Coordination Chemistry Reviews, 27 (1979) 195—222

COMPUTER EVALUATION OF COMPLEX EQUILIBRIA

F. GAIZER

Department of Chemistry, College of Food Technology, Hódmezővásárhely (Hungary)

(First received 14 September 1977; in revised form 7 November 1978)

CONTENTS

SYMBOLS

a	Neutralization value [26,63]
a_{H^+}	Hydrogen ion activity
A	Light absorption reduced to 1-cm layer thickness
β_{qp}	Overall formation constant of the species M_qL_p: $\beta_{qp} = [M_qL_p] \times [M]^{-q} \times [L]^{-p}$
d	Difference between measured and calculated values
D	Distribution ratio [127]: $D = [M]_{org}/[M]_{aq}$
E	Electromotive force
ϵ	Molar extinction coefficient
η	Log $([M]/T_M)$
H_NL	N-basic acid
K	Stepwise equilibrium constants or acid dissociation constant
$[L]$	Free ligand concentration
$[M]$	Free metal ion concentration
M_qL_p	General formula of complex species

n	Number of experimental points
$\bar{n}$	Average coordination number [16]: $\bar{n} = (T_L - [L])/T_M$
N	Maximum number of ligands
p	Number of ligands in a complex species
q	Number of metal ions in a complex species
(q,p)	Abbreviated notation for the composition of a complex species
Q	Reaction heat
r	Number of hydrogen ions in a complex species
S	Sum of the squares of residuals between calculated and measured values (error-square sum)
σ	Standard deviation
σ^2	Variance: $\sigma^2 = S/(n - 1 -$ number of parameters)
$\sigma(\beta)$	Error (standard deviation) of parameter
T_H	Total acid concentration
T_L	Total ligand concentration
T_M	Total metal ion concentration
w	Weighting factor
Z	Average number of protons released (taken up): $Z = ([H^+] - T_H)/T_M$
$\hat{}$	(Hat) calculated value of a quantity

A. INTRODUCTION

Electronic computers have been used for the calculation of complex equilibria for nearly 20 years. The first programmes, which have not been surpassed even today, were published at the beginning of the 1960's. Since then, more than 100 publications have appeared on original or variously modified computer procedures. Accordingly, at present the computer calculation of equilibrium and other constants can be regarded as solved in virtually every field of complex chemistry, and in several fields it is even possible to choose between a number of reliable programmes. Nevertheless, comparatively few papers have been published in which experimental results have been subjected to computer procedures. One of the causes of this is, in all probability, the fact that surprisingly few reviews have appeared on the application of computers in complex chemistry [1,2]. Another source that may be mentioned is Gans' review [3] on the numerical methods for data-fitting problems. In the present monograph an account is given of those computer procedures reported in the literature up to about 1976, and the practical observations relating to the evaluation procedures, including the difficulties, are dealt with. The treatment is confined to the computer evaluation of the results of the most frequently used potentiometric, pH-metric, polarographic, spectrophotometric, calorimetric, solvent-extraction and ion-exchange examinations.

B. EQUILIBRIUM STUDIES

The primary aim of equilibrium studies is to establish the compositions and equilibrium constants of species formed during the association or dissociation of the same or different ions and/or molecules. A number of excellent books discuss the constants used to characterize the stabilities of the various possible species (mononuclear, polynuclear, mixed ligand, protonated or deprotonated, hydroxo complexes and their combinations), and the experimental methods suitable for the determination of these. The present treatment is based primarily on the works of Beck [4] and Rossotti and Rossotti [5], where the chemical origin of the computational problems can be consulted.

Complex stability constants and other parameters (e.g. molar extinctions, formation enthalpies, etc.) can be calculated from functions which are relatively simple from a mathematical viewpoint. The total concentrations of the reacting components (T_M, T_L, T_H, etc.) are in general known quantities, and are the independent variables in the basic functions. The dependent variables of the functions in certain methods are the free species ([M], [L], [H^+], etc.), while in other procedures they are possibly these, together with the measured intensity values (e.g. extinction, etc.). The difficulties in the solution of the functions, i.e. in the calculation of the unknown values of the parameters, usually stem from the fact that some of the characteristic quantities of the complex formation cannot be measured, or not with sufficient accuracy.

Solution of the complex formation functions essentially falls into the sphere of regression analysis, and in very many cases treatment of the functions can be carried out with methods elaborated for regression analysis [6—13], or with standard computer programmes prepared on the basis of these. Although this solution has been employed by several authors, a far greater number favour those procedures which are based on treatments of the equations describing the given concrete problem that are relatively simpler mathematically, but can be followed more readily by the chemist.

C. COMPUTER PROCEDURES FOR EVALUATION OF COMPLEX EQUILIBRIA

The computer evaluation procedures employed in complex chemistry can be classified in accordance with a number of aspects. The most useful of these classifications is that in accordance with the mathematical method utilized in the programmes. This is essentially the classification followed in the present discussions, although it had to be dispensed with in some cases as it would have led to the creation of too many groups. A further compromise was necessary in that certain programmes could have been treated in several groups.

(i) Non-statistical programmes

(a) Procedures based on computer solution of equation systems. The most simple means of calculating the parameters was proposed by Thamer [14], and first applied successfully by Roth and Bunnett [15]. In this method five suitable experimental points are chosen on the A vs. [H^+] curve obtained

during study of the acid H_2L, and the extinction conditions are written for each point. Solution of the resulting simultaneous equations gives the values of the dissociation equilibrium constants K_1 and K_2 and the molar extinctions ϵ_L, ϵ_1 and ϵ_2. Naturally, constants may be calculated [17] with this principle on the basis of other functions too, e.g. the following Bjerrum function [16]

$$\sum_{j=0}^{N} (\bar{n} - j) \times \beta_j \times [L]^j = 0 \qquad (1)$$

($\beta_0 = 1$). However, constants obtained from only a few points in this way can not be regarded as reliable values. The procedure of Litchinsky et al. [18] is an attempt to eliminate this drawback of the Thamer principle. The randomly selecting sub-routine of the programme of these authors forms numerous equation system combinations by utilizing every experimental point (e.g. 14—20 from 100 experimental points). The authors take the final values of the parameters to be the arithmetical means of the individual parameter values obtained by solution of the equation systems.

One of the programmes of Romary et al. [19] is also based on solution of the Bjerrum equation (1). The procedure of Hendrickson [20] for the calculation of acid association constants and the stability constants of metal complexes of polyprotonic ligands is essentially the solution of simultaneous equations by a matrix inversion technique. The same method is used in the SPANA (Spektralanalyse) programme of Kaden and Zuberbühler [21], which calculates the molar extinctions of the species in a system containing at most 20 complexes. When the SPANA programme is combined with the authors' VARIAT programme, a full spectrophotometric evaluation can be carried out.

With the method of Chattopadhyaya and Singh [22] the results of spectrophotometric measurements can be evaluated in the case of the formation of a single complex species of composition ML. The method is based on solution of quadratic equations obtained by combination of the extinction values measured for each experimental point, the material balance equations and the measured molar extinctions of the reacting components. One of the roots of the equation gives the value of $\epsilon_0 (= \epsilon_{ML} - \epsilon_M - \epsilon_L)$ for each experimental point. The unknown β_1 can readily be calculated with the average of the ϵ_0 values. Considerable similarity to this method is to be found in the procedure reported by Siefker [23], likewise for one complex, which was also elaborated for calculation of the stability constants of the complexes ML, ML_2 or ML_3.

(b) Computer solution of graphical procedures. Some authors have constructed programmes to computerize well-known and successfully employed, but at times very laborious graphical procedures. Such a method is the procedure of McMasters and Schaap [24] for the evaluation of polarographic measurements, which in all probability is the first published computer procedure for the calculation of complex equilibria. The programme, following the method of Rossotti and Rossotti [25] and yielding 4 stability constants, first calculates

approximate values of all the constants, and then using the appropriate region of the data, calculates the constant for the next largest by successively assuming the constants for the smaller species.

On the basis of the graphical evaluation method of Schwarzenbach et al. [26], Datta and Grzybowski [27] developed a computer procedure for calculating the stability constants of the complexes ML and ML_2 from potentiometric measurements. The method is based in essence on the solution of pairs of equations obtained via the variables X and Y, calculated from the experimental values $\bar{n}$ and [L]; in the event of n experimental points, this results in $n/2$ values of K_1 and K_2. As regards the treatment of the constants found in the individual solutions, this programme is in effect a statistical procedure. The weighting takes into account the fact that the most reliable constants are obtained when the straight lines representing the two equations are perpendicular to each other. The weighting factor then has the maximum value of 1, while its value is proportionally lower for intersection at angles less than 90°. The final values of the constants are obtained by a further weighting, which depends on the extent to which the values of the constants resulting from the previous weighting differ from the average value.

The method was extended by Grzybowski et al. [28] to the determination of the dissociation contants of polybasic acids (e.g. citric acid) which dissociate in an overlapping manner. The SOLVIN procedure of their programme solves the equations in pairs, while similarly to the preceding programme the WTMEAN sub-routine computes the weighted mean values of the constants.

Thun et al. [29] developed a procedure involving computer treatment of the Fronaeus function [30]. The programme computes the integral terms of the function by parabolic integration from three consecutive points, and yields the values of the individual constants for each experimental point. The programme was applied to calculate four complex products or acid dissociation constants. Compared with the original Fronaeus method, the computer procedure has the special advantage that the errors do not accumulate, as extrapolation is performed only once.

Some authors have also used computers to calculate the curves of the graphical evaluation procedures. Examples are the GRAPE programme of Johansson et al. [31] and the PSEUDOPLOT programme of Corrie et al. [32], which facilitate calculation of the normalized functions. Gorton and Jameson [33] studied the fit of machine-computed p[L] vs. $\bar{n}$ curves and their experimental data. Evaluations with machine-calculated curves were similarly made by Wozniak and Nowogrocki [34].

(c) Some simpler programmes. A procedure with a short running time was developed for the Bjerrum function [16] by Barcza [35]. From the individual $\bar{n}$ and [L] value pairs the programme computes the values of the complex products via the equation

$$\beta_i = \sum_{\substack{j=0 \\ j \neq i}}^{N} \{(\bar{n} - j) \times \beta_j \times [\mathrm{L}]^j\} / \{(i - \bar{n}) \times [\mathrm{L}]^i\} \tag{2}$$

using only those experimental points for which $|\bar{n} - i| \geqslant 0.05$, and for which β_i has an effect on the value of $\bar{n}$.

The programme of Nagypál et al. [36] is based on a curve-reduction procedure elaborated by development of the Bjerrum formation function. Its essence is that not only the central metal ion M, but also the complexes ML, ML_2, etc. are treated as central species, for each of which a separate $\bar{n}_j$ vs. [L] formation curve can be written.

$$\bar{n}_{j+1} = \frac{\bar{n}_j^*}{1 - \kappa_j} - 1 \tag{3}$$

The individual values in this equation can be calculated via the Fronaeus equation [30], and from these the individual stepwise equilibrium constants can be obtained. The elaborated programme has the particular advantage that it permits establishment of the measurement interval(s) for which the basic assumption does not hold, i.e. the formation merely of mononuclear complexes, and those intervals in which the formation of hydrolyzed and/or mixed complexes, etc. must be reckoned with.

(ii) Procedures based on the least-squares principle

The vast majority of computer procedures developed for evaluation of the results of complex equilibrium investigations are based on the least-squares principle. These methods in some way calculate the "best" values of the desired parameters, i.e. those values for which the weighted or non-weighted sum of the squares of the residuals of the measured and calculated quantities is minimum. Of the methods used to calculate equilibria, these have the greatest importance, mainly because several of them can also be employed as general methods of evaluation. In addition to the programmes to be discussed below, there are some which will not be dealt with in detail, but are merely mentioned for the sake of completeness [37—41].

(a) Approximations with linear and polynomial equations. A relatively simple means of calculating the parameters is fitting with linear or polynomial functions. The special advantage of these methods is that the evaluation can also be performed with computer library programmes, while the errors in the parameters can also be calculated from the measure of the fit.

A linear approximation was employed by Ramette [42] to calculate the dissociation constant of a species of composition HL and the molar extinctions of the species L^- and HL. The programme searches for the linear equation best fitting the experimental values by varying the molar extinction. The method can also be applied to a complex ML if [M] can be measured and the molar extinction of ML is zero.

In their PRČEK (= DWARF) programme, Havel and Kuban [43] similarly use an approximation to linearized equations with the best fitting straight line. Although the method was applied for spectrophotometric evaluation of the dissociation process $H_NL \rightleftharpoons H_{N-q}L + qH^+$, other linearized equations can also be treated with the programme. The programme can be used to calculate the correct q value, the equilibrium constant of the process and the molar extinctions. It has been employed to carry out a number of successful evaluations [44,45]. The results obtained agreed well with those of the LETAGROP-SPEFO programme [46] although the values of S and $\sigma(A)$ proved somewhat larger than those given by the LETAGROP. The PRČEK programme is applicable to the formation of 2 species, or 3 species in the event of exceptionally favourable extinction conditions.

Briggs and Stuehr [47,48] utilized linearized acid—base titration curves to prepare a programme for calculation of the dissociation constants of at most tribasic acids. They use an iteration procedure and optimize to $\Sigma(\Delta \log a_{H^+})^2$.

The FAJAF 45 programme of Cumming et al. [49] calculates with linear transformations of the Bjerrum function [15]. The values of the parameters can be calculated as intercepts and/or slopes. The authors carry out the minimization with the following function

$$S = \sum_{i=1}^{n} [w_{X_i} \times (X_i - \hat{X}_i)^2 + w_{Y_i} \times (Y_i - \hat{Y}_i)^2 + 2w_{X_iY_i} \times (X_i - \hat{X}_i) \times (Y_i - \hat{Y}_i)] \tag{4}$$

where X_i and Y_i are the coordinates of the individual experimental points, and $\hat{X}_i$ and $\hat{Y}_i$ are the coordinates of the i-th point on the line fitted, closest to the i-th experimental point.

The constants of the species ML and ML_2, and HL and H_2L, can be determined with the FAJAF 45 programme.

It should be mentioned that the standard procedure of regression analysis for establishment of the equation of the best fitting straight line was applied by Tsentovskii and Tsentovskaya [50] to evaluate conductivity measurements.

Calculation of constants from the coefficients of polynomial equations describing the experimental results offers very broad applicability. Fundamental and extensive investigations with this method were first made by Rydberg and Sullivan [51—55]. These authors used a programme [52] calculating the coefficients a_n of a polynomial of formula

$$y = \sum_{n=0}^{N} a_n x^n \tag{5}$$

to evaluate the results of solvent-extraction experiments. Calculation of the β_n values of the Leden function [56], the Fronaeus function [30] and the Olerup function [57] is also possible on this principle.

A number of authors, primarily Varga [17] have made use of the Rydberg—Sullivan method. It was employed by McBryde [58] in his NLMAX programme for the evaluation of spectrophotometric equilibrium measurements. The experiments were carried out according to the method of corresponding solutions, as in this way the Bjerrum formation curve could be constructed even without a knowledge of the molar extinctions. Heys et al. [59] also used the same method to calculate the pK values of polybasic acids from spectrophotometric measurements.

A programme involving fitting to a power series was used by Rush et al. [60] to evaluate the results of their hydrolysis studies. Bad'yanov et al. [61] also made approximations with polynomials.

Some authors have utilized various library programmes and sub-routines for evaluation. An example is the method of Irving and Stacey [62], based on the least-squares solution of polynomials with a standard sub-routine. The equation to be solved is the Bjerrum equation (1), expanded by taking into consideration the dissociation of the ligand. The method of evaluating pH-metric measurements is in effect a procedure requiring 3—12 iterations; as the condition for termination, the authors chose the constancy of the standard deviation [63].

Momoki et al. [64] used a Taylor expansion series approximation of the DeFord—Hume equation [65] to evaluate polarographic measurements. An S4801 sub-routine solving simultaneous equations was employed, whereby the variance could also be calculated.

Irani and Callis [66] calculated the values of acid dissociation constants with a modified least-squares non-linear programme on an IBM-704 computer. An IBM-704 library programme was similarly used by Rabideau and Moore [67] to evaluate potentiometric equilibrium measurements. Vinogradov [68] used an IBM-1620 library programme constructed for the least-squares solution of general polynomial equations in order to approximate to a polynomial-transformed equation describing extinction.

(b) The Gauss method. One of the relatively simple and frequently applied methods of refining the parameters is the Gauss method [69—71]. Mathematically, this means the determination of the values of the parameters β_j in a function of the type

$$y = f(x; \beta_1, \beta_2, \ldots, \beta_N) \qquad (6)$$

In eqn. (6) y is some measured intensity value, and x the independent variable of the experiments (or a combination of these). The weighted sum of the squares of the residuals of the measured and calculated intensity values can be given by means of the following equation:

$$S = \sum_{i=1}^{n} w_i[y_i - f(x_i; \beta_1, \beta_2, \ldots \beta_N)]^2 = \sum_{i=1}^{n} w_i[y_i - y_i^0]^2 \qquad (7)$$

The task is to establish the best parameter values giving the minimum of the function for S, starting from the initial parameter values β_j^0. The conditions $\partial S/\partial\beta_j = 0$ ($j = 1, \ldots, N$) result in a total of N equations of the following type.

$$-2w_i[y_i - f(x_i; \beta_1, \beta_2, \ldots, \beta_N)]\partial f_i/\partial\beta_j = 0 \qquad (8)$$

The Gauss method approximates the function to be determined by a Taylor series with terms higher than first order neglected. When the appropriate substitutions are made, this approximation results in a normal equation system, which can be more easily surveyed and treated if written in matrix form.

$$\mathrm{G^TWG}\overset{\downarrow}{\mathrm{h}} = \mathrm{G^T}\overset{\downarrow}{\mathrm{d}} \tag{9a}$$

h in this equation is the vector of the parameter corrections, and can be obtained by carrying out the following matrix operations.

$$\overset{\downarrow}{\mathrm{h}} = (\mathrm{G^TWG})^{-1}\mathrm{G^TW}\overset{\downarrow}{\mathrm{d}} \tag{9b}$$

In eqn. (9), $\mathrm{W} = \mathrm{diag}(w_1, w_2 \ldots w_n)$, d is the error vector; its elements are the differences of the intensity values (y_i^0) calculated with the starting values of the parameters (β_j^0) and those measured at the individual experimental points (y_i). The elements ($\partial f_i/\partial \beta_j$) of the matrix G can be calculated by analytical or numerical differentiation; with the latter method it is most simple to increase the initial values of the parameters to be refined by appropriate small amounts (Δt_j) (always only one at a time), and intensity values (y'_{ij}) are calculated for every experimental point. The values of the individual matrix elements are given by $g_{ij} = (y'_{ij} - y^0)/\Delta t_j$. The SCOGS programme [72] uses the $g_{ij} = (y^+_{ij} - y^-_{ij})/2\Delta t_j$ elements (y^+_{ij} and y^-_{ij} are obtained by using the increments $\pm\Delta t_j$). This requires an extra calculation of y, but it provides improved precision of the numerical derivatives. Some problems occurring in the solving of eqn. (9b) are discussed in Gans' review [3].

The error in the individual parameters, $\sigma(\beta_j)$, can be calculated from the diagonal elements of the matrix $(\mathrm{G^TG})^{-1}$ and from the standard deviation calculated with the corrected parameter values, according to the following equation

$$\sigma(\beta_j) = [|g_{jj}^{\mathrm{inv}}| \times S/(n - 1 - \mathrm{NP})]^{1/2} \tag{10}$$

where NP is the number of refined parameters.

The Gauss method began to be applied in complex chemistry following the work of Moore and Zeigler [73]. Since then it has been used to solve very many types of problem, and several variants of the method are now known.

The method was probably earliest used by Rabideau and Kline [74] for the calculation of one equilibrium constant and two molar extinctions. Tobias and Hugus [75] studied the hydrolysis of tin(II) chloride via potentiometric measurement of the tin(II) activity. For computer evaluation the normal equations were produced by series expansion of the material-balance equation expressing the total tin concentration, while the minimization was performed with regard to the sum of the weighted squares of the residuals of the measured and calculated total tin concentrations. By evaluating pH-metric measurements, Tobias and Yasuda [76] calculated dissociation constants and complex products of hydrolyzed and non-hydrolyzed species. In the programme, minimization was performed for the total hydrogen ion concentration, and numerical differentiation was employed. Another early work was the programme of Rush et al. [77], with which Sillén [78] also carried out a comparative evaluation. Lansbury et al. [79] calculated the stability constants of the complexes ML and ML_2 by utilizing the Bjerrum function (1). Rosseinsky and Kellawi [80] performed the evaluation of spectrophotometric equilibrium measurements. The basic equation of light absorption was employed to construct the normal equations. With their programme they established the molar extinction and formation constant of a single species.

The general applicability of the Gauss method was pointed out by Gaizer; together with his colleagues, he has so far carried out evaluation of potentiometric [81,82], pH-metric [83,84], solubility [82] and spectrophotometric [71,85—87] measurements with one and the same algorithm. If eqn. (9) is examined, it can be seen that in the case of a filled matrix G and error vector **d**, the correction values of the parameters can be calculated, independently of what measurements the values of the individual elements originate from. The procedure used in the author's SPEFCA programme [71] has the further great advantage that it simplifies evaluation with equations that are difficult to treat mathematically. The parameters are treated as array elements, and are refined by the input of their serial numbers; in principle it is possible to solve very complex problems with any number of parameters.

The general applicability of the Gauss method was also pointed out by Karlson and Vestin [88]; the LINREG sub-routine of their GENREG (= Generalized Regression) programme is a regression procedure applicable to any problem. The use of the programme was presented on the results of a pH-metric study of the hydrolysis of thorium(IV).

An unusual and interesting application of the Gauss method is the procedure of Lingane and Hugus [89], by which spectrophotometric measurements at a number of wavelengths can be evaluated. This method also takes into account the correlation between the analytical errors. As the error-square sum, the authors take the squares of the weighted residuals of three terms: T_H, T_L and A. Utilizing the normal equations used by Lingane and Hugus, Magnell [90] constructed a programme (SPECTRO-1130) which is suitable for the treatment of 3 species, with measurements made at 5 wavelengths.

The Gauss method was further developed by Wentworth [91], thereby allowing the errors of both coordinates to be taken into consideration. Although in this way the solution of the problem is mathematically somewhat more complex than that of the original method, two possibilities were suggested for a relatively simple solution. Following the work of Box [92], Wentworth also proposed optimization of the vector containing the correction values of parameters as elements. This means a search for the scalar r_m, whereby, on multiplication with the correction vector, the greatest decrease of S can be achieved. The calculation of r_m essentially demands calculation of three S values: with the basic values of the parameters, and with the parameters modified by half of the correction values and by the total correction values.

The application of the Wentworth method was presented by the author and his colleagues by means of a spectrophotometric problem involving several wavelengths [93].

Perrin et al. developed their GAUSS programme [94—96] from the programme of Tobias and Yasuda [76] by incorporating the Wentworth modification. The GAUSS programme is suitable for computing acid dissociation constants, the stability products of metal complexes and hydroxo complexes, or combinations of these, and it may even be employed in the case of polynuclear species. In another paper [95] the authors reported their SCOSUS (= Stability Constants of Single Unknown Species) programme, elaborated for calculation of the stability constants of mixed complexes of types MLL′ and ML_2L'.

A programme developed in the Perrin school was published by Sayce [72,97]. The Sayce SCOGS (= Stability Constants of Generalized Species) programme is suitable for the calculation of the constants of as many as 20 species of type $M_qM'_{q'}L_pL'_{p'}(OH)_r$. This programme has been taken over by a very large number of research groups, and has even been subjected to certain modifications [98,99]. The most significant modification appears to be that of Nagypál [99]; the essence of this is that not only the free species ([M], [L], $[H^+]$), but any complex species may be treated as components of a complex in the programme. By this means the special role of the hydrogen ion disappears, and hence the programme becomes suitable for the calculation of constants from polarographic, distribution and spectrophotometric equilibrium measurements. The SCOGS programme was modified by Leggett and McBryde [100] to make it suitable for the evaluation of spectrophotometric measurements. Their SQUAD (= Stability Quotients from Absorbance Data) programme is convenient for the calculation of the stability products of species of any composition formed from two different metal ions and two different ligands, if the formation of the complexes is pH-dependent.

An interesting variant of the Gauss method was elaborated and employed by Sabatini and Vacca in their LEAST programme [101], which is suitable for the calculation of the stability constants of complexes of type $M_qL_pH_r$, on the basis of pH-metric measurements. An interesting feature of the LEAST programme is that the unknowns [M] and [L] of the material balance equations are treated with the desired parameters at the same level. In addition, in place of the somewhat less accurate and slower numerical differentiation, it employs analytical differentiation. The two variants of the programme, the GN (= Gauss—Newton) and the NR (= Newton—Raphson), both have very favourable running times. The NR variant also takes into account the $\partial^2 f_i/(\partial x_j \times \partial x_k)$ values.

By further development of the LEAST—GN programme, Sabatini et al. [102] constructed the MINIQUAD (= Minimi Quadrati) programme, which in principle is suitable for the treatment of complexes formed from any number of species, even in the case of measurements with several potentiometric electrodes. The MINIQUAD programme was later modified [103] in order to enhance the speed and to make the convergence more reliable. In one modification, after a refinement only the correction of the β values is performed, and not that of the free species; the values of these latter giving the minimum of the squares of the residuals for all the material balance equations are calculated with a separate sub-routine. The other modification was the introduction of the Marquart parameter [104], which rotates the correction vector in the direction of steepest descent. The authors did not find either modification in itself to be perfect, and therefore both are used in the improved variant of the programme, the MINIQUAD 75 [103].

(c) Procedures based on examination of the square of residuals function. Besides the general application of the Gauss method, establishment of the position of the minimum in $S = f(\beta)$ is the most general method for calculation of the "best" values of the parameters. In the following section, two examples of this will be presented.

Simple searching procedures. The simplest possibility for establishing the position of the minimum in the S function is by the systematic variation of the values of the unknown parameters. A common feature of programmes operating in this way is the fairly long running time, but they have the advantage that the finding of the minimum in S is guaranteed, even in the event of very complex functions or very uncertain starting parameter values.

One group of the programmes uses only simple value changes, by a constant amount, in the range of parameter values to be expected. Such approaches are to be found in the programmes of Unvin et al. [105] and Meites and Meites [106]. The former was used to calculate the parameters of the Bjerrum function [16], and the latter to evaluate spectrophotometric and pH-metric measurements.

The principle of the programmes in the other group is that the value of the parameter sought is varied in one direction so long as the value of S decreases; it is next varied in the opposite direction, but in smaller steps. This is then repeated until a certain condition for termination is fulfilled (e.g. the value of the increment of the parameter decreases to a given value). Such a method was employed by Conrow et al. [107] to calculate the stability constant and molar extinction of a single species from spectrophotometric measurements. The similar, SWING programme of Nagano and Metzler [108] is suitable for the calculation of the formation constants of the species H_2L^{2+}, HL^+ and LOH^- on the basis of the spectra recorded as a function of pH in a broad wavelength range. This programme varies only the values of the stability constants, and uses a least-squares method to calculate the molar extinctions. The JOBCON programme of Likussar [109] is applicable in the case of formation of a single complex of moderate stability, in the event of spectrophotometric and solvent-extraction equilibrium measurements in accordance with the Job method. The essence of the new procedure of Likussar and Boltz [110] is the search for the maximum absorbance $A_{\max}$ which it would be possible to measure assuming non-dissociation of the complex formed. The programme searches for this value essentially by a least-squares procedure.

A very effective direct search method, known as Powell 64 [111], in which each parameter is varied in turn, has been applied successfully in cases where other methods failed [112].

Mention should also be made here of the LEASK programme of Sarkar and Kruck [113], which in effect is the least-squares solution of simultaneous equation systems by iteration. This programme optimizes the individual stability products in turn. The VARIAT (= Variation von Konstanten) programme of Kaden and Zuberbühler [21] operates in the same way; this is suitable for the calculation of the constants of mixed ligand, polynuclear and hydrolyzed (protonated) species from potentiometric measurements.

The "pit-mapping" technique. The approach developed by Sillén et al. [114—116] is based on examination of the function $U = f(\mathbf{k})$ *. For $N = 1$ it is a parabola, for $N = 2$ this function is an elliptic paraboloid, while for $N > 2$ it is an $(N + 1)$-dimensional space. The essence of the method (called by the authors a "pit-mapping" procedure) is that the initial values $(\mathbf{k}_c)$ of the parameters are varied by appropriate steps h_i, the U values being calculated at a total of $\frac{1}{2}(N + 1)(N + 2)$ points. By this means the shape of the surface is mapped out as it were, and is next approximated with an $(N + 1)$-dimensional second-order surface. The coordinates of the pit, the deepest point of the U function, are computed from the coefficients of the equations describing the approximate surface. As emerges from the above, the Sillén method is a completely general optimizing procedure, which can also be employed in the case of inhomogeneous parameters (e.g. β values, molar extinctions, concentrations of stock solutions, etc.).

In the course of the use of their first, LETAGROP (= search for the pit) programme [116], Ingri and Sillén found that the U surface is very frequently not symmetrical; indeed, at times it is extremely distorted. With such cleft-like surfaces, difficulties arose in the calculations. When the selected steps h_i were too large, the U values lay too high on the wall of the surface, and terms higher than the second-order ones, neglected in the approximation, became appreciable. In the case of small h_i values, on the other hand, reliable correction values could not be calculated for the parameters because of the rounding errors in the computer. The procedure was therefore modified so that the values of the parameters were varied along the main axis of the pit in the second and subsequent refinement cycles. This is achieved with the aid of a twist matrix. This variant of the LETAGROP programme, which is still used at present, is known as the LETAGROP—VRID ** programme [117] (vrida = twist, turn).

We shall not dwell at length on an account of the mathematical bases of the programme published by Ingri and Sillén [118], but merely refer to the original publications [115—117]. We present only an abbreviated version of

---

* The original notations are followed in this section.

** This nomenclature is not used in this publication.

the calculation procedure followed in the programme, performing the simultaneous refinement of three parameters, but in the experience of the present author this can be applied very extensively if the individual parameters are treated as array elements [71].

In the case $N = 3$, calculation of the following U values is necessary: U_c, U_1, U_{-1}, U_2, U_{-2}, U_3, U_{-3}, $U_{1,2}$, $U_{1,3}$ and $U_{2,3}$ (the subscript gives the serial number of the parameter varied by the increment h_j, while the sign denotes the direction of the variation). The constant series k, with which each U value can be calculated, can be produced with the following equation.

$$\overset{\downarrow}{\mathbf{k}} = \overset{\downarrow}{\mathbf{c}} + \mathbf{SH}\overset{\downarrow}{\mathbf{v}} \tag{11}$$

The vector **c** in the equation contains the starting values of the parameters to be refined as elements. In its principal diagonal the step matrix **H** contains the increments h_i of the parameters to be refined; the other elements of the matrix have values of zero. The twist matrix **S** is a triangular one, in which the values of the elements under the principal diagonal are zero, while those of the elements in the principal diagonal are 1; in the first refinement cycle the values of the other elements are 0, while in the subsequent cycles they assume values depending on the symmetry of U. The variational vector **v** directs the variation in the values of the parameters. The values of its elements are either 0, +1 or —1. Having the set of U values, it is necessary to calculate a vector p

$$p_i = 0.5(U_{-i} - U_i)$$

and also the elements of the following matrix R.

$$r_{ii} = 0.5(U_i + U_{-i}) - U_c$$

$$r_{ij} = r_{ji} = 0.5(U_{ij} - U_c) + (p_i + p_j) + 0.5(r_{ii} + r_{jj})$$

By formation of the vector $\mathbf{v}_0 = \mathbf{pR}^{-1}$, the value of the minimum in the function U is

$$U_0 = U_c - \mathbf{p}\overset{\downarrow}{\mathbf{v}}_0 \tag{12}$$

while the vector of the improved parameter values is

$$\overset{\downarrow}{\mathbf{k}}_c = \overset{\downarrow}{\mathbf{c}} + \mathbf{SH}\overset{\downarrow}{\mathbf{v}}_0 \tag{13}$$

The value of the error $\sigma(k_j)$ in the individual parameters is

$$\sigma(k_j) = [|a_{jj}^{\text{inv}}| \times U_0/(n - N)]^{1/2} \tag{14}$$

where a_{jj}^{inv} are the values of the diagonal elements of the matrix $\mathbf{A}^{-1} = \mathbf{SHR}^{-1}(\mathbf{SH})^{\mathrm{T}}$.

If the improved parameter values are regarded as starting values, a new refinement cycle may be carried out, but for this it is necessary to calculate the upper right elements of the twist matrix. If this new matrix is denoted by

S' then

$$\mathbf{S}' = \mathbf{SHWH}^{-1} \tag{15}$$

In eqn. (15), similarly to **S** the matrix **W** is a triangular matrix, the elements of which can be calculated with the elements of the matrix **R**

$$w_{12} = -r_{12}/r_{11}$$

$$w_{13} = (r_{12}r_{23} - r_{13}r_{22})/D$$

$$w_{23} = (r_{13}r_{21} - r_{23}r_{11})/D$$

where $D = (r_{11}r_{22} - r_{12}r_{21})$.

If S' is taken in place of S in eqn. (11), the new refinement cycle may be carried out.

In the ALGOL variant [118] of the LETAGROP programme, published in 1964, Ingri and Sillén introduced several modifications and expanded the programme. Of these changes, only a few of the more essential ones will be mentioned here. With one of the modifications [119] the running time of the programme was reduced. Another extension [120,121] allows treatment of the parameters on two levels, i.e. as parameters relating only to the individual point groups, and as parameters relating to every point. A further very important extension is the species selector procedure [122], which serves to select the model(s) formally best describing the experimental data.

A picture of the main applications to date of the LETAGROP programme is given in Table 1.

Very many experiments have been successfully evaluated with the LETAGROP programme. It has generally been applied without change, but publications have also appeared on some modifications. It was modified by Ulmgren and Wahlberg [123] to take into account the diffusion potential and the activity constants. It has been employed by Meloun and Čermak [124] to evaluate chelatometric titrations in the presence of metal indicators, and it has also been adapted for small-memory computers [125].

(d) Selection of the quantity to be minimized and weighting of the experimental points in least-squares procedures. A fundamental criterion of the least-squares treatment is the correct calculation of the sum of the squares of the residuals. In the linear least-square case (when the parameters are not in exponential or logarithmic expressions), two decisions must be taken: what should be the quantity with which the error-square sum is calculated, and with what weight should the individual experimental points be taken into account?

The quantity to be minimized must primarily be one which can be measured with satisfactory accuracy, and which characterizes well the process of complex formation. A survey of the quantities selected for minimization in the individual programmes is given in Table 1. It can be seen that in the majority of the programmes refinement is performed for the sum of the squares of the residuals of only a single quantity (one of the coordinates). Some authors

TABLE 1

Quantities chosen for minimization in least-squares programmes

Author(s) and name of programme	Reference	Experimental method and measured quantity	Minimized quantity	Weighting [a]
Rydberg and Sullivan	[51,52]	Distr. $[M]_{org}/[M]_{aq}$	D	+
Irving and Stacey	[62]	pH	a	
Rabideau and Moore	[67]	Pot., [M]	E	
Tobias and Hugus	[75]	Pot., [M]	T_M	
Tobias and Yasuda	[76]	pH	T_H	+
Sillén et al.,	[115]	pH	Z	
LETAGROP	[118]	Pot., [M]		
		Pot., [L]	$\bar{n}$	
	[121]	Pot.	E	
	[46]	Spef.	A	
	[138]	Spef. and pot.	$A + E$	
	[126]	Cal.	Q	
	[127]	Distr.	$\log D$	
	[137]	Vapour pr.		
Conrow et al.	[107]	Spef.	A	
Inman et al.	[128]	Pot., [M]	$T_M/[M]$	+
Lansbury et al.	[79]	pH	$\bar{n}$	+
Paoletti et al.	[129]	Cal.	Q	
Vacca et al.	[130]	pH	E	
Perrin and Sayce	[94]	pH	T_H	+
Rosseinsky and Kellawi	[80]	Spef.	A	
Prasad and Peterson	[131]	Spef.	A	
Lingane and Hugus	[89]	Spef.	$T_M + T_L + A$	+
Gaizer,	[85—87]			
SPEFCA	[71,156]	Spef.	A	
	[81,82]	Pot., [M]	E	
	[83,84]	pH	E	
Gould and Sutton	[132]	Pot.	$\bar{n}$	
Sayce, SCOGS	[72]	pH	Titrant vol.	
Sabatini and Vacca, LEAST	[101]	pH	$T_M + T_L + T_H$	
Cumming et al. FAJAF 45	[49]	Pot., $\bar{n}$ — [L]	X and Y normalized variables	+
Karlson and Vestin	[88]	pH	Z	
Zur Nedden et al.	[133]	pH	T_{OH}	
Massoumi et al.	[134]	pH	α	
Márton and Inczédi	[135]	Ion exchange	D	
Gans and Vacca, STEW	[136]	pH	T_H	
Sabatini et al. MINIQUAD	[102]	Pot.	All material balance equations	
Gans et al. MINIQUAD 75	[103]			
Leggett and McBryde, SQUAD	[100]	Spef.	A	

[a] + means the weighting $w_i \neq 1$.

use in their programme the "rigorous" least-square treatment of experimental data. The term rigorous is applied to those problems where the independent variable is also subject to experimental errors. The first such programme was worked out by Wentworth [91]. In the programme of Lingane and Hugus [89] the minimization of spectrophotometric measurements is carried out with regard to the error-square sum of three quantities: T_M, T_L and A. Cumming et al. [49] consider that in a linear plot the experimental points are scattered along both coordinates; the standard deviation varies from point to point, but in addition the difference may attain several orders of magnitude. Hence, the authors carry out the minimization with eqn. (4).

The importance of the weighting of the individual experimental points was emphasized by the authors of even the earliest publications. There is no doubt that in certain cases failure to weight the experimental points or incorrect calculation of the values of the weighting factors may lead to completely false constants. Following the work of Deming [8], Wentworth [91] pointed out that one of the frequent errors arises on linearization of the functions, when the new variable obtained by the linearization may take on some weird relationships with the observed variable. In such a case it is absolutely necessary to assess the need for weighting.

Most authors have taken over the use of the weighting factors $w_i = 1/\sigma_i^2$ used in mathematical statistics. This was done by Rydberg and Sullivan [53] in their pioneering computer procedure for the evaluation of the results of their solvent-extraction experiments. They found that one of the sources of the error σ_i in their measured distribution ratio is the error $\sigma_D(\text{I})$ arising from the impulse count, while another is $\sigma_D(\text{L})$ arising from the measurement of the free ligand concentration. On the basis of a χ^2 test, they found that $\sigma_i = 2\sigma_D(\text{I}) + 0.75\ \sigma_D(\text{L})$. In addition, they used the percentage error estimated from the standard deviation of D in the distribution ratio. This weighting was also taken over by Varga et al. [139,140].

Calculation of the weighting factors of the individual experimental points in accordance with the theory of mathematical statistics [8] was first proposed and employed in complex chemistry by Hugus [141]. It would be difficult to establish the variances of the individual experimental points in complex chemical measurements by repeating the experiments several times. Instead of this, Hugus regards another method of error calculation [8] as suitable; this is based on the approximation that the change ΔF in the value of a function $F(x,y)$ as a result of the errors Δx and Δy in the variables is proportional to the first partial differentials of the function with respect to the variables.

$$\Delta F = (\partial F/\partial x)\Delta x + (\partial F/\partial y)\Delta y \tag{16}$$

The Hugus weighting was also employed by Varga and Hume [139]. In a later publication [17], Varga used another calculation, setting out from the consideration that the error in potentiometric measurements can be ascribed to both $\bar{n}$ and [L]. The Varga standard deviation values are $\sigma(\bar{n}) = P\bar{n}$ and

$\sigma([\mathrm{L}]) = P[\mathrm{L}]$, where the quantity P is the fractional error in the solution series.

Lansbury et al. [79] similarly take the reciprocal of the variance of the experimental points as the weighting factors; for the calculation of these, only the estimated errors in the electromotive force and total ligand concentration measurements are taken into consideration, while the errors in the weight measurement, electrode calibration, etc. are regarded as negligible. The following equation is obtained for the values of the variances of the experimental points

$$\sigma_i^2 = 2P^2\sigma^2(E) + Q^2\sigma^2(T_\mathrm{L}) \quad (17)$$

where the quantities P and Q can be calculated via the experimental quantities T_M, T_L and [L] and the stability products. Since these latter are not known, an iteration procedure is employed, starting from approximate values calculated from the half-value points of the Bjerrum complex formation function.

The rigorous method of Cumming et al. [49], used in the FAJAF-45 programme, gives the weighting factors of the individual experimental points, according to the following equation.

$$w_i = 1/[r^2\sigma^2(X_i) + 2s\,\mathrm{cov}(X_iY_i) + s^2\sigma^2(Y_i)] \quad (18)$$

The quantities r and s in eqn. (18) are the coefficients of the linearized equations, $\sigma^2(X_i)$ and $\sigma^2(Y_i)$ are the variances of the coordinates, and $\mathrm{cov}(X_iY_i)$ is the covariance of the coordinates. The independent variable generally has a much smaller error than the observed intensity. When the expressions have been rearranged and the dependent and independent quantities are mixed together as in the case of $\bar{n}$ methods, the pK values may be erroneous by more than one pK unit, if an improper weighting is used. The values of the constants obtained with this procedure agree well with the results obtained with the LETAGROP programme; this is also confirmation of the correctness of the simpler unit-weighting employed in the LETAGROP programme.

In the case of pH-metric titration, Tobias and Hugus [75] refined with respect to $\Sigma(\Delta T_\mathrm{H})^2$, using the Hugus [142] weighting. Perrin and Sayce [94] evaluated their studies of a similar nature both with weighting factors calculated according to Hugus [141,142], and with unit-weighting, and found that the values of the constants obtained with the unit-weighting factors corresponded within the limits of error with those obtained with the Hugus weighting. Therefore, they do not consider the use of weighting factors to be necessary. However, their finding can certainly not be generalized, but can only be utilized in the case of measurements in a fairly narrow pH interval. Neglect of the weighting in refinement with respect to $\Sigma(\Delta T_\mathrm{H})^2$ is considered by Sillén [115] to be definitely erroneous.

In the Sayce SCOGS programme [72], the minimization is performed with unit weighting factors with respect to $\Sigma(\Delta v)^2$, the square of the residuals of the titrant solution volume. The results of calculations refined with respect to $\Sigma(\Delta v)^2$ and to $\Sigma(\Delta \mathrm{pH})^2$ were compared by Nagypál [99], who demonstrated that a substantial difference between the constants obtained with the two minimization procedures is to be expected only if the pH vs. v titration curve is not linear. Nagypál proposed combined minimization, i.e. refinement with respect to the square of the perpendicular residuals. Δm can be calculated from the pH vs. v curve from simple geometrical considerations, according to the following formula.

$$\Delta m = \Delta v \Delta \mathrm{pH}/[(\Delta v)^2 + (\Delta \mathrm{pH})^2]^{1/2} \qquad (19)$$

Sabatini and Vacca [101] consider a good estimation of the variances of the observations to be a basic requirement for a statistically correct weighting; in the case of pH-metric titrations, for example, this demands the accurate estimation of the measurement errors in the total concentrations, the volume and the electromotive force. To take the errors in so many quantities into consideration automatically makes the calculation more difficult. This is further enhanced by the fact that, although the individual errors in these quantities (e.g. the standard potentials of the electrodes, etc.) are random, they appear as a systematic error in the individual experimental curves. The authors therefore rather recommend the use of unit weighting factors. They also carried out evaluation with weighting independently of these, when (by analogy with force-constant determinations, and sometimes, with crystallographic calculations) they made calculations with $w_i = 1/f_i^0$ values (f_i^0 are the quantities measured in the individual experiments). This weighting had otherwise already been used earlier by Momoki et al. [64] in the evaluation of polarographic measurements.

Finally, it should be mentioned that Inman et al. [128] calculated in the case of potentiometric measurements the following weighting factors.

$$w = \frac{1}{(T_\mathrm{M}/[\mathrm{M}])^2} \times \frac{g^2}{2.303^2} \times \frac{1}{\Delta E} \qquad (20)$$

($g = 59.16$ mV at 25°C)

(e) Some other evaluation procedures. Apart from those already presented, programmes operating on different mathematical bases have also been reported in the literature. Gans and Vacca [136] used the Davidon—Fletcher—Powell (DFP) approximation [143—145] in their STEW programme. The **h** vector of the correction values of the parameters can be calculated with the following equation

$$\overset{\downarrow}{\mathbf{h}} = -\mathbf{H}\overset{\downarrow}{\mathbf{g}} \qquad (21)$$

where the matrix **H** is a positive approximation of the inverse of the matrix built up of the elements $g_{ij} = \partial U^2/(\partial\beta_i\partial\beta_j)$, while the values of the elements of the vector g are $g_j = \partial U/\partial\beta_j$. The performance of the STEW programme corresponds to that of the Sabatini LEAST programme.

In a successfully applied variant of the DFP method [146] the gradient vector **g** and the inverse Hessian $\mathbf{H}^{-1}$ are built up numerically.

Feldberg et al. [147] used their programme to evaluate experimental spectrophotometric results on mixed ligand complexes. The method employed, essentially the steepest descent one, is based on a survey of the vicinity of the minimum of the error-square sum function, and eliminates that difficulty in the Sillén method which arises in the event of too flat or too asymmetric surfaces.

Kaneda and Martell [148] developed a new statistical method for the calculation of the constants of protonated and/or deprotonated mononuclear complexes from pH-metric measurements. They defined the function

$$Y = \Sigma(\Delta K_i)^2 = f(K_2, K_3, ..., K_N) \tag{22}$$

where K_i is the difference of the individual K_i values from the mean value. The programme essentially minimizes this function for the individual constants $K_2, K_3, ..., K_N$.

On the basis of the principle of internal linearity, Griffith and Potts [149] elaborated a new method for the calculation of equilibrium constants and enthalpies ΔH from digital spectra. In the system examined they studied the equilibrium of two species.

The computer method of Kankare [150] is suitable for evaluation of the results of spectrophotometric examinations at 27 wavelengths. In the programme, the method of factor analysis is employed to calculate the extinction matrix and the values of the equilibrium constants. The final values of the constants of 6 complexes were established, as were their errors, with the Chandler [151] STEPIT programme.

Finally, some methods will be mentioned which give information on the number of species present in an examined system, and possibly on their compositions. With the method of Vadasdi [152], the number and compositions of the species M_qL_p formed in a two-component system can be established. The applicability of this valuable method is somewhat restricted by the fact that a knowledge of both [M] and [L] is necessary for construction of the matrix equations.

Wallace and Katz [153] proposed a simple method to establish the rank of an extinction matrix (arranged in the form of a matrix of extinctions measured at several wavelengths in a solution series); the rank of the matrix gives the number of light-absorbing species. This method has also been used by Varga and Veatch [154], who plotted the matrix data as a three-dimensional contour map, and in this way checked the rank of the matrix.

Chriswell and Schilt [155] developed three possibilities for the identification of weak complexes, based on the mole ratio method. The methods of these authors gave good results even in cases where the customary procedures proved ineffective.

D. CONCLUSIONS

Computerized methods of evaluation have become very important procedures in the calculation of complex equilibria. With their aid a virtually unlimited number of experimental data can be processed, and the results can be evaluated in accordance with several models, thereby strengthening the reliability of the conclusions. These methods also permit the treatment of complex systems which with the known numerical and graphical methods would not be evaluable, or only with difficulty. Further, some of the procedures of computer evaluation also yield valuable information on the experimental data, this not being given by other methods.

Of the elaborated procedures for computer evaluation, the statistical programmes are of the greatest importance, partly because these are suitable for the solution of more problems, and partly because they provide more information.

In the past decade the preparation of larger and more general programmes has been the predominant feature in the construction of programmes. This undoubtedly has numerous advantages, such as their general utility. With the SCOGS [72] or MINIQUAD [102,103] programmes, for example, the majority of potentiometric problems can be solved. A further advantage is that if sophisticated techniques are used in them these programmes become very efficient. A problem has arisen, however, in that the làrge programmes have become difficult to understand and even baffling for the chemist, and in certain programmes the input system too has become over-complex. In addition, because of their general utility they are not always as fast as the special purpose-written programmes. The non-statistical or smaller programmes have nevertheless not lost their importance, and it is not inconceivable that, with desk computers becoming widespread, there will be a period of renaissance for these fast-running programmes with their low memory requirements.

In certain cases the computer solution of previously known and applied evaluation procedures had the advantage that it diminished some of the failings of the original method. It was pointed out by Rydberg [54] that, in the event of insufficiently accurate data, the Bjerrum method may result in more constants than the experimental data actually contain. In the Leden [56] and Fronaeus [30] methods as well the errors made in the extrapolations accumulate, and in an unfavourable case this may lead to an additional, non-existing "constant". The computer solution elaborated by Thun et al. [29] for the Fronaeus method eliminates this error. Naturally, all this does not mean that the earlier graphical procedures have been completely superseded by the computer solutions. Even in the age of the computer, they may be very useful, e.g. to provide preliminary information, while in the detection of systematic errors in measurements they are virtually indispensable.

Application of the statistical method necessitates the consideration of a number of fundamental questions. A basic assumption in the least-squares treatment, for instance, is the normal distribution of the errors. Disregarding the examinations by Rydberg [55] and the test used in the MINIQUAD programme [102] for the Gaussian distribution of residuals, in effect this question is still not unambiguously solved even today. Independently of this, in complex chemistry Gaussian distribution of the errors is taken as an accepted fact, but generally only because nothing seems to contradict this. In fact, the errors arise rather from the failure to take into account that if the error in a quantity x displays normal distribution, then the error distribution for its non-linear functions ($\log x$, x^2, etc.) is not normal.

Systematic errors in the experiments may mean a very serious danger in the course of the evaluations. In a simpler case, these only distort the values of the constants, but it may also occur that the systematic errors are compensated for by a constant ascribed to a complex that does not exist in reality; indeed, it cannot be excluded that the experimental results are described by a completely erroneous model. It must be accepted as a fact that it is very difficult to create experimental conditions under which systematic errors do not arise. These errors, however, can be attributed definitely to the experiments, and even without a computer it is only possible to draw erroneous conclusions from erroneous measurements. The most certain method to combat systematic errors is to carry out as many independent experiments as possible in the widest feasible concentration range. The advantage of computer evaluation is that a very large mass of such data can be processed without any particular difficulty.

The question of the selection of the function to be minimized and the calculation of the weighting factors was touched on earlier. Attention is drawn here to the aspect that, if there is a possibility for selection, minimization should be carried out with regard to the residuals of quantities for which the experimental points can be taken into consideration with equal weightings. From the literature data so far published, it appears that the experimental points can be treated with the same weight in refinements relating to values of $\bar{n}$, Z, η, E, $\log D$ and A.

A number of authors have raised the question of whether it is more reasonable to carry out the optimization with respect to the stepwise equilibrium constants, or to the stability products. Although Sillén [115] considered optimization of the K values to be better, most authors have performed calculations with the overall formation constants or their logarithms. Tobias and Yasuda [76] found calculation with the $\log \beta$ values to be more advantageous in the Gauss method. Gaizer too came to a similar conclusion [71]. A further advantage is that in this case the elements of matrix **G** (eqn. 9) are nearly of the same order of magnitude. The problems in solving eqn. (9), discussed by Gans [3] do not occur and in the calculation of spectrophotometric measurements, $\log \beta$'s and molar absorptivities, can be refined together [156]. On the other hand, Gans and Vacca [136] recommend calculation with the β values, since in this way they find establishment of the minimum in the error-square sum function to be more certain. Although their conclusion is justified, it should be taken into account only when the refinement must be performed setting out from very unreliable initial parameter values.

The selection of the model describing the system under study is generally based on chemical considerations. It is unconditionally advisable, however, to inspect other conceivable models too. It may happen that (in poorly-planned experiments, or in the case of measurements in a concentration interval which is not sufficiently wide because of experimental difficulties) the experimental curve can be described satisfactorily with a number of models. It is obvious that worthwhile conclusions cannot be drawn from such experiments.

On evaluation with statistical programmes, the assessment of the satisfactory or unsatisfactory nature of the fit requires great circumspection. For this the χ^2 test or the t test are used in a number of programmes. Vacca et al. [157] recommend the calculation of the R factor, similarly, as in crystallographic investigations. Although all these may be important in the assessment of the fit, there is no guarantee that the formally satisfactory, or even excellent fit would itself be conclusive. A statistically unsatisfactory fit indisputably casts doubt on the correctness of the selected model, but an excellent fit may also be a freak of chance. In every case, therefore, the results obtained must be subjected to a careful chemical checking, and agreement with other independent methods must be regarded as a substantiating circumstance.

A very great advantage of the application of statistical methods is that they provide a possibility for the exact calculation of the errors in the parameters. At any event, the principle is correct that the estimate of the error in a parameter is at least as important as the estimate of the value of the parameter [91]. In spite of this, there are relatively few publications in which the values of the individual parameters are given with their objectively calculated errors. Although there are many reasons for the origin of the surprisingly large number of unreliable equilibrium constants published in the literature [158,159], the unreliability of many constants would have been detected even in the course of the evaluation by the exact calculation of the errors in the constants. It appears that calculation of the errors in the parameters on the basis of the reproducibility alone would be misleading. Much more information is given by calculation of the "sharpness" as recommended by Conrow et al. [107] (but see also [80]). This is the percentage change in the sum of the squares of the residuals as a result of a 10% variation in the value of the parameter. In fact, really valuable information on the errors in the parameters is only given by methods based on statistical considerations (eqns. (10) and (14)).

In connection with the calculation of the errors, Gaizer [71,156] drew attention to the fact that it is advisable to take into account how many of the experimental points actually give information on the value of each parameter, and hence on its error. For measurements in a wide concentration interval, where a number of species are formed, this cannot be left out of consideration. For the calculation of the error in each parameter, therefore, the author takes into consideration only those experimental points for which the formation of the species relating to the parameter attains a reasonable threshold value. If the error in the parameter calculated in this way originates from too few experimental points (as compared to the total number), then, as a poorly determined constant, this parameter should not be accepted, or only with great caution.

A question which frequently arises in the comparison of the individual programmes is the running time of a programme. From a study of the information to date and the time requirements of the individual programme parts, it appears that the running time in statistical programmes is determined by the time for solution of the material balance equations; the time demands of the other operations are not appreciable. Primarily this must be taken into consideration, therefore, in a comparison of the running times of the various programmes. The Newton—Raphson [6,119] method is the fastest for the solution of the equations, but this may diverge in the case of uncertain starting values. Of the other procedures that always converge, the best is the iteration procedure with decreasing increment values, employed in the LETAGROP [116] and SPEFCA [71] programmes. The method used by Perrin et al. [99,160] requires an essentially longer running time than these.

As expected mathematically, comparative evaluations with the various least-squares programmes [97,101,136,161] resulted in essentially identical parameter values. The differences showed up rather in the errors in the parameters, but this probably originates partly from the method used in the differentiation, and partly (and mainly) from the fact that the values of the constants are calculated by the various programmes at different reliability levels. Surprisingly large differences were demonstrated by Hume [162] in the evaluation of given experimental data by various calculation methods. Similar results were obtained by Bond [163], although in this case the question arises as to whether 3 formation constants can be calculated by a statistical method from 15 equally-weighted experimental points differing by several orders of magnitude.

In conclusion, it must be pointed out that computer evaluation is one method for the calculation of equilibrium constants, and this method also has its advantageous and less advantageous features. Nowadays, the former seem to predominate. However, it should not be left out of consideration that it is very dangerous to accept the results obtained by means of the computer without due criticism; in all cases the results obtained must be supported by thorough chemical considerations, for only thus can a factual understanding of the system under investigation be reached.

REFERENCES

1 C.W. Childs, P.S. Hallman and D.D. Perrin, Talanta, 16 (1969) 1119.
2 F.J.C. Rossotti, H.S. Rossotti and R.J. Whewell, J. Inorg. Nucl. Chem., 33 (1971) 2051.
3 P. Gans, Coord. Chem. Rev., 19 (1976) 99.
4 M.T. Beck, Chemistry of Complex Equilibria, Akadémiai Kiadó, Budapest, 1970.
5 F.J.C. Rossotti and H.S. Rossotti, The Determination of Stability Constants, McGraw-Hill, New York, 1961.
6 J.B. Scarborough, Numerical Mathematical Analysis, The John Hopkins Press, Baltimore, 1930.
7 H. Cramér, Mathematical Methods of Statistics, Princeton Univ. Press, Princeton, 1946.
8 W.E. Deming, Statistical Adjustment of Data, Wiley, New York, 1948.
9 W.J. Youden, Statistical Methods for Chemists, Wiley, New York, 1951.

10 W.C. Hamilton, Statistics in Physical Science, Ronald Press Co., New York, 1964.
11 N.R. Draper and H. Smith, Applied Regression Analysis, Wiley, New York, 1966.
12 K. Eckschlager, Errors, Measurements and Results in Chemical Analysis, Van Nostrand Reinhold, London, 1969.
13 L.C.W. Dixon, Nonlinear Optimisation, English Universities Press, London, 1972.
14 B.J. Thamer, J. Phys. Chem., 59 (1955) 450.
15 B. Roth and J.F. Bunnett, J. Am. Chem. Soc., 87 (1965) 334.
16 J. Bjerrum, Metal Ammine Formation in Aqueous Solution, Haase, Copenhagen, 1941.
17 L.P. Varga, Anal. Chem., 41 (1969) 323.
18 D. Litchinsky, N. Purdie, M.B. Tomson and W.D. White, Anal. Chem., 41 (1969) 1726.
19 J.K. Romary, D.L. Donnelly and A.C. Andrews, J. Inorg. Nucl. Chem., 29 (1967) 1805.
20 H.S. Hendrickson, Anal. Biochem., 24 (1967) 176.
21 Th. Kaden and A. Zuberbühler, Talanta, 18 (1971) 61.
22 M.C. Chattopadhyaya and R.S. Singh, Anal. Chim. Acta, 70 (1974) 49.
23 J. Siefker, Anal. Chim. Acta, 52 (1970) 545.
24 D.L. McMasters and W.B. Schaap, Proc. Indiana Acad. Sci., 67 (1958) 111.
25 F.J.C. Rossotti and H.S. Rossotti, Acta Chem. Scand., 9 (1955) 1166.
26 G. Schwarzenbach, A. Willi and R.O. Bach, Helv. Chim. Acta, 30 (1947) 1303.
27 S.P. Datta and A.K. Grzybowski, J. Chem. Soc. (A), (1966) 1059.
28 A.K. Grzybowski, S.S. Tate and S.P. Datta, J. Chem. Soc. (A), (1970) 241.
29 H. Thun, F. Verbeek and W. Vanderleen, J. Inorg. Nucl. Chem., 29 (1967) 2109.
30 S. Fronaeus, Acta Chem. Scand., 4 (1950) 72.
31 K. Johansson, P. Ulmgren and O. Wahlberg, Chem. Comm Univ. Stockholm, Nr. X., (1971)
32 A.M. Corrie, G.K.R. Makar, M.L.D. Touche and D.R. Williams, J. Chem. Soc. Dalton Trans., (1975) 105.
33 J.E. Gorton and R.F. Jameson, J. Chem. Soc. Dalton Trans., (1972) 304.
34 M. Wozniak and G. Nowogrocki, Bull. Soc. Chim. Fr., (1974) 435.
35 L. Barcza, Magy. Kém. Foly., 75 (1969) 513.
36 I. Nagypál, A. Gergely and P. Jékel, J. Inorg. Nucl. Chem., 31 (1969) 3447.
37 P. De Weer and A.G. Lowe, J. Biol. Chem., 248 (1973) 2829.
38 A.P. Filippov, I.V. Khyarsing and K.B. Yatsimirskii, Ukr. Khim. Zh., 40 (1974) 787.
39 M.E. Lamm and D.M. Neville, J. Phys. Chem., 69 (1965) 3872.
40 C. Luca and G. Berthon, J. Pure Appl. Chem., 31 (1972) 503.
41 P. Sutre, C. R. Acad. Sci. Ser. C., 268 (1969) 2258.
42 R.W. Ramette, J. Chem. Educ., 44 (1967) 647.
43 J. Havel and V. Kuban, Scripta Fac. Sci. Natur. Univ. Purkyniynae Brun., 1 (1971) 87.
44 J. Havel, Trans. R. Inst. Technol. Stockholm, Nr. 277, (1972) 371.
45 V. Kuban and H. Havel, Acta Chem. Scand., 27 (1973) 528.
46 L.G. Sillén and B. Warnqvist, Ark. Kémi, 31 (1969) 317.
47 T.N. Briggs and J.E. Stuehr, Anal. Chem., 46 (1974) 1517.
48 T.N. Briggs and J.E. Stuehr, Anal. Chem., 47 (1975) 1916.
49 G.L. Cumming, J.S. Rollett, F.J.C. Rossotti and R.H. Whewell, J. Chem. Soc. Dalton Trans., (1972) 2652.
50 V.M. Tsentovskii and V.S. Tsentovskaya, Elektrokhimiya, 8 (1972) 1636.
51 J. Rydberg and J.C. Sullivan, Acta Chem. Scand., 13 (1959) 186.
52 J.C. Sullivan, J. Rydberg and W.F. Miller, Acta Chem. Scand., 13 (1959) 2023.
53 J. Rydberg and J.C. Sullivan, Acta Chem. Scand., 13 (1959) 2057.
54 J. Rydberg, Acta Chem. Scand., 14 (1960) 157.
55 J. Rydberg, Acta Chem. Scand., 15 (1961) 1723.
56 I. Leden, Z. Phys. Chem. A., 188 (1941) 160.
57 H. Olerup, Sv. Kem. Tidskr., 85 (1943) 324.
58 W.A.E. McBryde, Talanta, 21 (1974) 979.
59 G. Heys, H. Kinns and D.D. Perrin, Analyst, 97 (1972) 52.

60 R.M. Rush, J.S. Johnson and K.A. Kraus, Inorg. Chem., 1 (1962) 378.
61 B.N. Bad'yanov, Yu.F. Kolupaev and A.V. Morozova, Tr. Mosk. In-ta Elektron. Mashinostr., (50) (1975) 117; Chem. Abstr. 84 (1976) 141476f.
62 H. Irving and M.H. Stacey, J. Chem. Soc., (1961) 2019.
63 G. Schwarzenbach, Helv. Chim. Acta, 33 (1950) 947.
64 K. Momoki, H. Sato and H. Ogawa, Anal. Chem., 39 (1967) 1072.
65 D.D. DeFord and D.N. Hume, J. Am. Chem. Soc., 73 (1951) 5321.
66 R.R. Irani and C.F. Callis, J. Phys. Chem., 65 (1961) 934.
67 S.W. Rabideau and R.H. Moore, J. Phys. Chem., 65 (1961) 371.
68 S.N. Vinogradov, Can. J. Chem., 41 (1963) 2719.
69 C.F. Gauss, Theoriae Combinationis Observationum Errobus Minimis Obnoxiae, 1821, In Gauss' Werk, Bd. IV., Göttingen, 1873.
70 H. Kim, J. Chem. Educ., 47 (1970) 120.
71 F. Gaizer, Magy. Kém. Foly., 83 (1977) 250.
72 I.G. Sayce, Talanta, 15 (1968) 1397.
73 R.H. Moore and R.K. Zeigler, The Solution of the General Least-Squares Problem with Special Reference to High Speed Computers, LA-2367, March 4, 1960.
74 S.W. Rabideau and R.J. Kline, J. Phys. Chem., 64 (1960) 680.
75 S.T. Tobias and Z.Z. Hugus, Jr., J. Phys. Chem., 65 (1961) 2165.
76 S.T. Tobias and M. Yasuda, Inorg. Chem., 2 (1962) 1307.
77 R.M. Rush, J.S. Johnson and K.A. Kraus, ORNL-3278, Feb. 7, 1963.
78 L.G. Sillén, J. Pure Appl. Chem., 17 (1968) 55,
79 R.C. Lansbury, V.E. Price and A.G. Smeeth, J. Chem. Soc., (1965) 1896.
80 D.R. Rosseinsky and H. Kellawi, J. Chem. Soc. (A), (1969) 1207.
81 F. Gaizer, Magy. Kém. Foly., 78 (1972) 54.
82 F. Gaizer and I. Szalai, Magy. Kém. Foly., 78 (1972) 488.
83 K. Burger, F. Gaizer, B. Noszál, M. Pékli and G. Takácsi-Nagy, Magy. Kém. Foly., 83 (1977) 241; J. Bioinorg. Chem., 7 (1977) 335.
84 K. Burger, F. Gaizer, I. Zay, M. Pékli and B. Noszál, Magy. Kém. Foly., 83 (1977) 235; J. Inorg. Nucl. Chem., 40 (1978) 725.
85 K. Burger, F. Gaizer, E. Papp-Molnár and Tran Thi Binh, J. Inorg. Nucl. Chem., 36 (1974) 863.
86 F. Gaizer, P. Buxbaum, E. Papp-Molnár and K. Burger, J. Inorg. Nucl. Chem., 36 (1974) 849.
87 F. Gaizer, Tran Thi Binh and K. Burger, J. Inorg. Nucl. Chem., 36 (1974) 1601.
88 F. Karlson and R. Vestin, Chem. Scr., 2 (1972) 207.
89 P.J. Lingane and Z.Z. Hugus, Jr., Inorg. Chem., 9 (1970) 757.
90 K.R. Magnell, J. Chem. Educ., 50 (1973) 619.
91 W.E. Wentworth, J. Chem. Educ., 42 (1965) 96.
92 G.E.P. Box, Bull. Inst. Intern. Statistique, 36 (1958) 215.
93 W.E. Wentworth, W. Hirsch and E. Chem, J. Phys. Chem., 71 (1971) 218.
94 D.D. Perrin and I.G. Sayce, J. Chem. Soc. (A), (1967) 82.
95 D.D. Perrin, I.G. Sayce and V.S. Sharma, J. Chem. Soc. (A), (1967) 1755.
96 D.D. Perrin and I.G. Sayce, J. Chem. Soc. (A), (1968) 53.
97 I.G. Sayce, Talanta, 18 (1971) 653.
98 W.A.E. McBryde and J.L. McCourt, Talanta, 19 (1972) 1486.
99 I. Nagypál, Magy. Kém. Foly., 80 (1974) 49.
100 D.J. Legett and W.A.E. McBryde, Anal. Chem., 47 (1975) 1065.
101 A. Sabatini and A. Vacca, J. Chem. Soc. Dalton Trans., (1972) 1693.
102 A. Sabatini, A. Vacca and P. Gans, Talanta, 21 (1974) 53.
103 P. Gans, A. Sabatini and A. Vacca, Inorg. Chim. Acta, 18 (1976) 237.
104 D.W. Marquart, J. Soc. Ind. Appl. Math., 11 (1963) 431.
105 E.A. Unvin, R.G. Beimer and Q. Fernando, Anal. Chim. Acta, 39 (1967) 95.
106 T. Meites and L. Meites, Talanta, 19 (1972) 1131.

107 K. Conrow, G.D. Johnson and R.E. Bowen, J. Am. Chem. Soc., 86 (1964) 1025.
108 K. Nagano and D.E. Metzler, J. Am. Chem. Soc., 89 (1967) 2891.
109 W. Likussar, Anal. Chem., 45 (1973) 1926.
110 W. Likussar and D.F. Boltz, Anal. Chem., 43 (1971) 1265.
111 M.D.J. Powell, Computer J., 7 (1964) 155.
112 R.N.F. Thorneley, A.G. Sykes and P. Gans, J. Chem. Soc. (A), (1971) 1494.
113 B. Sarkar and T.P.A. Kruck, Can. J. Chem., 51 (1973) 3541.
114 D. Dyrssen, N. Ingri and L.G. Sillén, Acta Chem. Scand., 15 (1961) 694.
115 L.G. Sillén, Acta Chem. Scand., 16 (1962) 159.
116 N. Ingri and L.G. Sillén, Acta Chem. Scand., 16 (1962) 173.
117 L.G. Sillén, Acta Chem. Scand., 18 (1964) 1085. ,
118 N. Ingri and L.G. Sillén, Ark. Kémi, 23 (1964) 97.
119 R. Arnek, L.G. Sillén and O. Wahlberg, Ark. Kémi, 31 (1969) 353.
120 L.G. Sillén and B. Warnqvist, Ark. Kémi, 31 (1969) 315.
121 P. Brauner, L.G. Sillén and R. Whiteker, Ark. Kémi, 31 (1969) 365.
122 L.G. Sillén and B. Warnqvist, Ark. Kémi, 31 (1969) 341.
123 P. Ulmgren and O. Wahlberg, Chem. Scr., 8 (1975) 126.
124 M. Meloun and J. Čermak, Talanta, 23 (1976) 15.
125 M. Meloun and J. Čermak, Sc. Papers of Univ. Chem. Technol. Pardubice, 30 (1973) 49.
126 R. Arnek, Ark. Kémi, 32 (1970) 81.
127 D.H. Liem, Acta Chem. Scand., 25 (1971) 1521.
128 D. Inman, I. Regan and B. Girling, J. Chem. Soc. (A), (1964) 348.
129 P. Paoletti, A. Vacca and D. Arenare, J. Phys. Chem., 70 (1966) 193.
130 A. Vacca, D. Arenare and P. Paoletti, Inorg. Chem., 5 (1966) 1384.
131 J. Prasad and N.C. Peterson, Inorg. Chem., 8 (1969) 1622.
132 R.O. Gould and H.M. Sutton, J. Chem. Soc. (A), (1970) 1184 and 1439.
133 P. Zur Nedden, E. Merciny and G. Duyckaerts, Anal. Chim. Acta, 64 (1973) 197.
134 A. Massoumi, P. Overvoll and F.J. Langmyhr, Anal. Chim. Acta, 68 (1974) 103.
135 A. Márton and J. Inczédi, Magy. Kém. Foly., 80 (1974) 145.
136 P. Gans and A. Vacca, Talanta, 21 (1974) 45.
137 B. Warnqvist, Chem. Scr., 1 (1971) 49.
138 L. Lyhamn, Chem. Scr., 10 (1976) 49.
139 L.P. Varga and D.N. Hume, Inorg. Chem., 2 (1963) 201.
140 L.P. Varga, W.D. Wakley, L.S. Nicolson, M.L. Madden and J. Patterson, Anal. Chem., 37 (1965) 1003.
141 Z.Z. Hugus, Jr., in S. Kirschner (Ed.), Advances in the Chemistry of the Coordination Compounds, Macmillan, New York, 1961.
142 Z.Z. Hugus, Jr., Proceedings of the International Conference on Coordination Chemistry, Stockholm, 1962.
143 W.C. Davidon, U.S. At. Energy Repts., ANL-5990, 1959.
144 R. Fletcher and M.J.D. Powell, Computer J., 6 (1963) 163.
145 P. Gans, J. Chem. Soc. (A), (1971) 2017.
146 G.J. Ford, P. Gans and L.D. Pettit, J. Chem. Soc. Dalton Trans., (1972) 1763.
147 S. Feldberg, P. Klotz and L. Newman, Inorg. Chem., 11 (1972) 2860.
148 A. Kaneda and A.E. Martell, J. Coord. Chem., 4 (1975) 137.
149 T.R. Griffiths and Ph.J. Potts, J. Chem. Soc. Dalton Trans., (1975) 344.
150 J.J. Kankare, Anal. Chem., 42 (1970) 1322.
151 J.P. Chandler, Minimum of a Function of Several Variables, Program 66.1, Quantum Chemistry Program Exchange, Indiana University, Bloomington, Ind., 1966.
152 K. Vadasdi, J. Phys. Chem., 78 (1974) 816.
153 R.M. Wallace and S.M. Katz, J. Phys. Chem., 68 (1964) 3890.
154 L.P. Varga and F.C. Veatch, Anal. Chem., 39 (1967) 1101.
155 C.D. Chriswell and A.A. Schilt, Anal. Chem., 47 (1975) 1623.

156 F. Gaizer and M. Máté, Acta Chim. Hung. Acad. Sci., in press.
157 A. Vacca, A. Sabatini and M.A. Gristina, Coord. Chem. Rev., 8 (1972) 44.
158 J.E. Prue, Chem. Br., 3 (1967) 344.
159 M.T. Beck, J. Pure Appl. Chem., 49 (1977) 127.
160 D.D. Perrin and I.G. Sayce, Talanta, 14 (1967) 833.
161 A. Dei, P. Paoletti and A. Vacca, Inorg. Chem., 7 (1968) 865.
162 D.N. Hume, Proc. 3rd Symp. Coord. Chem., Debrecen, Hungary, 1970., Akadémiai Kiadó, Budapest, 1971.
163 A.M. Bond, Coord. Chem. Rev., 6 (1971) 377.