

The Use of High Speed Digital Computers for the Least Squares Calculation of Complexity Constants*

JAMES C. SULLIVAN, JAN RYDBERG** and
WILLIAM F. MILLER

*Chemistry Division and Applied Mathematics Division, Argonne National Laboratory,
Lemont, Illinois, U.S.A.*

Some general aspects are given of the calculation of stability constants for mononuclear complexes by means of a weighted least squares technique. Computer (the IBM 704, and the GEORGE) programs are developed for the treatment of solvent extraction and potentiometric titration data. The results of the use of these programs are compared in a number of examples with results of earlier (usually graphical) calculations.

Despite the extensive amount of research that has been conducted during the past twenty years the problems encountered in the study of the association between cations and complexing ligands, in aqueous systems, are far from completely resolved. Most of the data in the literature has been interpreted on the basis of the model, originally suggested by N. Bjerrum¹, that the formation of complexes always appears to occur in step-wise fashion with the stabilities of the various species characterized by a series of mass action constants.

If the free energy data of complex ion systems are to be correlated with other properties of the constituents of such systems (as for example the radii of the cations) it is imperative that the original data be summarized in a manner efficient for comparative purposes.

Approximation techniques have been generally employed for the calculation of association constants^{2,3}. The increasing availability of high speed digital computers makes it feasible to use a standard objective treatment, *e.g.*, the least squares adjustment, in the calculation of such constants. Computations involved in such an adjustment become difficult when a desk type of calculator is used both because of the large amount of data generally accumulated in the study of a complex ion system and because of the fact that the normal equations are often illconditioned.

* Based on work performed under the auspices of the U. S. Atomic Energy Commission.

** On leave of absence from the University of Stockholm, and the Research Institute of National Defense, Stockholm, Sweden.

Along with the calculation of the best values for the association constants, the digital computer calculation makes it feasible to calculate the standard deviation to be assigned to each constant (based on the over-all fit of the calculated curve to the experimental points) as well as to provide some measure of the consistency of the model used (based on the same criteria) without recourse to *a priori* assumptions concerning the specific number of entities involved.

In the present communication we will present an outline of the methods that have been used to express the relevant equations in forms suitable for the digital computer computations and present the results obtained for a few selected systems.

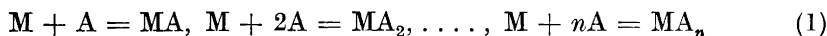
1. GENERAL ASPECTS OF THE CALCULATION OF SUCCESSIVE CONSTANTS OF MONONUCLEAR COMPLEXES BY A WEIGHTED LEAST SQUARES TECHNIQUE

Basic equations and definitions

$[M]$ = concentration of uncomplexed central ion or molecule

$[A]$ = concentration of unbound ligand

$[MA_n]$ = concentration of the n th complex



The equilibrium constant for the n th complex is

$$\beta_n = \frac{[MA_n]}{[M][A]^n} = \prod_1^n k_n \quad (2)$$

where the k 's are constants for the formation of the individual complexes formed in a stepwise manner. The total central ion concentration C_M is

$$C_M = [M] + \sum_{n=1}^N [MA_n] \quad (3)$$

and the total ligand concentration C_A is

$$C_A = [A] + \sum_{n=1}^N n [MA_n] \quad (4)$$

where N is the maximum number of ligands bound per metal atom. The ligand number, \bar{n} , is defined² as the average number of ligands combined with each central atom

$$\bar{n} = \frac{C_A - [A]}{C_M} = \frac{\sum_{n=1}^N n \beta_n [A]^n}{\sum_{n=0}^N \beta_n [A]^n} \quad (5)$$

The degree of formation ² of a given complex, MA_c, is

$$\alpha_c = \frac{[\text{MA}_c]}{C_M} = \frac{\beta_c [\text{A}]^c}{\sum_{n=0}^N \beta_n [\text{A}]^n} \quad (6)$$

Various combinations of the above basic equations have been used for the estimation of association constants³. Underlying eqn. 2 is the basic assumption that the ratio of the activity coefficients of the various species involved in the equilibrium are constant. A test of the validity of the model specified by eqn. 2 is not included in this investigation.

Some general aspects of the calculations

When, as in our examples below, either a potentiometric titration or a solvent extraction procedure is used to investigate the formation of complex ions, the particular equations that are relevant for the specific procedure can be expressed in terms of a power series. The parameters of this power series (which are to be determined by a least squares procedure) are constants of the type defined in eqn. 2 or related to that type of constant by an elementary transformation. On the basis of the model we are using, negative parameters and those whose standard deviation spans zero will be rejected, that is, we shall determine the best set of parameters in the least squares sense subject to the restriction that each parameter be positive, in fact, that each parameter minus one standard deviation be positive.

An outline of the least squares analysis used as a basis for the computer programming is given in Appendix I.

The equations that we are to subject to the least squares adjustment generally merit the term of opprobrium, "ill-conditioned"⁴. One can understand how a condition sufficient to bring about this state of the equations arises from a consideration of the physical aspects of the problem.

An experiment to determine association constants involves the variation of the concentration of the complexing ligand under consideration over as wide a range possible concomitant with the conditions of observational feasibility and invariance of activity coefficient ratios. The usual result is the variation of such a concentration (or independent variable) over many orders of magnitude with a much smaller range covered by the measured property (or dependent variable). Such a disparity in the range of the dependent and independent variables is sufficient to bring about ill-conditioned systems.

From the computational viewpoint this ill-conditioning means that there are a large number of nearly correct solutions to the equations. The existence of this wide variety of nearly correct solutions leads to difficulties in obtaining numerical values for the parameters. Although it is feasible to compute a measure of the degree of conditioning of the equations to be solved^{4,5} since we are primarily concerned with the parameters obtained, and their standard deviation, we have not performed any such calculations in the course of this investigation.

The capacity of the digital computers is generally large enough so that the effect of rounding off errors on the matrix inversion is minimal for the computations to be described. This tends to eliminate one of the sources of difficulty encountered when the matrix inversion technique is used for the solution of the normal equations.

Prior to any computer calculation, the data is (of course) studied graphically to determine: (a) the highest power of the independent variable in the series that should be considered (*i.e.* N); (b) that a smooth curve can indeed be drawn through the experimental points.

After the calculation has been completed the computed curve is plotted and the scatter of the experimental points around this curve is examined to determine if there are systematic deviations, *i.e.*, a large number of consecutive points which lie on one side of the curve. In addition, since a log-log scale (with its attendant distortion) is generally most convenient for a graphical representation of the type of data considered here, a point by point comparison between experimental and calculated values is carried out to insure the adequacy of the functional form used in the computation.

Weighting procedures

In the analysis outlined in the Appendix it has been explicitly assumed that only the measured distribution coefficients (in the case of solvent extraction experiments) and the measured free ligand concentration (for the potentiometric titration investigations) have associated random errors. The validity of this assumption is securely based on the experimental procedures used in these investigations. It is also apparent from considerations of the experimental procedures that the individual observations have different variances over the range of free ligand concentrations studied. This necessitates weighting the individual measurements.

With the foregoing in mind an estimate of the weight of an individual observation in experiments of the type to be considered here is most reasonably obtained by detailed considerations of the specific investigation. For example, in the case of the $\text{Np(IV)}-\text{HSO}_4^-$ system⁶, since the measured distribution was always adjusted to approximately unity and replication indicated the same precision at all points, the use of $1/y$ as a weighting parameter is reasonable.

For other distribution studies the initial approximation of the weights may be obtained from the measured standard deviation at free ligand concentrations for which replicate measurements were made or by considering the error to be feared in the distribution coefficient caused by the error in the analysis of each phase. For studies in which a radiometric assay procedure is used this latter quantity is readily estimated using standard formulas⁷. Caution must be exercised in assigning weights based on the variance from replicate measurements unless the number of measurements at each point is large. When only a few sets of duplicate measurements are available consideration should be given to using an average percentage uncertainty as first approximation to the individual weights.

For potentiometric titration experiments similar considerations may be used as the basis for initial weight approximations. For example in the absence of replicate experiments the type of precision to be expected in an emf measurement might be used as the basis of assigning weights to an individual observation. To avoid prolixity, further discussion of this point will be deferred until the individual examples are considered.

In the calculations to be described use has been made of the errors of external consistency⁸ for the computation of the standard derivations of the parameters and as a test of the validity of the hypothesis under consideration. This approach is a reasonable one to use if we consider the type of data that is to be treated. For example, it is easy to visualize a solvent extraction experiment with systematic errors in a certain region of free ligand concentration. Errors of this type would never be observed by replication and consequently weights based on the apparent precision of measurements at a point in this region would not be realistic for the calculation of the weights of the parameters. Therefore the square of the actual deviation between experimental and calculated points is a logical scale to use for the calculation of the standard deviations of the parameters.

In addition, the reasonableness of an arbitrary assignment of weights to individual observations can be justified by the fact that if one is fitting with the correct functional form and has assigned the correct weights, then the mean value of $\chi^2/k - 1 = 1$ *.

Computer programs

Three separate programs were used in the course of this study, two of them being specifically developed for this investigation.

The digital computer "George", developed at Argonne National Laboratory⁹, has available a program for the least squares calculation of the a_n in the polynomial

$$y = \sum_0^N a_n x^n \quad (7)$$

The input data for this program are y , x and the weight of y . This program is suitable for the computation of the β_n in the polynomials used by either Leden¹⁰ or Olerup¹¹ for graphical estimations of the parameters.

$$\frac{C_M - [M]}{[M][A]} = \sum_{n=1}^N \beta_n [A]^{n-1} \quad (\text{Leden}) \quad (7a)$$

$$\frac{C_M}{[M]} = \sum_{n=0}^N \beta_n [A]^n \quad (\text{Olerup}) \quad (7b)$$

As illustrated by the treatment of the $\text{Np(IV)} - \text{HSO}_4^-$ data, the form of the input is altered depending on the type of measurements made; cf. Ref.¹².

* See Appendix for a definition of this quantity.

A program was developed for the IBM 704 for use in solvent extraction investigations. This procedure calculates the best (in the least squares sense) a_n in the polynomial

$$yx^N = z = \sum_0^N a_n x^n \quad (8)$$

The input data are x , y , and estimated errors in the determination of the concentrations in both phases from which weights are calculated¹². This program can, *e.g.*, be used for data that can be summarized by eqn. 6. The development of the program in the specific form used in the present investigation emphasizes the underlying philosophy of machine computations, *i.e.*, the minimization of transformations applied to the original experimental data prior to computer input.

The \bar{n} relation of Bjerrum² (eqn. 5) can be rearranged

$$\sum_{n=0}^N (\bar{n}-n) \beta_n [A]^n = \sum_{n=0}^N (C_A - [A] - n C_M) \beta_n [A]^n = 0 \quad (9)$$

A program was developed for the IBM 704 to calculate the best values for the β_n from the equation:

$$0 = \sum_0^N (y - x - nz) \beta_n x^n \quad (9a)$$

The input data are $y = C_A$, $x = [A]$ and $z = C_M$ as well as the estimated standard deviation that occurs in the measurement of x .

Output data that are common to all three programs includes: the values of the parameters with their standard deviation (based on external consistency); the value of $S_{\min.}$ *, calculated values of the function, using the derived constants, for the original experimental x values.

2. EXAMPLES OF THE TREATMENT OF SOLVENT EXTRACTION DATA

The System Np(IV)-HSO₄⁻⁶. The use of theonyltrifluoroacetone (TTA) as a chelating agent for the extraction of cations from an aqueous into an organic phase has been used by a number of investigators¹³ to obtain data that may be interpreted in terms of ionic association. The assumptions upon which this interpretation is based have been clearly stated^{13a} and will be used in the following treatment. The partition ratio of the concentration of Np(IV) in the organic phase to the total concentration of Np(IV) in the aqueous phase, Q , under conditions of constant acidity and ionic strength and corrected to unit activity of TTA in the organic phase can be expressed as,

$$Q = Q_0 / \sum_0^N \beta_n [\text{HSO}_4^-]^n \quad (10)$$

where $\beta_0 = 1$, and $N = 2$.

* See Appendix.

Simple transformations lead to

$$y = a_0 + a_1 [\text{HSO}_4^-] + a_2 [\text{HSO}_4^-]^2 \quad (11)$$

where $y = 1/Q$, $a_0 = 1/Q_0$ (Q_0 is the normalized distribution coefficient of Np(IV) in the absence of HSO_4^-), $a_1 = \beta_1/Q_0$, $a_2 = \beta_2/Q_0$. This puts the original data in a form suitable for the least squares calculation of the a_n using the computer program based on eqn. 7.

The weight of the individual observation was taken as $1/y$ as previously indicated. In this case a test of the internal consistency of the data is made possible by comparison of the calculated value of Q_0 with the independent experimental determination of this quantity. The relevant output data are

Table 1. Parameters describing the system Np(IV)- HSO_4^- , 25°C, $\mu = 2.00$, $[\text{H}^+] = 2.00$.

	Q_0	k_1	k_2
This work	$(2.94 \pm .23) \times 10^4$	118 ± 14	16 ± 2
Original ⁶	$(2.84 \pm .04) \times 10^4$	135 ± 14	16 ± 2

summarized in Table 1, along with the results of the original computations made by the limiting value method of Leden¹⁰. The uncertainties listed in the above are standard deviations as are all such uncertainties stated in this work.

The values for the parameters obtained by the least squares adjustment are in reasonable agreement with the original graphical estimates. It is of interest to note that the original estimate of the errors (%) in the parameters k_1 and k_2 are of the same order of magnitude as those determined by the more objective computation described here. However the estimation of the uncertainty in Q_0 , based on the experiments summarized in Table 2 of the original investigation, is seen to be significantly smaller than the standard error based on eqn. 8, Appendix. This emphasizes the logic of the choice of external consistency as a basis for weighting the parameters determined by the least squares adjustment.

Pu(IV)-acetylacetonone. A large number of solvent extraction investigations have been conducted for the purpose of determining association constants between metal ions and organic ligands¹⁴. The assumptions basic to such an interpretation have been outlined¹⁵ and are implicitly accepted in the following treatment.

We will consider the least squares adjustment of the data presented for the Pu(IV)-acetylacetonone system³¹ as an example of the utility of digital computers in the interpretation of such distribution results. Considering only uncharged species in the organic phase, and under conditions of constant ionic strength from considerations similar to those previously discussed, we may write eqn. 8 as

$$y = [\text{Aa}]^4/Q = a_0 + a_1 [\text{Aa}] + a_2 [\text{Aa}]^2 + a_3 [\text{Aa}]^3 + a_4 [\text{Aa}]^4 \quad (12)$$

where Q is the measured partition ratio of the concentration of PuAa_4 in the organic phase to all species of Pu in the aqueous phase, $[\text{Aa}]$ is the free

Table 2. Parameters for the system Pu(IV)-acetylacetonone, $\mu = 0.1$, 25°C.

Computational	$-\log a_0$	$-\log a_1$	$-\log a_2$	$-\log a_3$	$-\log a_4$
Least squares		26.22 ± 0.11		8.45 ± 0.06	2.54 ± 0.07
Two parameter *	36.4 ± 1.3	25.5 ± 0.8	16.2 ± 0.7	8.5 ± 0.4	2.44 ± 0.20
Ligand number *	36.0 ± 1.3	25.0 ± 0.8	16.9 ± 0.7	8.4 ± 0.4	2.50 ± 0.20
Limiting value *	36.6 ± 1.0	25.8 ± 0.8	16.2 ± 0.7	8.6 ± 0.4	2.56 ± 0.20

* A description of the approximation techniques as used is given in Ref. ⁸ⁱ

acetylacetonone ligand concentration in the aqueous phase, $a_4 = 1/\lambda_4$ where λ_4 is the distribution coefficient of PuAa_4 , $a_3 = \beta_3/\lambda_4\beta_4$, $a_2 = \beta_2/\lambda_4\beta_4$, $a_1 = \beta_1/\lambda_4\beta_4$, and $a_0 = 1/\lambda_4\beta_4$. From the published data one may estimate a weight for the original observations based on a 20 % standard error in the measured partition ratio Q . This type of estimate is obtained from a consideration of the results of replicate experiments since the analytical scheme does not readily afford a basis for choosing weights based on the uncertainties to be expected in the determination of Q .

The relevant output data is summarized in Table 2 along with the graphical estimates of the parameters computed in the original publication.

The most interesting aspect of the results obtained is that the least squares treatment of the data allows a complete description of the results using only three parameters. Attempts to force a fit to the data in terms of five parameters, *i.e.*, in terms of the *a priori* assumption that the formation of the complex ions PuAa^{+3} , PuAa_2^{+2} , PuAa_3^+ , PuAa_4 , proceeds in a regular manner, always resulted in values for the parameters a_0 and a_2 which were inconsistent with the specified model. This example emphasizes one of the most useful features of the computational method under consideration that it provides a description of the results in a manner consistent only with the precision of the observations.

Chemical bias coupled with a highly ill-conditioned system made possible the results obtained originally by the limiting value and ligand number methods. The results obtained with the two parameter method are a reflection of the assumptions used in that approach and emphasize the limitations imposed by Dyrssen and Sillen in their development of this technique ^{3b}, *i.e.*, that the equations developed are exact only for the cases where one or two species are formed.

From the chemical viewpoint this result implies that, at equilibrium, in the experiments considered, the aqueous phase concentration of the species PuAa_2^{+2} was not large enough to have any observable influence on the measured distribution coefficient. The data are adequate to provide reliable values of the distribution coefficient, the product k_2k_3 and k_4 . Any further information adduced from these results must be viewed with the reservation accorded postulates of unknown scientific validity.

3. POTENTIOMETRIC TITRATION DATA

Eqn. 9 is the basis for computation of complexity constants in systems where the total metal, ligand and free ligand concentrations are the measured quantities. Data from the pioneering work of J. Bjerrum² will be used to demonstrate applications of this computer program.

The Systems Ag(I)–NH₃ and Cu(II)–NH₃. The results obtained for the first mentioned system with an estimated weight of an individual observation corresponding to a standard deviation of 0.7 % in the free ligand concentra-

Table 3. Association constants for the system Ag(I)-NH₃, $\mu = 2.0$ (NH₄NO₃), $t = 30^\circ\text{C}$.

	log k_1	log k_2
Bjerrum ²	3.20	3.83
This work	3.181 \pm 0.008	3.848 \pm 0.008
Irving and Rossotti ^{3c}	3.19	3.84

tion (yielding $\chi^2/k - 1 = 1.00$) are presented in Table 3. For comparison we also include the values obtained from the computations in the original paper as well as a set of values for the association constants computed by Irving and Rossotti^{3c}. The later investigators made their computations by rearranging what is essentially eqn. 9 into a form linear in β_1 and β_2 and performing an unweighted least squares computation. The results obtained in this example emphasize a number of important aspects of the present investigation.

(1) When experiments of high precision (as in this case) are available a weighted least squares treatment provides the best summary of such data. This is demonstrated by the fact that the precision of the original data justifies the use of one more significant figure than was originally presented. Another question is if the third decimal figure is of any significance because of systematic errors; however, this will not be discussed here.

(2) When graphical adjustments are properly carried out they will provide values for the constants consistent with the results obtained by the analytical methods being discussed here.

(3) The values given by the third treatment tabularized points out a pitfall for the unwary. The maximum amount of informations consistent with the precision of the experiments is not obtained unless the least squares treatment of the data is carried out in a proper manner.

It is instructive to consider a second example from the same work² in order to point out difficulties that may arise when using the present computer program. It is obvious that if any two rows or columns of the matrix **W** (eqn. 5, Appendix) are nearly multiples of one another a situation will exist that makes for near indeterminacy with the consequent difficulties in a computation of the parameters. Another more subtle point that will be troublesome can be shown by the following example.

For the system Cu–NH₃ calculations based on all the observations failed to yield significant parameters. Inspection of the data will show that at the point with $\bar{n} = 4.002$ there is a change in the value of χ by approximately a factor of 10. The introduction of this number into the matrix **W** may req-

Table 4. Association constants for the system Cu(II)-NH₃, $\mu = 2.0$ (NH₄NO₃), $t = 30^\circ\text{C}$.

	log k_1	log k_2	log k_3	log k_4
Bjerrum ²	4.19	3.50	2.88	2.19
This work	4.134 \pm 0.004	3.528 \pm 0.008	2.87 \pm 0.01	2.15 \pm 0.02

quire a range of numerical values exceeding the capacity of the computer. The use of other techniques, such as orthogonal polynomials is under investigation in an attempt to extend the computations over the complete range of the variables. A computation excluding this point gives the results summarized in Table 4. A weight proportional to a standard deviation in the free ligand concentration of 0.72 % will yield a value of $\chi^2/k-1 = 1.00$ for this system.

The conclusions to be drawn from these results are self-evident in the light of our previous discussion.

CONCLUSIONS

If quantitative comparisons between the association constants of complex ion systems are to be made in terms of electrostatic effects or any other of the fundamental properties of these systems such as suggested by J. Bjerrum ² it is necessary not only that precise data be available but also that such data is summarized in the most efficient possible manner. Such a summary is provided by the least squares adjustment of experimental observations. It should be emphasized that the least squares adjustment of data for the calculation of complexity constants is readily feasible at the present because of the development of high speed computers. The concepts involved have been well known for over a hundred years.

Attention should be directed to the limitations on the information that is provided by such an adjustment. This particular method of treating the data can only compensate for the ubiquitous random errors which plague the experimentalist. The use of this method cannot compensate for inadequate data.

However, when properly applied, the least squares treatment does provide the best summary of experimental observations. We hope that the availability of the programs for the IBM 704 will be of assistance to investigators who do desire to use this treatment of their data.

APPENDIX

Presented below, in outline form, is the analysis upon which the computer programs we have described are based. Consider the equation

$$y = \sum_{n=0}^N a_n x^n \quad (1)$$

where the dependent variable is subject to random errors. The weighted least squares method requires that a set of a_n be determined which minimizes

$$S = \sum_{i=1}^I w_i \left(\sum_0^N a_n x_i^n - y_i \right)^2 \quad (2)$$

The choice of

$$w_i = \frac{1}{\sigma_i^2} \quad (3)$$

where σ_i^2 is the variance of the measurement y_i , corresponds to the Chi-squared minimum method*.

The normal equations resulting from the minimization of S with respect to variations in the a_n are:

$$\begin{bmatrix} \sum_i w_i & \sum_i w_i x_i & \dots & \sum_i w_i x_i^N \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \sum_i w_i x_i^N & \sum_i w_i x_i^{N+1} & \dots & \sum_i w_i x_i^{2N} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \cdot \\ \cdot \\ a_N \end{bmatrix} = \begin{bmatrix} \sum_i w_i y_i \\ \sum_i w_i y_i x_i \\ \cdot \\ \cdot \\ \sum_i w_i y_i x_i^N \end{bmatrix} \quad (4)$$

or in matrix notation $\mathbf{WA} = \mathbf{Y}$. (5)

Eqn. 5 is solved by inverting \mathbf{W} and multiplying this inverse matrix into the vector \mathbf{Y} . Thus

$$\mathbf{A} = \mathbf{W}^{-1} \mathbf{Y} \quad (6)$$

yields the best values of the parameters a_0, a_1, \dots, a_N . The minimum value of S , that is the right hand side of eqn. 2 using the best values of the $\{a_n\}$, is then calculated, and the standard deviations of the parameters are calculated by external consistency *via*

$$\sigma_{a_i}^2 = \mathbf{W}_{ii}^{-1} \frac{S_{\min}}{k-1}, \quad (7)$$

where \mathbf{W}_{ii}^{-1} is the i th diagonal element of the inverse of \mathbf{W} , and k is the number of degrees of freedom for the system (4)**. In this formulation the number of degrees for the system is $[I - (N + 1)]$ ***. Furthermore the sample values of S_{\min} for a system of k degrees of freedom have the Chi-squared distribution for $k-1$ degrees of freedom****. Thus we should expect S_{\min} to have the χ^2 distribution for $(I - N)$ degrees of freedom and eqn. 7 becomes

$$\sigma_{a_i}^2 = \mathbf{W}_{ii}^{-1} \frac{S_{\min}}{I - N} \quad (8)$$

The consistency of the resulting parameters $\{a_n\}$, the measurements $\{y_i\}$, and the estimated errors of measurement $\{\sigma_i\}$ is tested by comparing the value of S_{\min} with the expected value of χ^2 for $(I - N)$ degrees of freedom.

* Harald Cramér, *Mathematical Methods of Statistics*, Princeton University Press, Princeton 1946, § 18.1 and § 30.3.

** Dumond and Cohen, *Rev. Modern Phys.* **25** (1953) 691.

*** Harald Cramer, *Mathematical Methods of Statistics*, § 29.2, p. 379 ff.

**** *Ibid.*, § 29.3, p. 381 ff.

With the following modifications, eqn. 9, as applied to the titration data, can be given the same treatment.

Let

$$U = \sum_{n=0}^N (y - x - nx) \beta_n x^n \quad (9)$$

In the titration experiment, one should have $U = 0$. One can determine that set of $\{\beta_n\}$ which, for a given set of $\{x_i, y_i, z_i\}$, makes U nearest to zero in the least squares sense by minimizing

$$S = \sum_{i=1}^I w_i U^2(x_i, y_i, z_i) \quad (10)$$

with respect to variation of the parameters $\{\beta_n\}$. The weights, w_i , should be taken as

$$w_i = \frac{1}{\delta U^2} \quad (11)$$

where δU is the variation in U for fluctuations in the variables (x, y, z) . If most of the variation in U is due to variation in x , then

$$\delta U = \frac{\partial U}{\partial x} \sigma_x \quad (12)$$

Since one needs to know the β_n in order to calculate $\partial U/\partial x$, an iterative procedure was employed here. First a set of approximate β 's was calculated for $w_i = 1$. Using these approximate β 's, the weights *via* eqnns. 12 and 11 were calculated. Then an improved set of β 's are calculated. Generally three such iterations yield a consistent solution.

Acknowledgement. The authors wish to express their appreciation to Dr. J. C. Hindman whose advice and criticism have enhanced the development of this investigation. In addition we wish to thank the personnel of the Applied Mathematics Division, in particular the programmers, Mrs. M. Hanson, Mrs. S. Steidl and Mr. W. Greenhow, for their technical assistance.

REFERENCES

1. Bjerrum, N. *Kgl. Kgl. Danske Videnskab Selskab. Skrifter, naturvidenskab. math. Afdel.* (1) **12** (1925).
2. Bjerrum, J. *Metal Ammine Formation in Aqueous Solution*, P. Hasse and Son, Copenhagen 1941.
3. For a review of the standard techniques: (a) Sullivan, J. C. and Hindman, J. C. *J. Am. Chem. Soc.* **74** (1952) 6091. Other methods are discussed by: (b) Dyrssen, D. and Sillén, L. G. *Acta Chem. Scand.* **7** (1953) 663. (c) Irving, H. and Rossotti, H. S. *J. Chem. Soc.* **1953** 3397. (d) Block, B. P. and McIntyre, G. H. Jr. *J. Am. Chem. Soc.* **75** (1953) 5667. (e) Rossotti, F. J. C. and Rossotti, H. S. *Acta Chem. Scand.* **9** (1955) 1166.

- (f) Irving, H., Rossotti, F. J. C. and Williams, R. J. P. *J. Chem. Soc.* **1955** 1906.
- (g) Gryder, J. W. *J. Am. Chem. Soc.* **77** (1955) 6196.
- (h) Hearon, J. Z. and Gilbert, J. B. *Ibid* **77** (1955) 2594.
- (i) Rydberg, J. *Arkiv Kemi* **9** (1956) 109.
- (j) Colleter, J.-C. and Romain, P. *Bull. soc. chim. France* **1958** 858.
- 4. Turing, A. M. *Quart. J. Mech. & Applied Math.* **1** (1948) 287.
- 5. Box, G. E. P. and Hunter, J. S. *Biometrika* **41** (1954) 190.
- 6. Sullivan, J. C. and Hindman, J. C. *J. Am. Chem. Soc.* **76** (1954) 5931.
- 7. Friedlander, G. and Kennedy, T. W. *Nuclear & Radiochemistry*, J. Wiley & Sons, New York 1955.
- 8. Deming, W. E. *Statistical Adjustment of Data*, J. Wiley & Sons Inc., New York 1943
- 9. *J. Assn. Computing Machinery* **3** (1956) 45.
- 10. Leden, I. *Z. physik. Chem.* **A 188** (1941) 160.
- 11. Olerup, H., Diss. Lund 1944.
- 12. Rydberg, J., and Sullivan, J. C. *Acta Chem. Scand.* **13** (1959) 2057.
- 13. For examples see:
 - (a) Connick, R. E. and McVey, W. H. *J. Am. Chem. Soc.* **71** (1949) 3182.
 - (b) Connick, R. E. and Reas, W. H. *J. Am. Chem. Soc.* **73** (1951) 1171.
 - (c) Betts, R. H. and Leigh, R. *Can. J. Research* **B 28** (1950) 514.
 - (d) Zebroski, E. J., Alter, H. W. and Heumann, F. K. *J. Am. Chem. Soc.* **73** (1951) 5946.
 - (e) Day, R. A., Jr. and Powers, R. M. *J. Am. Chem. Soc.* **76** (1954) 3895.
- 14. For a comprehensive list of references in this field see:
 - (a) Rydberg, J. *Svensk Kem. Tidskr.* **67** (1955) 499.
 - (b) Dyrssen, D. *Svensk Kem. Tidskr.* **68** (1956) 212.
- 15. Rydberg, J. *Acta Chem. Scand.* **4** (1956) 1503.

Received August 6, 1959.