Computational Methods for the Determination of Formation Constants

Edited by
DAVID J. LEGGETT

Computational Methods for the Determination of Formation Constants

Edited by
DAVID J. LEGGETT
Dow Chemical USA
Texas Division
Freeport, Texas

Plenum Press • New Y

Library of Congress Cataloging in Publication Data

Main entry under title:

Computational methods for the determination of formation constants.

(Modern inorganic chemistry)
Bibliography: p.
Includes index:
1. Phase rule and equilibrium – Data processing. I. Leggett, David J. II. Series.
QD503.C66 1985 541.3'63'0285 85-16991
ISBN 0-306-41957-2

A Division of Plenum Publishing Corporation
233 Spring Street, New York, N.Y. 10013

Printed in the United States of America

MODERN INORGANIC CHEMISTRY

Series Editor: John P. Fackler, Jr.
Texas A&M University

METAL INTERACTIONS WITH BORON CLUSTERS
Edited by Russell N. Grimes

HOMOGENEOUS CATALYSIS
WITH METAL PHOSPHINE COMPLEXES
Edited by Louis H. Pignolet

THE JAHN-TELLER EFFECT AND
VIBRONIC INTERACTIONS IN MODERN CHEMISTRY
I. B. Bersuker

MÖSSBAUER SPECTROSCOPY
APPLIED TO INORGANIC CHEMISTRY, *Volume 1*
Edited by Gary J. Long

CARBON-FUNCTIONAL ORGANOSILICON COMPOUNDS
Edited by Václav Chvalovský and Jon M. Bellama

COMPUTATIONAL METHODS FOR THE DETERMINATION
OF FORMATION CONSTANTS
Edited by David J. Leggett

A Continuation Order Plan is available for this series. A continuation order will bring delivery of each new volume immediately upon publication. Volumes are billed only upon actual shipment. For further information please contact the publisher.

Contributors

Alex Avdeef • Syracuse University, Department of Chemistry, Syracuse, New York 13210

Josef Havel • Department of Analytical Chemistry, J. E. Purkyne University, Kotlarska 2, 611 37 Brno, Czechoslovakia

David J. Leggett • Dow Chemical USA, Texas Division, Freeport, Texas 77541

Peter M. May • Department of Applied Chemistry, University of Wales, Institute of Science and Technology, Cardiff CF1 3NU, United Kingdom

Milan Meloun • Department of Analytical Chemistry, University of Chemical Technology, Leninovo nam. 565, 532 10 Pardubice, Czechoslovakia

I. Nagypal • Computer Center, Institute of Inorganic and Analytical Chemistry, Lajos Kossuth University, H-4010, Debrecen. Hungary

D. D. Perrin • Medical Chemistry Group, The John Curtin School of Medical Research, The Australian National University, P.O. Box 334, Canberra City, A.C.T., 2601, Australia

Antonio Sabatini • Dipartimento di Chimica dell' Università di Firenze and I.S.S.E.C.C. - CNR, Florence, Italy

H. Stunzi • Universite de Neuchatel, Institut de Chimie, Avenue de Bellevaux 51, CH-2000 Neuchatel, Switzerland

Alberto Vacca • Dipartimento di Chimica dell' Università di Firenze and I.S.S.E.C.C. - CNR, Florence, Italy

David R. Williams • Department of Applied Chemistry, University of Wales, Institute of Science and Technology, Cardiff CF1 3NU, United Kingdom

L. Zekany • Computer Center, Institute of Inorganic and Analytical Chemistry, Lajos Kossuth University, H-4010, Debrecen, Hungary

Preface

This volume is concerned with methods that are available for the calculation of formation constants, in particular computational procedures. Although graphical methods have considerable value in the exploration of primary (raw) data they have been overtaken by computational methods, which, for the most part, take primary data and return the refined formation constants. Graphical methods are now considered complementary to these general computational procedures.

This volume brings together programs that span the lifetime of computer-assisted determination of formation constants. On one hand the reader will find listings of programs that are derived from LETAGROP (b.1961) and the GAUSS-G/SCOGS (b.1962) families. On the other hand programs are presented that are the newest members of the SCOGS lineage and from the on-going MINIQUAD series. One program is presented that describes a computational approach to the classical Hedstrom–Osterberg methods; another that takes care of electrode calibration in a simple yet rigorous manner.

Potentiometry and spectrophotometry are the most popular experimental techniques for equilibrium studies, and the programs in this volume reflect this. Four programs handle potentiometric data, two will process spectrophotometric data, and one makes use of both types of data separately or in combination.

This collection of programs also represents a collection of well-tested numerical algorithms for the general problem of data fitting. As such the volume will have value to users not necessarily devoted to full-time solution equilibrium studies. Each program employs a different, and in several instances novel, method for data input.

Every effort has been made to ensure that the codings are accurate. Of course, each program has been run on the author's computer. More importantly the programs have also been tested on a Honeywell 66/60 and on a Vax 11/780, under the FORTRAN 66 compiler. The input data, supplied by each author and used to test the programs, have been included. Minor changes were made to some of the programs to work around local features of the original machines. The program listings were printed out using a daisy-wheel printer after being down-loaded from the Vax 11/780 to a DecMate word processor. The listings, so produced, clearly distinguish between a zero and upper case "oh"; *D* and "oh" are also obviously different.

I would like to take this opportunity to acknowledge the patience and forebear-

ance of a number of people in what has turned out to be a longer than expected time to produce the volume: the contributors to this volume; Plenum as publishers; my wife.

It would be remiss of me not to acknowledge the positive guidance of Professor W.A.E. (Pete) McBryde during my graduate student days; but for him none of this would have been possible.

David J. Leggett

Contents

3 MAGEC: A Program for the Definite Calibration of the Glass Electrode

PETER M. MAY and DAVID R. WILLIAMS

4 SCOGS2: A Nonlinear Least-Squares Program for the Evaluation of Formation Constants of Metal Complexes

D.D. PERRIN and H. STUNZI

7 General Computer Programs for the Determination of Formation Constants from Various Types of Data

JOSEF HAVEL and MILAN MELOUN

8 PSEQUAD: A Comprehensive Program for the Evaluation of Potentiometric and/or Spectrophotometric Equilibrium Data Using Analytical Derivatives

L. ZEKANY and I. NAGYPAL

9 STBLTY: Methods for Construction and Refinement of Equilibrium Models

ALEX AVDEEF

1

The Determination of Formation Constants

An Overview of Computational Methods for Data Processing

DAVID J. LEGGETT

1. INTRODUCTION

The interaction of metal ions and ligands, in solution, giving rise to equilibrium mixtures of coordination complexes has been studied systematically since 1941,[1] although many references atest to 30 or more years of earlier research. Until about 1960 data obtained from these studies were subjected to graphical analysis. Many valuable contributions were made by Hazel and Francis Rossotti (their book[2] remains the definitive text on the subject) and by Lars Gunnar Sillen.[3] The impact of Sillen and co-workers on solution equilibria has not been restricted to innovative graphical procedures. In 1961[4] the first general computer program for the evaluation of formation constants from equilibrium data was announced. Many programs have followed from several independent research groups, including the LETAGROP[5] family of computational techniques.

This review will cover the literature up to the end of 1983. The programs will be classified according to the principal algorithm used. Several reviews, discussing the subject from different viewpoints, have been published.[6-11] Allen and McMeeking[12] have also reviewed the general topic of optimization methods.

The "best fit" of formation constants to the observed data may be expressed in terms of the least-squares criterion. The quantity U, the sum of squares of residuals, is defined according to

$$U = \sum_{k=1}^{\mathrm{NPTS}} (Y^k_{\mathrm{obs}} - Y^k_{\mathrm{calc}})^2 \qquad (1)$$

DAVID J. LEGGETT • Dow Chemical U.S.A., Texas Division, Freeport, Texas 77541

where Y^k_{obs} is the observed data and Y^k_{calc} is the value of the objective function defined in general terms as $f(\mathbf{x}, \mathbf{t})$. The parameter vector, $\mathbf{x}$, will always comprise the formation constants and may also include additional parameters relating to the model (e.g., molar absorptivities, enthalpy of formation, etc.) and/or parameters relating to the experimental conditions (e.g., analytical concentrations, electrode calibration constants, etc). This vector is commonly known as the dependent variable vector. The vector $\mathbf{t}$ contains the variables determined by the experimenter, known as the independent variables. Thus we seek to locate the set of dependent variables that will minimize U. Inherent to this approach is the need to define the equilibrium model before undertaking the minimization process. A general method that does not require predefinition of the model will be described in Section 4.

2. DIRECT SEARCH AND UNIVARIATE METHODS

Several such methods are available. They are, as the name suggests, techniques wherein one or more parameters are systematically varied. The function may be minimized by monitoring the impact of these parameter variations. The simplest search methods are Grid, Star, and Composite design; Fibonacci, and Golden Section. More elaborate methods include Simplex, Hooke and Jeeves, and interpolation techniques.

2.1. Grid, Star, and Composite Designs

These three related techniques involve the establishment of an n-dimensional figure, the minimum being located within the figure. The objective function is evaluated at each point of the figure. Although the Star and Composite design figures are more often encountered in experimental design protocols,[13] they can serve as useful alternatives to the simple Grid Search. While these methods are slow to converge, they provide a uniform overview of the hypersurface and hence can reveal multiple minima. Izatt *et al.*[14] employed a Grid Search of calorimetric data and used the coordinates of the located minimum ($\log \beta_1$ and $\log \beta_2$) as initial estimates for Davidon's [15] cubic interpolation. Nagano and Metzler adopted the same approach as Izatt in order to locate suitable starting estimates for their program PITMAP.[16] It was noted that the grid method required approximately 60 times more CPU resources than did Davidon's algorithm. Leussing[17] used a contracting grid approach in an ingenious manner to determine three formation constants for Schiff's base complexes with nickel. Natansohn *et al.*[18] employed the contracting grid method using a steepest descent alogorithm to reposition the grid under certain conditions. This publication describes an interesting approach to the application of experimental design techniques and error analysis to equilibrium modeling. The information given by Natansohn is presented in a manner that is not seen in the majority of publications relating to the estimation of formation constants. SQUAD[19] (described in chapter 6) has an option for a Star design search to obtain initial estimates for the sought parameters.

2.2. Fibonacci and Golden Section

The Fibonacci sequence, F_n, discovered by Leonardo of Pisa (1175–1230), has been used to minimize a unimodal function. The incorporation of this sequence into the minimization procedure is due to Kiefer.[20] The method is reliable for functions with a single minimum but may not converge for general functions.

The Golden Section is based on the fact that for any pair of Fibonacci numbers the relationship

$$\lim_{n\to\infty} \frac{F_n}{F_{n+1}} = \frac{\sqrt{5}-1}{2} \tag{2}$$

is valid. Consequently the Golden Section algorithm is a limiting form of the Fibonacci search. The Golden Section method has been employed for the determination of a single constant.[21]

2.3. Hooke and Jeeves, Simplex

The method of Hooke and Jeeves involves a two-step procedure. In the first step each parameter axis is explored in order to locate the best parameter vector in the exploratory space. Next, the search is moved in the direction of each vector. These two steps alternate, the search length being reduced as the minimum is approached.

The Simplex algorithm, originated by Spendley *et al.*[22] and improved by Nelder and Mean,[23] is essentially the hypersurface-sensitive adaptive movement of an n-dimensional figure through hyperspace. The Simplex technique has been employed in several areas of chemistry, and has been admirably reviewed and demonstrated by Deming.[24] In terms of numerical minimization Nelder and Mead's version has been described by Walsh[25] as "...one of the most efficient pattern search methods currently available...." May and Williams use the Simplex algorithm as the major refinement method in their electrode calibration program MAGEC, described in a later chapter.

Although the computer code requirements for these search methods are small, making them ideal for mini- and microcomputer use, convergence is slow when close to the minimum.

2.4. Quadratic and Cubic Interpolation Methods

Consider the linear search problem as the minimization of the function $f(x)$ along the line

$$x = x_k + \lambda d \tag{3}$$

where x_k is the current point and d is the given search direction. Minimization may be achieved by minimizing the value of λ using quadratic interpolation[26] or cubic interpolation.[15] Powell's[26] method, i.e., quadratic fitting of the objective function

at three values of λ followed by location of the turning point of the quadratic equation, provides values of λ that may be the minimum value. The process locates the minimum iteratively. Quadratic interpolation is a one-dimensional search and may be incorporated into other algorithms requiring an efficient linear search, such as Powell 65.[27]

Alternatively, the minimum along a descent line may be located by means of a cubic interpolation procedure.[15] Davidon's method requires that the function value and the gradient are available at two points which bracket the minimum. Similar to Powell's method, Davidon's technique can be used in algorithms that require a linear search. The procedure is an integral part of the variable metric minimization methods, the most widely used of which is the Davidon–Fletcher–Powell (DVP) method.[28]

The principal difference between the two approaches is that while the former uses function values only, the latter requires the evaluation of the function and its derivatives. Thus Davidon's approach is more elaborate than Powell's but usually locates the minimum of $f(x)$ in fewer iterations.

Powell's algorithm has been employed by Sabatini *et al.*[29] in MINIQUAD, where the normal equation matrix (factorized) is linearly optimized by the subroutine LIMIN.

The application of the DVP algorithm to the computation of force constants was illustrated by Gans.[30, 31] Subsequently, the DVP method has been incorporated into Gans and Vacca's program STEW.[32] The authors noted improved convergence speed and numerical stability compared to SCOGS[33] and LETAGROP[5]; but STEW was comparable to LEAST.[32] It was also demonstrated that the refinement of the logarithm of the formation constant is intrinsically ill conditioned.

In a study of aluminum fluoride interactions, using a fluoride ion-specific electrode, Baumann[35] processed the data using the DVP algorithm.

Christensen *et al.*[14] have adapted Davidon's basic variable metric algorithm to permit the evaluation of formation constants from calorimetric data.

2.5. Summary

The availability of computers, together with faster CPU, make the search and univariate methods described here increasingly attractive. Sections 2.1, 2.2, and 2.3 briefly describe the more popular search techniques that have the distinct advantage of being straightforward in their methods of operation. Other methods that warrant usage in the area of minimization by search techniques include the secant and dichotomous search methods; the algorithm of Davies, Swann, and Campey, all for single-variable optimizations[36]; and the direct search method of Chandler[37] as implemented by Kankare[38] in the spectrophotometric determination of formation constants.

The linear search techniques of Powell and Davidon are of a more complex nature and correspondingly more powerful. Jacoby, Kowalik, and Pizzo[39] suggest that Davidon's method be used when gradients are available, otherwise that of Davies, Swann, and Campey is recommended for the initial interpolation, switching to Powell's for subsequent moves.