Coordination Chemistry Reviews, 120 (1992) 389–405
Elsevier Science Publishers B.V., Amsterdam

Mathematical algorithms and computer programs for the determination of equilibrium constants from potentiometric and spectrophotometric measurements

A. Sabatini and A. Vacca
Dipartimento di Chimica, Università di Firenze, Via Maragliano 77, 50144 Florence (Italy)
P. Gans
School of Chemistry, The University, Leeds LS2 9JT (UK)

(Received 8 October 1991)

CONTENTS

A. INTRODUCTION

During the last thirty years there have been many developments in the mathematical algorithms, computer programs and experimental techniques used to determine equilibrium constants in solution [1,2]. Developments are continuing apace, spurred on by the fact that the experimental investigations are being directed at ever more complex systems. In particular, the identification and characterization of species present in biofluids, surface and underground waters, effluents, etc. constitutes an important present-day problem which demands the use of modern analytical techniques and computer programs that can be used to determine the formation constants of all the species present in these natural solutions.

Our contribution in this field has been to develop a series of programs to be used for the refinement of formation constants based on potentiometric titration data. The first of these was called LEAST [3]. In LEAST, a least-squares function was minimized which permitted, for the first time, the use of analytical expressions

0010-8545/92/$15.00

for the partial derivatives required by the minimization. Mainly as a result of this, LEAST was much faster and more stable than the other programs available at that time, which used numerical expressions for the requisite derivatives. However, LEAST was restricted to systems containing at most one metal ion and one ligand. In MINIQUAD [4], we have developed a program based on the same minimization procedure as LEAST but with more general applicability; MINIQUAD can handle systems with two metal ions and one ligand (competing equilibria) or two ligands and one metal ion (ternary complexes) and also offers the possibility of using more than one ion-sensitive electrode. MINIQUAD has been used extensively all over the world. Various improvements were made in MINIQUAD75 [5] which resulted in greater execution speed.

The experience acquired in the application of this program was of considerable value in the development of the next program in the series, SUPERQUAD [6]. As well as using a different algorithm, which allowed statistical testing to be properly applied for the first time in any potentiometric system, new facilities were introduced for the refinement of the amount of substance used and of the standard electrode potential. The new facilities were needed in order to quantify equilibria involving ligands which cannot be satisfactorily purified, such as peptides isolated from living systems. SUPERQUAD also incorporated a species selection procedure in order that it may be used as a tool, not only for equilibrium constant determination, but also for selection of the best chemical model for a system.

To move forward in the field of solution equilibria, the development of new experimental techniques and new programs must go hand in hand. We feel that a useful experimental direction to take lies in the simultaneous acquisition of data by more than one technique. To go in this direction, a new computer program will be needed capable of handling both potentiometric data and data from another source such as spectrophotometry. We have written such a program and called it HYPERQUAD. This program will offer all the facilities for model selection on multicomponent equilibria that were built into SUPERQUAD; it is at present undergoing testing and development of the user–data interface.

In this paper we will describe in detail the mathematical algorithms used in the above-mentioned programs and the services they provide. Our object is to provide a clear picture of the possibilities and limitations inherent in the programs. We hope that this will be useful to potential and actual users by giving them a deeper understanding of the methodology in order that they may use the programs more effectively. It should also provide a basis for new developments.

B. THE METHOD OF LEAST-SQUARES

All the programs that we shall discuss are based on the method of least-squares [7]. As an introduction to the various algorithms, we must first give a brief summary of this method outlining the formalism that will be used. We begin by defining a model of the chemical system. Suppose that n observations have been made. Corre-

sponding to each observed quantity, f_i^0, there is a calculated value, $f_i(v_i, p_1, ..., p_m)$, which is a function of the ith independent variable, v_i, and a set of m parameters, $p_1, ..., p_m$. This functional relationship defines the model. We define the residuals, r_i, in eqn. (1).

$$r_i = f_i^0 - f_i(v_i, p_1, ..., p_m); \quad \mathbf{r} = \{r_1, ..., r_n\}^T \tag{1}$$

The method of least-squares is used to calculate that value of the m-vector, $\mathbf{p}$, of parameters that minimizes U, the weighted sum of squares and products of residuals (eqn. (2)). U is called the objective function.

$$U = \sum_{h=1,m} \sum_{i=1,m} W_{hi} r_h r_i; \quad U = \mathbf{r}^T \mathbf{W} \mathbf{r} \tag{2}$$

W_{ij} is an element of the weight matrix $\boldsymbol{W}$, which should be the inverse of the variance–covariance matrix of the observations. When, as is often the case, the observational errors are uncorrelated, $\boldsymbol{W}$ is diagonal and each of its elements, W_{ii}, is the reciprocal of the variance of the ith observation; in this case the objective function simplifies to a weighted sum of squared residuals

$$U = \sum_{i=1,m} W_{ii} r_i^2 \tag{3}$$

The minimum of the objective function is to be found by setting its partial derivatives with respect to each parameter to zero.

$$\frac{\partial U}{\partial p_j} = -\sum_h \sum_i W_{hi} \left(r_h \frac{\partial f_i}{\partial p_j} + r_i \frac{\partial f_h}{\partial p_j} \right) = 0; \quad \boldsymbol{g} = -2\boldsymbol{J}^T \boldsymbol{W} \boldsymbol{r} = 0 \tag{4}$$

$\boldsymbol{g}$ is called the gradient vector, and contains the m elements $\partial U/\partial p_j$. $\boldsymbol{J}$ is known as the Jacobian matrix whose elements are the partial derivatives $\partial f_i/\partial p_j$.

In linear systems, f_i is a linear combination of the parameters, $\boldsymbol{f} = \boldsymbol{J}\boldsymbol{p}$, so the residuals become $\boldsymbol{r} = \boldsymbol{f}^0 - \boldsymbol{J}\boldsymbol{p}$. If this expression for the residuals is substituted into eqn. (4), we obtain the so-called normal equations

$$\begin{aligned} &-2\boldsymbol{J}^T \boldsymbol{W} (\boldsymbol{f}^0 - \boldsymbol{J}\boldsymbol{p}) = \boldsymbol{0} \\ &\boldsymbol{J}^T \boldsymbol{W} \boldsymbol{J} \boldsymbol{p} = \boldsymbol{J}^T \boldsymbol{W} \boldsymbol{f}^0 \end{aligned} \tag{5}$$

However, in equilibrium systems, the models are non-linear and a method of successive approximations must be used. It is first necessary to estimate values for the parameters. Let the vector of parameter estimates be $\boldsymbol{p}^0$. Ideally these parameter estimates should be close to their optimal values, i.e. close to those values for which the gradient vector is zero. Each element of the gradient vector is now expanded as a first-order Taylor series about $\boldsymbol{p}^0$.

$$\begin{aligned} &g_j = g_j^0 + \sum_{k=1,n} \frac{\partial g_j}{\partial p_k} (p_k - p_k^0) \\ &\boldsymbol{g} = \boldsymbol{g}^0 + \boldsymbol{H}\boldsymbol{s} \end{aligned} \tag{6}$$

Since g is zero at the minimum, eqn. (6) can be written as

$$\boldsymbol{H}\boldsymbol{s} = -\boldsymbol{g}^0 \tag{7}$$

and the normal equations as

$$\boldsymbol{H}\boldsymbol{s} = 2\boldsymbol{J}^T\boldsymbol{W}\boldsymbol{r} \tag{8}$$

s is known as the shift vector. The matrix $\boldsymbol{H}$ is known as the Hessian and its elements are obtained by differentiating eqn. (4).

$$\frac{\partial g_j}{\partial p_k} = -\sum_h \sum_i W_{hi}\left(r_h \frac{\partial^2 f_i}{\partial p_j \partial p_k} + r_i \frac{\partial^2 f_h}{\partial p_j \partial p_k} - \frac{\partial f_h}{\partial p_k}\frac{\partial f_i}{\partial p_j} - \frac{\partial f_i}{\partial p_k}\frac{\partial f_h}{\partial p_j}\right) \tag{9}$$

$$C_{jk} = \sum_h \sum_i W_{hi}\left(r_h \frac{\partial^2 f_i}{\partial p_j \partial p_k} + r_i \frac{\partial^2 f_h}{\partial p_j \partial p_k}\right) \tag{10}$$

$$\boldsymbol{H} = 2\boldsymbol{J}^T\boldsymbol{W}\boldsymbol{J} - 2\boldsymbol{C} \tag{11}$$

On substituting eqn. (11) into the normal eqn. (8), we obtain eqn. (12), the classical Newton–Raphson equations.

$$(\boldsymbol{J}^T\boldsymbol{W}\boldsymbol{J} - \boldsymbol{C})\boldsymbol{s} = \boldsymbol{J}^T\boldsymbol{W}\boldsymbol{r} \tag{12}$$

The residual vector $\boldsymbol{r}$ represents the difference between the observed values and the values calculated with the initial parameter estimates.

If, now, we set the matrix $\boldsymbol{C}$ equal to zero (as it is in a linear system where $\boldsymbol{f} = \boldsymbol{J}\boldsymbol{p}$) we obtain eqn. (13), the normal equations of the Gauss–Newton method.

$$\boldsymbol{J}^T\boldsymbol{W}\boldsymbol{J}\boldsymbol{s} = \boldsymbol{J}^T\boldsymbol{W}\boldsymbol{r} \tag{13}$$

Solution of these equations yields the shift vector, s. If the first-order Taylor series expansion (6) is valid, then adding the shift vector to the vector of parameter estimates gives those parameters for whose values the gradient vector is zero. If the expansion is only approximate, the sum $s + \boldsymbol{p}^0$ becomes a new initial parameter estimate, and the process must be repeated iteratively. It continues until some convergence criterion is satisfied. At that point the objective function has its minimum value, $U_{\min}$.

Upon convergence, the variance–covariance matrix of the parameters is obtained by error propagation as

$$\boldsymbol{M}(p) = \frac{U_{\min}}{n-m}(\boldsymbol{J}^T\boldsymbol{W}\boldsymbol{J})^{-1} \tag{14}$$

Hence the parameter standard deviations and correlation coefficients are given by

$$\begin{aligned} \sigma_i &= [M(p)_{ii}]^{1/2} \\ \rho_{ij} &= [(M_{ij}/M_{ii}M_{jj})]^{1/2} \end{aligned} \tag{15}$$

(i) Concentrations of species in solution

Let there be a set of n_r reactants A, B, ... in solution. The total concentration, T_A^0, of reactant A at any point in a titration is given by the quantity (in mmoles if the volume is in ml) present initially plus the quantity added during the titration, divided by the total volume.

$$T_A^0 = \frac{(\text{initial quantity})_A + (\text{burette concentration})_A \times \text{volume added}}{\text{initial volume} + \text{volume added}} \tag{16}$$

There will be similar expressions for the other reactants. In the case of a coulometric titration, the total hydrogen ion concentration would be given by

$$T_H^0 = \frac{(\text{initial quantity})_H - (\text{rate of hydroxide mass generation}) \times \text{time}}{\text{volume}} \tag{17}$$

We use the convention that the total concentration of hydrogen ions in pure water is zero, and that adding hydroxide is equivalent to subtracting protons.

Let the reactants form n_k species in solution with stoichiometric coefficients $a, b, \ldots$

$$a\mathrm{A} + b\mathrm{B} + \ldots \rightleftharpoons \mathrm{A}_a\mathrm{B}_b\ldots \tag{18}$$

The concentration, C_j, of the jth species is defined in terms of the coefficients $a_j, b_j, \ldots$ and a cumulative formation constant β_j.

$$C_j = [\mathrm{A}_{a_j}\mathrm{B}_{b_j}\ldots] = \beta_j[\mathrm{A}]^{a_j}[\mathrm{B}]^{b_j}\ldots \tag{19}$$

Here, any electric charges carried by the species have been omitted for convenience. The total concentration of each reactant at equilibrium is constrained by a condition of mass balance.

$$\begin{aligned} T_A &= [\mathrm{A}] + \sum a_j C_j \\ T_B &= [\mathrm{B}] + \sum b_j C_j \\ &\ldots \quad \ldots \quad \ldots\ldots\ldots \end{aligned} \tag{20}$$

Note that, if the formation constants are known, the set of n_r eqns. (20) can be solved for the "free" concentrations [A], [B], ... and hence all the species concentrations can be obtained from eqn. (19).

Potentiometric measurements are made with ion-selective electrodes, such as the glass electrode which is sensitive to protons. The response of the electrode is generally assumed to follow a Nernstian type of law.

$$E = E^0 + \text{slope} \times \log_e[\mathrm{A}] \tag{21}$$

E is the measured potential, E^0 is the standard potential of the electrode and *slope*

is a factor equal to RT/nF in the ideal case. If the electrode gives pH readings, an arbitrary standard potential can be assumed.

The experimental data consists of burette readings with corresponding electrode potentials. At each point in the titration, the "free" concentrations that satisfy the conditions of mass balance are unknown, but calculable, given a set of formation constants. An alternative view is that an electrode reading gives the concentration of the corresponding free ion, and that the other free-ion concentrations are unknown.

A major problem is the derivation of initial estimates for the formation constants. This problem will not be discussed here. However, success in the refinement of formation constants is sometimes critically dependent on the quality of the initial estimates.

C. THE PROGRAMS LETAGROP AND SCOGS

LETAGROP [8–10] was the first computer program to be applied to general equilibrium problems. The method of minimization was the Newton–Raphson method (eqn. (12)). When the objective function is the sum of squared residuals in electrode potential, it is not immediately obvious how the derivatives $\partial U/\partial p_j$ can be calculated. Ingri and Sillén [9] therefore devised a method for approximating these derivatives by a difference formula

$$\frac{\partial U}{\partial p_j} \approx \frac{U(p_j+\delta)-U(p_j)}{\delta} \tag{22}$$

In LETAGROP, the "free" concentrations were found by solving the n_r non-linear equations of mass-balance (eqns. (20)).

The next general purpose program to be published was SCOGS [11–15]. In this program, the objective function was based on the titre volume; hydrogen ion concentration was taken from the electrode reading and the free concentrations of the other species were obtained by solving the other $n_r - 1$ mass-balance equations. Given the free concentrations and formation constants, one may calculate the total concentrations from eqns. (20) and hence the titre volume from eqn. (16). Once again, the derivatives, in this case $\partial v/\partial p_j$, were obtained by means of a difference formula.

D. THE PROGRAMS LEAST AND MINIQUAD

In view of the unsatisfactory nature of the numerical approximations for the derivatives required in least-squares minimizations of the objective function in LETAGROP and SCOGS, we looked for an objective function which would make it easy to give expressions for the elements of the Jacobian. The function chosen was the weighted sum of squared residuals in total concentration

$$U = \sum_{i=1,n} w_i[(T_A^0 - T_A)^2 + (T_B^0 - T_B)^2 \ldots] \tag{23}$$

T_A^0, T_B^0 are the concentrations calculated from eqn. (16), whilst T_A, T_B, ... are the concentrations calculated from eqns. (20). As implemented in LEAST [3], n_r may only be either 2 or 3. Let us look at the latter case, and denote the three reactants as A, B and H. The free proton concentration was taken from the electrode reading, leaving [A] and [B] unknown at each titration point. These unknown free concentrations were designated as parameters of the refinement. The required derivatives are found by differentiating the eqns. (20).

$$[\mathrm{B}]\frac{\partial T_A}{\partial [\mathrm{B}]} = \delta_{AB}[\mathrm{B}] + \sum a_j b_j C_j; \quad \delta_{AB} = 0 \text{ when } A = B, \text{ and } 1 \text{ when } A \neq B \tag{24}$$

$$[\beta_j]\frac{\partial T_A}{\partial [\beta_j]} = \sum a_j C_j \tag{25}$$

Note that the derivatives have been scaled by multiplying each one by the relevant parameter. This scaling serves two purposes: it gives the expressions for derivatives as linear combinations of the species' concentrations, and hence gives derivatives of comparable magnitude. The form of the Jacobian for $n_r = 3$ is shown in eqn. (26).

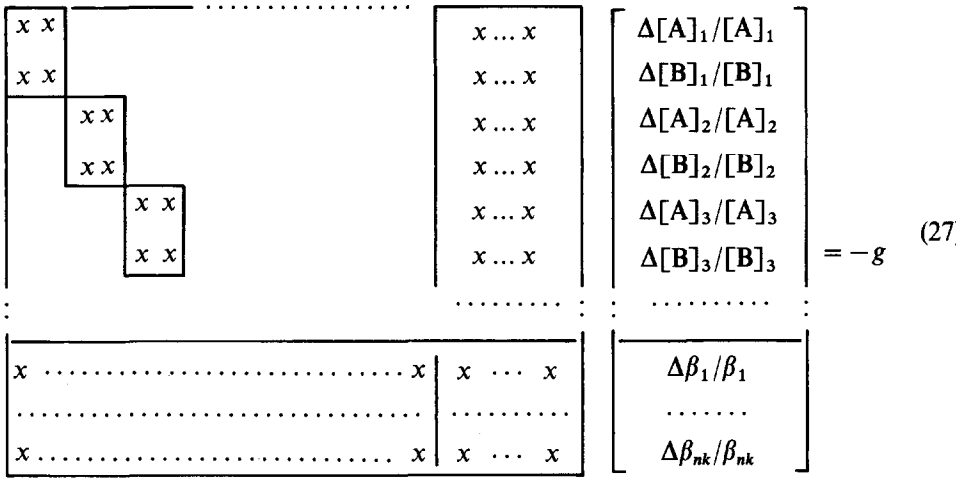
$$\begin{array}{|cc|cc|cc|c|}
\hline
a_1 & b_1 & & & & & g_{11} \cdots g_{n_k 1} \\
c_1 & d_1 & & & & & h_{11} \ldots h_{n_k 1} \\
e_1 & f_1 & & & & & i_{11} \ldots i_{n_k 1} \\
\hline
 & & a_2 & b_2 & & & g_{12} \cdots g_{n_k 2} \\
 & & c_2 & d_2 & & & h_{12} \ldots h_{n_k 2} \\
 & & e_2 & f_2 & & & i_{12} \ldots i_{n_k 2} \\
\hline
 & & & & a_3 & b_3 & g_{13} \cdots g_{n_k 3} \\
 & & & & c_3 & d_3 & h_{13} \ldots h_{n_k 3} \\
 & & & & e_3 & f_3 & i_{13} \ldots i_{n_k 3} \\
\hline
 & & & & & & \cdots\cdots \\
\hline
 & & & & & a_n \; b_n & g_{1n} \cdots g_{n_k n} \\
 & & & & & c_n \; d_n & h_{1n} \ldots h_{n_k n} \\
 & & & & & e_n \; f_n & i_{1n} \ldots i_{n_k n} \\
\hline
\end{array} \tag{26}$$

$a = [\mathrm{A}]\,\partial T_A/\partial[\mathrm{A}]$ $\quad b = [\mathrm{B}]\,\partial T_A/\partial[\mathrm{B}]$ $\quad g_j = \beta_j \partial T_A/\partial \beta_j$

$c = [\mathrm{A}]\,\partial T_B/\partial[\mathrm{A}]$ $\quad d = [\mathrm{B}]\,\partial T_B/\partial[\mathrm{B}]$ $\quad h_j = \beta_j \partial T_B/\partial \beta_j$

$e = [\mathrm{A}]\,\partial T_H/\partial[\mathrm{A}]$ $\quad f = [\mathrm{B}]\,\partial T_H/\partial[\mathrm{B}]$ $\quad i_j = \beta_j \partial T_C/\partial \beta_j$

Each derivative has a further subscript which indexes the titration point. The normal equations derived from this Jacobian are shown in eqn. (27).



(27)

The normal equations matrix is of order $2m + n_k$ and sparse. However, writing the normal equations in the block form

$$\begin{pmatrix} \boldsymbol{B}_1 & \boldsymbol{B}_2 \\ \boldsymbol{B}_2^T & \boldsymbol{B}_3 \end{pmatrix} \begin{pmatrix} \boldsymbol{s}_1 \\ \boldsymbol{s}_2 \end{pmatrix} = \begin{pmatrix} -\boldsymbol{g}_1 \\ -\boldsymbol{g}_2 \end{pmatrix} \tag{28}$$

the solution can be written as

$$\boldsymbol{s}_2 = (\boldsymbol{B}_3 - \boldsymbol{B}_2^T \boldsymbol{B}_1^{-1} \boldsymbol{B}_2)^{-1} (\boldsymbol{B}_2^T \boldsymbol{B}_1^{-1} \boldsymbol{g}_1 - \boldsymbol{g}_2) \tag{29}$$

and

$$\boldsymbol{s}_1 = -\boldsymbol{B}_1^{-1} (\boldsymbol{g}_1 + \boldsymbol{B}_2 \boldsymbol{g}_2) \tag{30}$$

The advantage of this solution is that it involves inversion of matrices of order 2 and n_k, which are much smaller than the full normal equations matrix.

Having implemented this algorithm, the next stage was to compare its performance with other programs [3]. Using three sets of experimental data, LEAST was compared with versions of SCOGS (SCOGS/1) and LETAGROP (LG/3). It was found that the numerical values of the parameters obtained using all three programs were within the limits of experimental error. Also, whilst LG/3 used about 60% of the time used by SCOGS/1, LEAST used about 30% of the time needed by LG/3. A version of LEAST which used the Newton–Raphson minimization was compared with a version using the Gauss–Newton minimization; it was found that the Newton–

Raphson method gave identical results but took 2–4 times as much time. Weights proportional to $(1/f^0)^2$ were also investigated; weighting had minimal effect. It was concluded that LEAST employed a sound algorithm.

After examining a pseudo-Newton method [16], we decided to use LEAST as a basis for a general-purpose program, which we called MINIQUAD, from the Italian MINImi QUADrati, meaning least-squares [4]. LEAST was re-written, and extended in various ways. The number of reactants allowed was 2–5. Either one or two electrodes, either one reading in pH or EMF could be used. The possibility of a coulometric titration was introduced. Mathematically, the minimization was protected against divergence by a process of optimizing the length of the shift vector. The program was tested on a variety of experimental data before being released; it then became widely adopted as the standard program for refinement of formation constants. Let us summarize its salient features.

(a) All the partial derivatives needed in the algorithm are obtained from analytical expressions. This removes the errors inevitably introduced into the computations by the process of numerical differentiation, and eliminates the problem of having to choose an increment δ for the formula (22). Also, numerical differentiation requires two sets of calculations, for $E(p+\delta)$ and $E(p)$, so any program using this procedure is expected to take at least twice as much time.

(b) The parameters refined are all the unknown quantities, including those "free" concentrations not determined potentiometrically.

(c) The objective function minimized involves all the mass-balance equations with equal weight, so that no one reagent is treated as different from the others. The total concentrations of all reactants are considered to be affected by experimental error.

(d) The use of scaled derivatives in the Jacobian confines the elements of the Jacobian to a limited interval, which confers numerical stability on the solution of the normal equations.

(e) It is possible, in principle, to analyse systems containing any number of reactant species, using any number of electrodes. Formation constants for mononuclear, polynuclear or hydrolysed complexes can be refined.

(f) Upon convergence, the residuals are plotted out, and can be analysed statistically to evaluate the goodness of fit. Species concentration plots may also be produced.

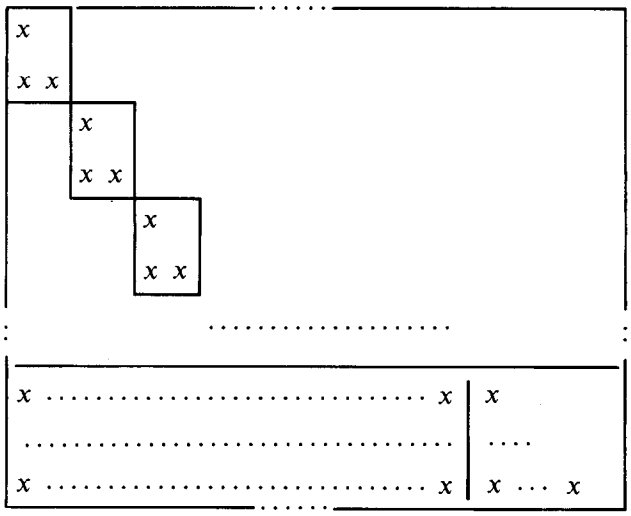
(g) The program was written in a "portable" subset of FORTRAN so that it could be compiled on any computer. The plots referred to above were line-printer plots chosen because they do not need any special plotting device. Because the program was designed for "export", it travelled well to many different laboratories.

E. MINIQUAD75

We found two ways in which to improve the algorithm used in MINIQUAD. Firstly, it was felt that the process of refinement of the unknown free concentrations

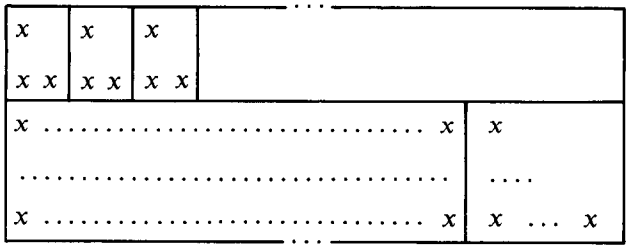
was wasteful. The free concentrations were eliminated as parameters as follows. The shift vector for the formation constants was calculated as before (eqn. (29)) and the formation constants were updated, but instead of calculating shifts for the free concentrations (eqn. (30)) the n_c unknown free concentrations were calculated by fitting the n_r mass-balance equations by the Gauss–Newton method. In this way, the unknown free concentrations are optimal for the new set of formation constants. It was found that this procedure resulted in a much faster convergence, since the unknown free concentrations approached their ultimate values much more quickly.

Secondly, the protection against divergence was improved by replacing the shift-length optimization by a rotation procedure based on a Marquardt parameter, which is now recognized by most numerical analysts as a superior procedure. The key to implementing this lay in the recognition that the normal equations (27) could be solved without having to consider the elements which are equal to zero. The Choleski factorization of the normal equations matrix, $\boldsymbol{L}\boldsymbol{L}^T = \boldsymbol{B}$ leads to factors which have the form



(31)

which can be stored in the compact form



(32)

Addition of the Marquardt parameter to the diagonal elements of the normal equations matrix makes no difference to the form of the factorization, so that the modified normal eqn. (33) can be solved directly.

$$(\boldsymbol{J}^T\boldsymbol{W}\boldsymbol{J} + \lambda\boldsymbol{D})\boldsymbol{s} = \boldsymbol{J}^T\boldsymbol{W}\boldsymbol{r} \tag{33}$$

λ is the Marquardt parameter, and $\boldsymbol{D}$ holds the diagonal elements of $\boldsymbol{J}^T\boldsymbol{W}\boldsymbol{J}$. The Marquardt parameter was chosen efficiently by using Fletcher's strategy [17].

These improvements are complementary. The first works well when the current estimates of the formation constants are close to the ultimate values, and the second works well when the current estimates are bad estimates. We incorporated both in the program MINIQUAD75 [5] by alternating a refinement cycle using the first improvement with five cycles using the second. This resulted in a running speed 2–3 times faster, together with greater tolerance to poor estimates of the formation constants.

F. THE PROGRAM BEST

One criticism that can be levelled at the programs described so far is that the refinements are not based on the residuals in the observed quantities, volume and pH, though SCOGS uses residuals in volume. This appears to be a motivation behind the program BEST [18], which minimizes the sum of squared deviations in pH. This program has a number of interesting features. The residuals are assigned weights equal to $(1/\Delta\text{pH})^2$, which, assuming the volume increments are equal, is equivalent to weighting the points in inverse proportion to the slope of the titration curve. The minimization is based on direct incremental variation of the parameters, thus avoiding the use of derivatives of the objective function altogether, at the expense of refinement speed. The program is highly interactive, and also permits the refinement of quantity for use in those cases that the substance cannot be purified. Listings on both paper and disk are included in ref. 2, as are details on the various options.

G. THE PROGRAM SUPERQUAD

We should assume that both titre volume and pH (or electrode potential) are subject to random error, since they are both experimental measurements. No published program hitherto included this assumption. In the light of this, we have sought to develop an alternative refinement procedure based on the objective function

$$U = \sum_{h=1,m} \sum_{i=1,m} W_{hi}(E_h^0 - E_h)(E_i^0 - E_i) \tag{34}$$

The problem which has to be solved before this function can be minimized by the Gauss–Newton method is how to obtain the derivatives $\partial E_i/\partial\beta_j$. The formation constants do not appear in the Nernstian eqn. (21), so implicit differentiation must

be used. Sillén and co-workers [8–10] adopted a numerical method for obtaining derivatives.

The derivatives of electrode potential with respect to any parameter, x, may be obtained as follows. Differentiation of eqn. (21) with respect to E^0 yields the derivative directly; differentiation with respect to a general parameter, x, yields eqn. (35).

$$\frac{\partial E}{\partial x} = \frac{\partial E}{\partial [\mathrm{A}]} \frac{\partial [\mathrm{A}]}{\partial x} = \frac{\text{slope}}{[\mathrm{A}]} \frac{\partial [\mathrm{A}]}{\partial x} \tag{35}$$

The partial derivative $\partial[\mathrm{A}]/\partial x$ is obtained [19] by solving the set of simultaneous equations

$$\begin{pmatrix} [\mathrm{A}]\dfrac{\partial T_\mathrm{A}}{\partial [\mathrm{A}]} & [\mathrm{B}]\dfrac{\partial T_\mathrm{A}}{\partial [\mathrm{B}]} \cdots \\ [\mathrm{A}]\dfrac{\partial T_\mathrm{B}}{\partial [\mathrm{A}]} & [\mathrm{B}]\dfrac{\partial T_\mathrm{B}}{\partial [\mathrm{B}]} \cdots \\ \cdots & \cdots \end{pmatrix} \begin{pmatrix} \dfrac{x}{[\mathrm{A}]}\dfrac{\partial [\mathrm{A}]}{\partial x} \\ \dfrac{x}{[\mathrm{B}]}\dfrac{\partial [\mathrm{B}]}{\partial x} \\ \cdots \end{pmatrix} = \begin{pmatrix} -x\dfrac{\partial T_\mathrm{A}}{\partial x} \\ -x\dfrac{\partial T_\mathrm{B}}{\partial x} \\ \cdots \end{pmatrix} \tag{36}$$

The coefficients on the left-hand side are found, as before, from eqn. (24). The right-hand side takes different forms according to the nature of the parameter. If it is a formation constant, eqn. (25) is used. Alternatively, the parameter may be initial quantity or burette concentration in which case the derivative is obtained by differentiating eqn. (16). The unknown free concentrations are found by solving the n_r eqns. (20) of mass balance. The solution consists of an iterative Newton–Raphson refinement, using

$$\begin{pmatrix} [\mathrm{A}]\dfrac{\partial T_\mathrm{A}}{\partial [\mathrm{A}]} & [\mathrm{B}]\dfrac{\partial T_\mathrm{A}}{\partial [\mathrm{B}]} \cdots \\ [\mathrm{A}]\dfrac{\partial T_\mathrm{B}}{\partial [\mathrm{A}]} & [\mathrm{B}]\dfrac{\partial T_\mathrm{B}}{\partial [\mathrm{B}]} \cdots \\ \cdots & \cdots \end{pmatrix} \begin{pmatrix} \dfrac{\Delta [\mathrm{A}]}{[\mathrm{A}]} \\ \dfrac{\Delta [\mathrm{B}]}{[\mathrm{B}]} \\ \cdots \end{pmatrix} = \begin{pmatrix} \Delta T_\mathrm{A} \\ \Delta T_\mathrm{B} \\ \cdots \end{pmatrix} \tag{37}$$

It should be noted that the matrix of coefficients in (37) is the same as that in (36); consequently, once the free concentrations have been refined, all the quantities required to calculate the derivatives of potential with respect to the parameters are immediately available.

The weights given in the objective function (34) are calculated according to the "rigorous" method of least squares [7] as follows. We assign constant values to the errors in potential, and in volume, denoted by σ_E and σ_V, respectively. Assuming that there is no correlation between the errors at different titration points, the variance–covariance matrix is block-diagonal, the order of the blocks being the

number of electrodes used. Hence, when two electrodes are used each block has the form

$$\begin{pmatrix} \sigma_E^2 + (\partial E/\partial v)^2 \sigma_V^2 & (\partial E/\partial v)(\partial F/\partial v)\sigma_V^2 \\ (\partial E/\partial v)(\partial F/\partial v)\sigma_V^2 & \sigma_F^2 + (\partial F/\partial v)^2 \sigma_V^2 \end{pmatrix} \tag{38}$$

Where E and F are the two electrode potentials. With one electrode, the variance of the observation is given by $\sigma_E^2 + (\partial E/\partial v)^2 \sigma_V^2$. This formula can also be derived from a consideration of the propagation of the titre error onto the measured potential. The slope $\partial E/\partial v$ is calculated by numerical differentiation of the experimental titration curve. It should be noted that in BEST, the weights are obtained by dropping the term σ_E^2.

These procedures were incorporated in the program SUPERQUAD [6], so named because it SUPERcedes miniQUAD. SUPERQUAD is an extremely powerful program for formation constant refinement, which can treat data from almost any conceivable kind of potentiometric titration. It offers a number of facilities not previously available, such as the refinement of initial quantity, burette concentration and standard electrode potential. Furthermore, constraints can be applied to the parameters, something which is particularly useful when dealing with substances of unknown purity. Another way in which SUPERQUAD differs is that formation constants are allowed to assume negative values during a refinement. The idea underlying this is that, if a refinement converges to a point where one or more formation constants are negative, then it is the model which is wrong, not the refinement process.

The complete solution of a chemical equilibrium problem requires that a model be identified which gives an adequate representation of the experimental data, within experimental error. This model also specifies the concentrations of all the species present in solution in any given set of experimental conditions. The refinement of the formation constants is only a stage in the more general problem of species selection. SUPERQUAD is the only program which incorporates an automatic species selection procedure. It works as follows. The data includes a "basis set" of formation constants which includes all which might be needed in the final model. Each formation constant is associated with a refinement key which may be 1 (refinable), 0 (constant) or -1 (ignored) in any given refinement. If the result of a refinement is a negative formation constant, its key is set to -1 and the new model is automatically refined. Otherwise, a new model is specified by supplying a new set of keys. A record is kept of the best model in terms of the refinement keys and the value of the objective function, which is, of course, the lowest attained with the models examined to that point. As mentioned above, models with negative formation constants are automatically rejected, as are models in which one or more formation constants is ill-defined, which we take to mean as having a standard deviation of more than 1/3 of its value.

Since the objective function (34) is based directly on measured quantities, with a weight matrix (approximately) equal to the inverse of the variance–covariance matrix of the observations, theory shows that $(U/(n-m))^{1/2}$ has an expectation value of 1 [7]. This expectation value is independent of the probability distribution function of the random experimental errors, though it implies both that the model is correct and that there are no systematic errors in the data. We have found that values close to one can be achieved, and that therefore one has an immediate criterion on which to judge whether a model is acceptable. However, this criterion is critically dependent on the precision of the error estimates σ_E and σ_V, and on the assumption that these errors are independent of volume or measured potential.

H. THE PROGRAM HYPERQUAD

As mentioned in the introduction, there is a need to be able to treat simultaneously data obtained by different experimental techniques; ideally the different sets of data should be complementary. Some programs have been published which can treat both potentiometric and spectrophotometric data [20]. Considering the advantages of SUPERQUAD over other programs used to reduce potentiometric data, we have decided to generalize the algorithm so that the data may be potentiometric and/or spectrophotometric.

Experimentally, we have connected a spectroscopic flow cell to the titration apparatus so that, after each addition of titrant, we may measure both the electrode potential and the optical density (at selected wavelengths) of the solution.

The principles of the new program, HYPERQUAD, which is currently under development, are the same as those of SUPERQUAD. The objective function for combined data is given in the equation

$$U = \boldsymbol{r}_E^T \boldsymbol{W}_E \boldsymbol{r}_E + \boldsymbol{r}_D^T \boldsymbol{W}_D \boldsymbol{r}_D \tag{39}$$

where $\boldsymbol{r}_E$ is a vector of residuals in EMF and $\boldsymbol{r}_D$ is a vector of residuals in optical density. The optical density, D, is calculated from the Beer–Lambert law

$$D = d \sum \varepsilon_j C_j \tag{40}$$

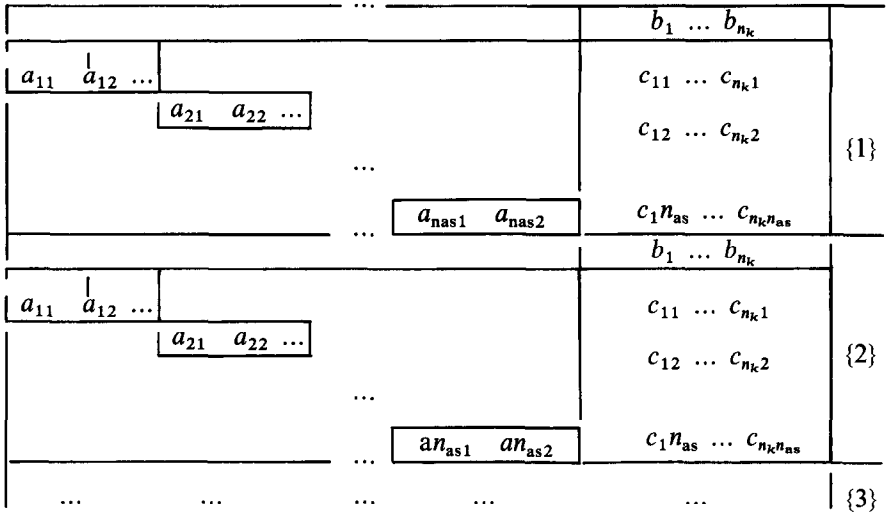
The summation is taken over all the absorbing species, and d is the optical path length. ε_j is the molar extinction coefficient of the jth species, whose molar concentration is C_j. Equation (40) may be expanded as

$$D = d\left[\sum \varepsilon_j \beta_j [A]^{a_j} [B]^{b_j} \ldots + \varepsilon_A [A] + \varepsilon_B [B] \ldots\right] \tag{41}$$

The parameters to refine are now both the formation constants, $\boldsymbol{\beta}$, and those molar extinction coefficients, $\boldsymbol{\varepsilon}$, which are unknown. It can be seen that the optical density, D, is a non-linear function of the parameters so that the unknown extinction coefficients must be refined. If spectrophotometric measurements are taken at more

than one wavelength, as is common practice, the unknown extinction coefficients must be determined at every wavelength; it is therefore important to choose only those wavelengths where the absorbance is most sensitive to changes in a species' concentration.

There are now three kinds of non-zero entry in the Jacobian, $\partial E/\partial\beta$, $\partial D/\partial\beta$, and $\partial D/\partial\varepsilon$. For the case of one electrode, the Jacobian has the form



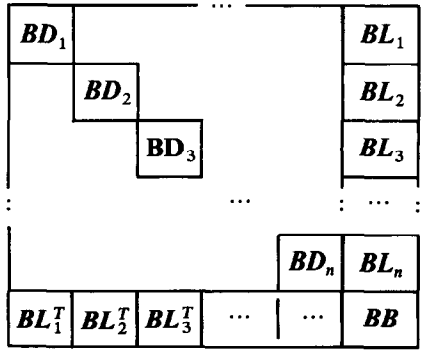
(42)

$a_{ij} = \varepsilon_j \partial D_i/\partial\varepsilon_j$

$b_k = \beta_k \partial E/\partial\beta_k$

$c_{ik} = \beta_k \partial D_i/\partial\beta_k$

The numbers at the right-hand side are the titration point numbers. The form of the normal equations matrix is shown in (43).



(43)

The blocks ***BD*** are square, of order equal to the number of refinable extinction coefficients, whilst ***BB*** has the order of the number of formation constants to be

refined. The form of the normal equations is similar to (27), and can be written as

$$\begin{pmatrix} \boldsymbol{BD} & \boldsymbol{BL} \\ \boldsymbol{BL}^T & \boldsymbol{BB} \end{pmatrix} \begin{pmatrix} s(\varepsilon) \\ s(\beta) \end{pmatrix} = \begin{pmatrix} g(\varepsilon) \\ g(\beta) \end{pmatrix} \tag{44}$$

$s(\varepsilon)$ is the vector of shifts on the ε parameters and $s(\beta)$ is the vector of shifts for the formation constants. $g(\varepsilon)$ is the portion of the gradient vector relating to the extinction coefficients, and $g(\beta)$ relates to the formation constants. The solution of these equations has the same form as (29) and (30).

$$s(\beta) = [\boldsymbol{BB} - \boldsymbol{BL}^T \boldsymbol{BD}^{-1} \boldsymbol{BL}]^{-1} [\boldsymbol{BL}^T \boldsymbol{BD}^{-1} g(\varepsilon) - g(\beta)] \tag{45}$$

BD is easy to invert because of its block-diagonal structure. The vector of formation constants is then upgraded and, rather than calculate shifts on the extinction coefficients by solving (44), the latter are found by a linear least-squares calculation based on (41). This procedure is similar to that used in MINIQUAD75 in that it constrains the extinction coefficients to be optimal for the current set of formation constants. Nevertheless, it should be noted that the shifts for the formation constants are calculated from a complete set of normal equations; in other programs such as SQUAD [21] the block ***BL*** is effectively set to zero. We believe that using the full set of normal equations will lead to faster convergence, and it will certainly give better estimates of the standard deviations on the parameters.

Since the potentiometric and spectrophotometric measurements have intrinsically different precision, it is of fundamental importance to have good estimates of all the experimental errors. The estimate σ_E is still valid with regard to the electrode potentials. The errors on the optical densities will depend very much on the spectrophotometer employed, and will, in general, depend both on the wavelength and on the actual optical density of each measurement. Moreover, since both types of measurement are made on the same solution, the "rigorous" method of least-squares requires that the weight matrix takes correlation of errors into account in dealing with a potentiometric/spectrophotometric titration.

I. CONCLUSIONS

The use of computer programs to reduce experimental data has resulted in great advances being made in the study of solution equilibria. The majority of research workers in the field now use MINIQUAD or SUPERQUAD. The success of these programs is no doubt due to the fact that they were designed by experimentalists for experimentalists. They have opened up many opportunities for new kinds of experimental work, though the option to use two electrodes has not been as fruitfully applied as we had hoped. This is due to the poor quality of most ion-selective electrodes with respect to the glass electrode.

In this review we have concentrated on the mathematical aspects of the data

reduction. The data itself must be of the highest quality, with careful attention given to all aspects of experimental procedure such as purification of materials and calibration of electrodes. Also, the results must be interpreted with great caution, particularly if "dangerous" parameters such as quantity of substance have been refined. In the final analysis, the results must be both mathematically and chemically satisfactory.

In developing HYPERQUAD, we hope to improve the species selection process by using the power of spectrophotometry to discriminate between species on the basis of where and how much they absorb light together with the power of potentiometry to establish the extent of protonation as a function of pH.

REFERENCES

1 D.J. Leggett (Ed.), Computational Methods for the Determination of Formation Constants, Plenum Press, New York, 1985.
2 A.E. Martell and R.J. Motekaitis, Determination and Use of Stability Constants, VCH Publishers, New York, 1988.
3 A. Sabatini and A. Vacca, J. Chem. Soc. Dalton Trans., (1972) 1693.
4 A. Sabatini, A. Vacca and P. Gans, Talanta, 21 (1974) 53.
5 P. Gans, A. Sabatini and A. Vacca, Inorg. Chim. Acta, 18 (1976) 239.
6 P. Gans, A. Sabatini and A. Vacca, J. Chem. Soc. Dalton Trans., (1985) 1195.
7 P. Gans, Data Fitting in the Chemical Sciences, Wiley, Chichester, 1992.
8 D. Dyrssen, N. Ingri and L.G. Sillén, Acta Chem. Scand., 15 (1961) 694.
9 N. Ingri and L.G. Sillén, Ark. Kemi, 23 (1964) 97.
10 P. Brauner, L.G. Sillén and R. Whiteker, Ark. Kemi, 31 (1969) 365.
11 D.D. Perrin and I.G. Sayce, Talanta, 14 (1967) 82.
12 I.G. Sayce, Talanta, 15 (1968) 1397.
13 I.G. Sayce, Talanta, 18 (1971) 653.
14 I.G. Sayce and V.S. Sharma, Talanta, 19 (1972) 831.
15 D.D. Perrin and H. Stunzi, in D.J. Leggett (Ed.), Computational Methods for the Determination of Formation Constants, Plenum Press, New York, 1985, Chap. 4.
16 P. Gans and A. Vacca, Talanta, 21 (1974) 45.
17 R. Fletcher, UKAEA Rep. AERE-R 6799, H.M.S.O., London, 1971.
18 R.J. Motekaitis and A.E. Martell, Can. J. Chem., 60 (1982) 2403.
19 I. Nagypal, I. Paka and L. Zekany, Talanta, 25 (1978) 549.
20 H. Gampp, M. Maeder and A.D. Zuberbuhler, Talanta, 27 (1980) 1037.
21 D.J. Leggett, in D.J. Leggett (Ed.), Computational Methods for the Determination of Formation Constants, Plenum Press, New York, 1985, Chap. 6.