. Chatterjee, S.; Price, B. Regression Analysis by Example; John Wiley: New York, 1977.

139. Robinson, E. A. Least Squares Regression Analysis in Terms of Linear Algebra; Goose Pond Press: Houston, TX, 1981.

140. Lee, J. D.; Lee, T. D. Statistics and Numerical Methods in BASIC for Biologist; Van Nostrand, Reinhold: Workingham, Berkshire, England, 1982.

141. Dowdy, S.; Wearden, S. Statistics for Research; John Wiley: New York, 1983.

142. Middle Tennessee State University Computer Center User's Guide; Lea, L., Ed.; Middle Tennessee State University: Murfreesboro, TN, Jan. 1984.

143. Cox, M. J.; Sullivan, K. B. Structuring Programs in Microsoft BASIC; Boyd and Fraser: Boston, 1986.

144. Mandell, S. L. BASIC Programming Today; West: New York, 1986.

145. Hamming, R. W. Numerical Methods for Scientists and Engineers; McGraw-Hill: New York, 1962.

146. Ref. 78, p 101, problem 4.26.

147. Steffensen, J. F. Interpolation, 2nd ed.; Chelsea: New York, 1950.

148. Dorn, W. S.; McCracken, D. D. Numerical Methods with FORTRAN IV Case Studies; John Wiley: New York, 1972; pp 155-175 and 324-330.

149. Warme, P. K. Scientific Plotter; Interactive Microware: State College, PA, 1980.

150. Corl, T. L.; Warme, P. K. Curve Fitter-PC; Interactive Microware: State College, PA, 1984.

151. Coulthard, M. A. Proc. Phys. Soc. London. 1967, 91, 44-49.

152. Cromer, D. T.; Mann, J. B. Acta. Crystallogr., Sect. A: Found. Crystallogr. 1968, 24, 321-324.

153. Mann, J. B. Los Alamos Scientific Report LA3691, 1968; Los Alamos Scientific Laboratory, Los Alamos, NM, p 196.

154. James, R. W. X-ray Crystallography 5th ed.; Methuen: London, 1961.

155. (a) Berghuis, J.; Haanappel, I. M.; Potters, M.; Loopstra, B. O.; MacGillavry, C. H., Veenendaal, A. L. Acta Crystallogr. 1955, 8, 478-483.   (b) Ibers, J. A. Acta Crystallogr. 1957, 10, 86.

156. Lavine, L. R.; Rollett, J. S. Acta Crystallogr. 1956, 9, 269-273.

157. Ahmed, F. R.; Cruickshank, D. W. J. Acta Crystallogr. 1953, 6, 765-769.

158. Sayre, D. Acta Crystallogr. 1951, 4, 362-367.

159. Ten Eyck, L. F. Acta Crystallogr., Sect. A: Found. Crystallogr. 1977, 33, 486-492.

160. Shelly, G. B.; Cashman, T. J. Computer Fundamentals for an Information Age; Anaheim: Brea, CA, 1984.

161. Brainerd, W. S.; Goldberg, C. H.; Gross, J. L. WATF IV-5 Fundamentals and Style; Boyd and Fraser: Boston, 1986; pp 1-10.

162. Davis, R. M. Science 1977, 195, 1096-1102.

163. Friedrich, O. Time 1983, 121(1), 4-24.

164. Cocks, J. Time 1983, 121(1), 25-27.

165. A Basic Library List; Mathematical Association of America: Washington, DC, 1961. For a compendium of CUPM recommendations, see ref. 230.

166. Chemistry: An Experimental Science; Pimentel, G. C., Ed.; W. H. Freeman: San Franscisco, 1963.

167. Merrill, R. J.; Ridgway, D. W. The CHEM Study Story; W. H. Freeman: San Francisco, 1969.

168. Physics; Physical Science Study Committee; D. C. Heath: Boston, 1960; ibid., 2nd ed., 1963.

169. Investigating the Earth; Earth Science Curriculum Project; Houghton-Mifflin: Boston, 1967.

170. Ref. 121, Chapter 12.

171. New Movements in the Study and Teaching of Chemistry; Daniels, D. J., Ed.; Temple Smith: London, 1975.

172. A Basic Library List for Four Year Colleges, 2nd ed.; The Mathematical Association of America: Washington, DC, 1976.

173. Tompkins, C. B.; Wilson, W. L. Elementary Numerical Analysis; Prentice-Hall: Englewood Cliffs, NJ, 1969.

174. Soltzberg, L. J. Computer Strategies for Chemistry Students; Houghton-Mifflin: Dallas, 1987.

175. Isenhour, T. L.; Jurs, P. C. Introduction to Computer Programming for Chemists: FORTRAN, 2nd ed.; Allyn and Bacon: Boston, 1979.

176. Beech, G. FORTRAN IV in Chemistry: An Introduction to Computer-Assisted Methods; John Wiley: London, 1975.

177. Dickson, T. R. The Computer and Chemistry; W. H. Freeman: San Francisco, 1968.

178. Wiberg, K. B. Computer Programming for Chemists; W. A. Benjamin: New York, 1965.

179. Meyer, S. L. Data Analysis for Scientists and Engineers; John Wiley: New York, 1975.

180. Youmans, H. L. Statistics for Chemistry; Charles E. Merrill: Columbus, OH, 1973.

181. Daniel, C.; Wood, F. S. Fitting Equations to Data: Computer Analysis of Multifactor Data, 2nd ed.; John Wiley: New York, 1980.

182. Kittrell, J. R.; Mezaki, R.; Watson, C. C. Ind. Eng. Chem. 1965, 57(12), 18-27.

183. Copeland, T. G. J. Chem. Educ. 1984, 61, 778-779.

184. Dombi, G. W. J. Chem. Educ. 1984, 61, 527-528.

185. Trindle, C. J. Chem. Educ. 1983, 60, 566-567.

186. De Marco, D.; Bellomo, A.; De Robertis, A. J. Chem. Educ. 1981, 58, 696.

187. Menéndez, V.; Acuña, A. U. J. Chem. Educ. 1976, 53, 98.

188. Anderson, R. H. J. Chem. Educ. 1975, 52, 776.

189. Dye, J. L.; Nicely, V. A. J. Chem. Educ. 1971, 48, 443-448.

190. Krieger, A. G.; Wallace, C. K. J. Chem. Educ. 1971, 48, 457.

191. Pollnow, G. F. J. Chem. Educ. 1971, 48, 518-519.

192. Pollnow, G. F. In Personal Computers in Chemistry; Lykos, P., Ed.; Wiley-Interscience: New York, 1981; Chapter 12.

193. Carter, D. W. J. Chem. Educ. 1985, 62, 497-498.

194. Adams, K. A. H.; Storer, A. C.; Cornish-Bowden, A. J. Chem. Educ. 1984, 61, 527.

195. Jaskólski, M. J. Chem. Educ. 1983, 60, 980-981.

196. Waldo, G. S.; Schulze, C. A.; Battino, R. J. Chem. Educ. 1984, 61, 530-531.

197. Gray, E. T.; Workman, H. J. J. Chem. Educ. 1981, 58, 239-241.

198. Nathan, L. C. J. Chem. Educ. 1975, 52, 438.

199. Moore, J. W.; Moore, E. A. J. Chem. Educ. 1986, 63, 248-251.

200. Kim, H. J. Chem. Educ. 1970, 47, 120-122.

201. Jurs, P. C. In ref. 207, Chapter 1.

202. Perone, S. P. J. Chem. Educ. 1970, 47, 105-111.

203. (a) Perone, S. P.; Eagleston, J. F. J. Chem. Educ. 1971, 48, 317-322; (b) Perone, S. P.; Eagleston, J. F. ibid. 1971, 48, 438-442.

204. Schilling, J. W. In ref. 207, Chapter 9.

205. DeToma, R. P. In ref. 207, Chapter 2.

206. Zebolsky, D. In ref. 207, Chapter 5.

207. Computer Education of Chemists; Lykos, P., Ed.; Wiley-Interscience: New York, 1984.

208. Geist, W.; Ripota, P. In Computers in Chemistry; Springer-Verlag: New York, 1973, pp 169-195.

209. (a) Computer Networks in the Chemical Laboratory; Levy, G. C.; Terpstra, D., Eds.; Wiley-Interscience: New York, 1981. (b) Supercomputers in Chemistry; Lykos, P.; Shavitt, I., Eds.' American Chemical Society: Washington, DC, 1981. (c) Data Processing in Chemistry, Hippe, Z., Ed.; Elsevier North-Holland: Amsterdam, 1981.

210. Hamilton, W. C. Statistics in Physical Science; Ronald Press: New York, 1964.

211. Soltzberg, L; Shah, A. A.; Saber, J. C.; Canty, E. T. BASIC and Chemistry; Houghton-Mifflin: Boston, 1975.

212. Johnson, K. J. Numerical Methods in Chemistry; Marcel Dekker: New York, 1980.

213. Norris, A. C. Computational Chemistry; John Wiley: New York, 1981.

214. Wilson, S. Chemistry by Computer: An Overview of the Applications of Computers in Chemistry; Plenum Press: New York, 1986.

215. Wilkins, C. L.; Perone, S. P.; Klopfenstein, C. E.; Williams, R. C.; Jones, D. E. Digital Electronics and Laboratory Computer Experiments; Plenum Press: New York, 1975.

216. Clarke, F. H. Calculator Programming for Chemistry and the Life Sciences; Academic Press: New York, 1981.

217. Ouchi, G. J. Personal Computers for Scientists: A Byte at a Time; American Chemical Society: Washington, DC, 1986.

218. Athey, T. H.; Day, J. C.; Zmud, R. W. Computers and End-User Software with BASIC; Scott, Foresman: Glenview, IL, 1987.

219. (a) Schwarz, K.; Schulz, H. Acta Crystallogr., Sect. A: Found. Crystallogr. 1978, 34, 994-999. (b) Schulz, H.; Schwarz, K. ibid., 999-1005. (c) Boebel, G.; Cortona, P. J. Phys. B 1983, 16, 349-357.

220. Schmidt, P. C.; Sen, K. D.; Weiss, A. Ber. Bunsenges. Phys. Chem. 1980, 84, 1240-1251.

221. Schmidt, P. C.; Weiss, A. Z. Naturforsch., A: Phys., Phys. Chem., Kosmosphys. 1979, 34, 1471-1481.

222. Thomas, G. D.; Jeffries, T. V.; Stewart, M. V. Presented at the 18th Annual Southeastern Regional Conference of Undergraduate Chemists, Clemson University, Clemson, SC, April 1986; paper 17; ibid., Abstracts of Papers, p 18.

223. Jeffries, T. V.; Stewart, M. V.; Thomas, G. D. J. Tenn. Acad. Sci. 1986, 61, 36.

224. Jeffries, T. V.; Stewart, M. V.; Thomas, G. D. Presented at the 95th Annual Meeting of the Tennessee Academy of Science, Middle Tennessee State University, Murfreesboro, TN, Nov. 1985.

225. Stewart, M. V. Presented at the Winter Meeting of the American Crystallographic Association, Eufaula, AL, March 1980; paper P5; ibid., ACA Program and Abstracts, Ser. 2 1980, 7(2), 39.

226. (a) Holden, A.; Morrison, P. _Crystals and Crystal Growing_, 1st MIT Press ed.; MIT Press: Cambridge, MA, 1982. (b) Holden, A.; Singer, P. _ibid._; Anchor Books, Doubleday: Garden City, NY, 1960.

227. Skoog, D. A. _Principles of Instrumental Analysis_, 3rd ed.; Saunders College Publishing: Philadelphia, PA, 1985.

228. Allendoerfer, R. D. _Explorations in Chemistry_; Prentice-Hall: Englewood Cliffs, NJ, 1985.

229. Schatz, P. F. _NMR Simulator for the IBM PC_; City Software: Milwaukee, WI, 1985.

230. _A Compendium of CUPM Recommendations_; The Mathematical Association of America: Washington, DC, 1974; 2 vols.

231. Jackson, D. _Fourier Series and Orthogonal Polynomials_; The Mathematical Association of America; The Collegiate Press, George Banta: Menasha, WI, 1941.