

Some Graphical Methods for Determining Equilibrium Constants II. On "Curve-fitting" Methods for Two-variable Data

LARS GUNNAR SILLÉN

Department of Inorganic Chemistry, Royal Institute of Technology, Stockholm 70, Sweden

A survey is given of graphical methods for determining a number of unknown constants (parameters) $p_1, p_2 \dots$ from a set of data $y(x)$ when the function $y(x, p_1, p_2 \dots)$ is known. If x, y , and p_1 may be so chosen that one or two parameters can be connected with the variables to form "normalized variables" $X = x + p_1, Y = y + p_2$, the parameters may be found by fitting normalized curves. Such "curve-fitting" is preferable to numerical or "elimination" methods (see Table 1). For instance, with three parameters, $y(x)$ may be fitted to a family of curves $Y(X, p_3)$; p_1 and p_2 are obtained from the shifts of the coordinate axes, and p_3 from the shape of the curve. — A linear plot is a special case of curve-fitting; linear plots, however, have limitations which often make curved plots preferable.

Some examples are given from the determination of one or two complexity constants, and the determination of coefficients in simple polynomials.

If there are more than 2–3 parameters, "sweeping" and successive approximation may be used; the first procedure should not be used without the second. The use of smoothed curves, instead of the experimental points, requires great caution.

The determination of the equilibrium constants in a complicated system is a special case of the problem of finding a set of unknown constants from a series of experimental data. By choosing as many experimental points as one has unknown parameters, one may obtain a set of equations and solve them straight-forwardly^{1, 2} or by successive approximations^{3, p. 36}. Graphical methods have the advantage over such numerical ones that the information from a much greater number of data may be considered simultaneously.

The usual tendency among chemists who apply graphical methods seems to be that of transforming the experimental data into a diagram of two variables that should give a *straight line*. From the slope and intercept of the best straight line, one may then determine two unknown constants. If there are

more than two unknowns, they may be obtained by a series of successive linear extrapolations, which is a special case of "sweeping" below.

In previous papers by Dyrssen⁴ and Sillén⁴⁻⁶ a somewhat different method has been applied to some special cases, namely fitting a diagram of the experimental data to a family of calculated curves. The shape and position of the curve with the best fit then gives the required information. — In the work on equilibria with polynuclear complexes to be published, various "curve-fitting" methods have proved useful. It has then seemed timely to classify what such methods may achieve.

The present paper will first give a rather general survey of graphical ways of determining one, two, or three unknown parameters from a given collection of experimental data $y(x)$. The application of methods of curve-fitting will then be shown for a number of special cases. A linear plot is only a special case of curve-fitting; it will be argued that the straight line is sometimes, but not always, the best curve.

The way of treatment given below may be common usage in some other branches of pure or applied science. However, it does not seem to be well known to all chemists.

SYMBOLS

a	concentration of free A (6)
b	concentration of free B (6)
f	experimentally determined function in (21) and following
K	equilibrium constant (6)
K_1, K_2	stepwise formation constants of complexes (text above eq. 12)
k_0, k_1, k_2	unknown coefficients in (21) and following
p_1, p_2, p_3	unknown constants (parameters) to be determined
v	auxiliary variable (7, 12a, 23, 27a, 28a)
X, Y	normalized values for x and y (2, 3)
x, y	variables determined experimentally
x_0, y_0	values for x and y coinciding with $X = 0, Y = 0$ at best fit (2b, 3a)
Z	average number of A bound per B (7, 12)
β_1, β_2	complexity products (12)
Subscript) _{x}	indicates that x is constant

THE PROBLEM

Table I gives a schematic survey of possible ways of treating a collection of experimental data (x, y) in order to determine one or more unknown constants (parameters) p_1, p_2, \dots . If the experimental "points" are sets of two variables (x, y), they may be plotted in a planar diagram $y(x)$. — The case of three experimental variables (x, y, w), which cannot be reduced to a two-dimensional form, will be treated in a following paper.

We shall assume that x and y can be calculated directly from the experimental data, although they may not be the quantities actually measured. Nor need p_1, p_2, \dots be the equilibrium constants etc. ultimately sought for, although

Table 1. Methods for finding unknown constants p_i from experimental data $y(x)$.

Number of parameters	Number of normalized variables	Compare (x, y) graph with calculated graph	Eliminate y to give graph	Two-dimensional projection
1 (p_1)	1(x)	X, y	x, p_1	Curve $y(X)$ and plot $y(x)$
	0			Plot $p_1(x)$
2(p_1, p_2)	2(x, y)	X, Y	x, p_1, p_2	Curve $Y(X)$ and plot $y(x)$
	1(x)	X, y, p_2		a) Family of curves $y(X)_{p_2}$ and $y(x)$ plot b) $p_2(X)_y$ and proj. strip $(x)_y$
	0			a) Horizontal line in family $p_2(x)_{p_1}$ b) Intersection of curves $p_2(p_1)_x$
3(p_1, p_2, p_3)	2 (x, y)	X, Y, p_3	x, p_1, p_2, p_3	Family of curves $Y(X)_{p_3}$
	1(x)	X, y, p_2, p_3		—
	0			—

one must be able to calculate the latter, once the p_i are known. — In fact, it may often require a fair amount of trial to devise the most suitable functions x and y of the primary data, and the most suitable parameters p_i .

For each special problem we know the mathematical relationship between y , x , and the p_i , e. g. in the form

$$\begin{aligned}
 y &= y(x, p_1) \\
 y &= y(x, p_1, p_2) \\
 \text{or } y &= y(x, p_1, p_2, p_3)
 \end{aligned}
 \tag{1}$$

We shall define our *primary* problem as finding the best possible values for the p_i , assuming that the function (1) is correct, and not as testing the validity of (1).

Table 1 shows two main ways of treating the problem graphically: eliminating one variable (say y), or comparing a plot of $y(x)$ with a calculated "normalized graph".

We shall not discuss graphical methods like that of Scatchard⁷ in which the intercepts and slopes of two limiting tangents of a curve give approximate values for four parameters which are, however, then refined numerically. This method does not seem easily adaptable for graphical refinement of the constants.

Eliminating y

One parameter. When the values (x, y) for the experimental points are inserted into $y(x, p_1)$, each point gives a value for p_1 . If p_1 is now plotted

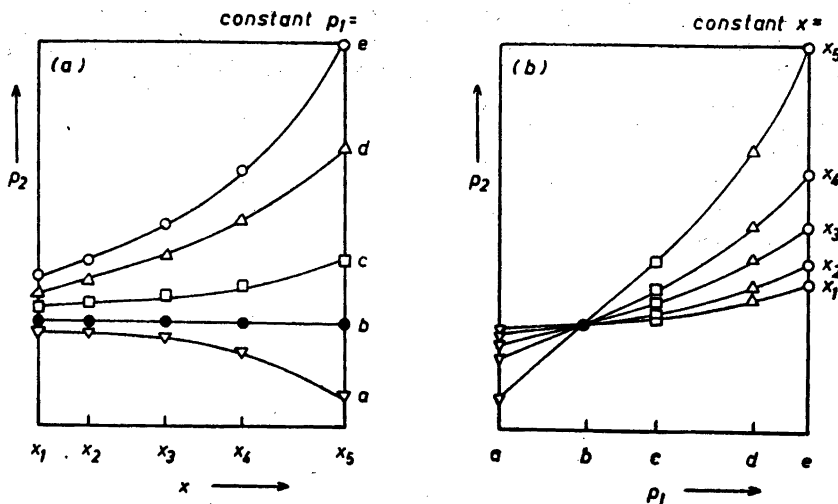


Fig. 1. Finding two parameters from $y(x)$ data by elimination: a) $p_2(x)p_1$ curves: horizontal line for correct value of p_1 ; b) $p_2(p_1)x$ curves: intersection at required point p_1, p_2 .

against x , one would expect to find the points scattered randomly around a horizontal line, since p_1 is really a constant. Usually no plot is made but the average value for p_1 is calculated numerically. The diagram $p_1(x)$ is drawn only if one wants to test for a trend in p_1 (which would show that the equation $y(x, p_1)$ in (1) is not quite correct), or to illustrate the spread in p_1 .

Two parameters. The insertion of a pair of experimental values (x, y) into $y(x, p_1, p_2)$ gives a functional relationship between p_1 and p_2 , which may be thought of as a planar curve $p_2(p_1)_x$ in (x, p_1, p_2) space. (The subscript " x " indicates that x is kept constant.) A number of these curves would give a three-dimensional surface $p_2(p_1, x)$. Since p_1 and p_2 are constants by definition, the correct set (p_1, p_2) must be possible for all values of x , and thus corresponds to a straight line which is parallel to the x axis and lies entirely on the surface $p_2(p_1, x)$.

This set may be found in two ways:

a) One may plot $p_2(x)p_1$, testing various constant values for p_1 , until one finds the value for p_1 that makes $p_2(x)$ a horizontal line (Fig. 1a).

b) One may also plot $p_2(p_1)_x$, each experimental point giving rise to one curve, and find the required set (p_1, p_2) at the common point of intersection of these curves (Fig. 1b). An example of the second method is given in (20) below (Schwarzenbach⁸). It is in general more expedient if one wants accurate values for p_1 and p_2 ; however, the first method may have advantages if one wants to test for systematic errors in certain parts of the range of x .

Three parameters. Each experimental point $y(x)$, when combined with $y(x, p_1, p_2, p_3)$ in (1), gives rise to a surface $p_1(p_2, p_3)$ in (p_1, p_2, p_3) space; this surface may be thought of as a section at constant x of a four-dimensional super-surface $p_1(p_2, p_3, x)$.

The set (p_1, p_2, p_3) sought for is the common point of intersection of all these surfaces $p_1(p_2, p_3)_x$. If one of the parameters, say p_3 , may be estimated rather accurately, so that only two or three values of p_3 need be considered, one might construct a plot $p_1(p_2)_{x,p_3}$ for each of these values of p_3 . By studying the range where these curves intersect in each section, one may locate the point of intersection of the surfaces $p_1(p_2, p_3)_x$.

However, in general it seems impractical to vary more than two parameters simultaneously using the elimination procedure. One would then rather resort to successive approximation (see below).

Normalized graphs

The calculations are much simplified if one may normalize one variable or both.

One parameter. Suppose, for instance, that by suitable choice of variables and parameter we may arrange that the function $y(x, p_1)$ contains as the only variable the sum of x and p_1

$$X = x + p_1 \quad (2)$$

The function, which we may thus write

$$y = y(X) \quad (2a)$$

($X =$ "normalized x "), may then be plotted as a graph (cf. Fig. 2a), and moved along the x axis of the experimental diagram $y(x)$ (cf. Fig. 2b) until the best fit is obtained (cf. Fig. 2c). In this position one reads the x value, x_0 , corresponding to $X = 0$, and finds

$$X = 0; x = x_0 = -p_1 \quad (2b)$$

as required by (2).

In this laboratory, such curve-fitting is generally made by superimposing the experimental and calculated diagrams, drawn on an ample scale, on a table of opalescent glass, 50 cm \times 60 cm, illuminated from below.

Two parameters. If it is possible to normalize both variables

$$X = x + p_1; Y = y + p_2 \quad (3)$$

the function $y = y(x, p_1, p_2)$ is reduced to the form $Y = Y(X)$. The curve $Y(X)$ is calculated, and moved on top of the experimental data — this time allowing parallel movement along both coordinate axes — until the best fit is obtained. The coordinates of the point (x_0, y_0) coinciding with the origin of the normalized graph ($X = 0, Y = 0$) are read

$$(X = 0, Y = 0); x = x_0 = -p_1; y = y_0 = -p_2 \quad (3a)$$

This procedure provides a great simplification compared with the "elimination" procedure above.

Even if one does not find a suitable way of normalizing two variables, the gain of normalizing one is also considerable. In this case we have a normalized function

$$y = y(X, p_2) \quad (4)$$

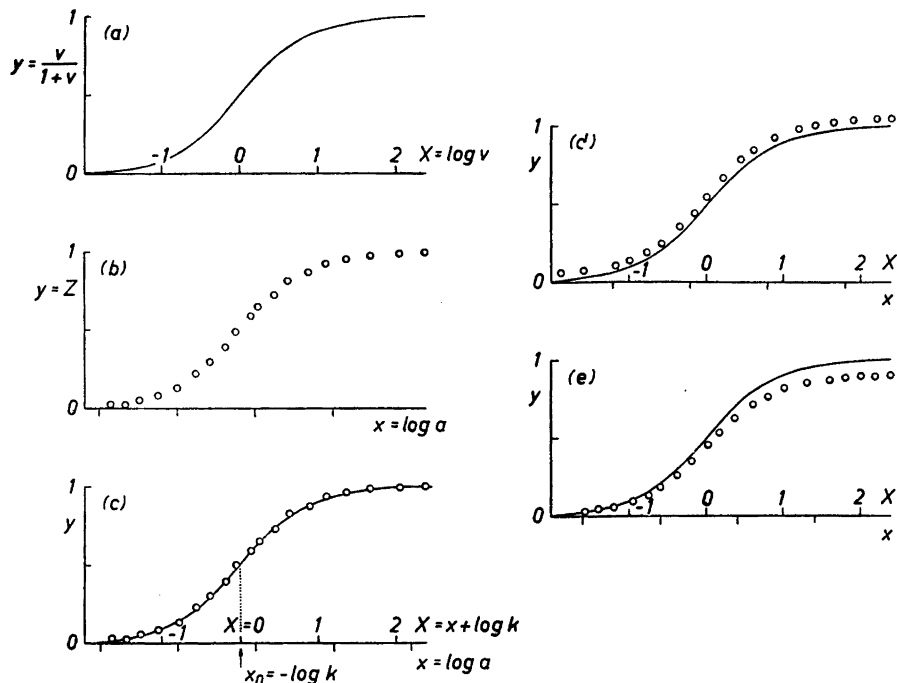


Fig. 2. Single complex: a) calculated curve $y(X)$, eq (8a) and (8b); b) imaginary experimental data $y(x)$, eq (8); c) best fit of a and b; d, e) effect of systematic analytical errors: d) constant error in assumed amount of A, e) constant error in assumed ratio of A to B.

to which the data $y(x)$ may be fitted. We may imagine $y(X, p_2)$ as a three-dimensional surface, to which the plane curve $y(x)$ may be fitted by shifts along the X and p_2 axes. One might consider working with a three-dimensional model, but two-dimensional plots are more easily handled and represented in print. Two approaches have been attempted:

a) In some earlier papers (Sillén⁶, Hietanen and Sillén⁹, Rossotti and Rossotti¹⁰, using the "two-parameter" approximation⁴) a family of curves $y(X)_p$ was compared with the experimental diagram $y(x)$. Then p_2 was obtained from the parameter of the curve that gave the best fit, and $-p_1$ from x_0 at the best position. For improving an approximate value of p_2 one may either draw curves for intermediate values of p_2 , or use some characteristic measure of the curve, such as the horizontal distance $x(y_2) - x(y_1)$ between two fixed values of y , or the slope dy/dx at a certain value of y .

b) A subsequent paper by Rossotti, Rossotti, and Sillén¹¹ will show the advantages of constructing curves $p_2(X)_y$ for a series of round values of y , and comparing them with a projection strip $(x)_y$ of the experimental data.

Three parameters, finally, may in principle be determined by fitting $y(x)$ to a three-dimensional surface $\bar{Y}(X, p_3)$, provided that both x and y can be normalized. Conveniently, $y(x)$ is compared with a family of normalized

curves $Y(X)_{p_i}$: p_1 and p_2 are found from the shifts along the coordinate axes (thus from x_0 and y_0), and p_3 from the shape of the curve. Examples are given in a paper by Dyrssen and Sillén⁴, and more will be given below.

If only one variable can be normalized, we have to fit $y(x)$ to a four-dimensional super-surface $y(X, p_2, p_3)$. This problem may be transformed to two-dimensional diagrams, *e. g.* by some method of successive approximation.

Linear plots are a special case of a family of normalized curves. If the two functions y and x are devised so as to give a linear plot, say

$$y = p_2x + p_1p_2 = p_2X \quad (5)$$

the construction of the best line through the data $y(x)$ may be considered as the fitting of $y(x)$ to a family of curves $y(X)_{p_i}$, which are in this special case a bundle of straight lines $y = p_2X$. In practice they are generally not constructed since a templet (*i. e.* a ruler) for these curves is always available.

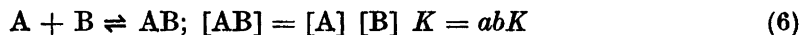
Fitting a straight line of a given slope to the $y(x)$ is a special case of fitting a single normalized curve $y(X)$.

In the following sections it will be shown by means of a few simple examples that the use of a suitably chosen set of non-linear normalized curves may have advantages compared with a linear plot. Finally some methods will be discussed — successive approximation, "sweeping" — which are resorted to when the number of parameters is so great that a single two-dimensional diagram will not suffice.

APPLICATIONS TO SIMPLE CASES OF COMPLEX FORMATION

Z (log a) data for single complex

If two reagents, A and B, form the single complex AB, the only constant to be determined is the equilibrium constant K



(If, for instance, $A = H^+$, $B = Ac^-$, and $AB = HAc$, then $K = K_{acid}^{-1}$ for HAc.) Let us assume that the concentration a of free A is measured for a series of solutions. For each solution the analytical data, together with a , give us Z , the number of A bound per B. The relationship is

$$Z = \frac{[AB]}{[B] + [AB]} = \frac{aK}{1 + aK} = \frac{v}{1 + v} \quad (7)$$

Data $Z(a)$ are often used for determining complexity constants and acidity constants. As seen from (7), Z is a function of the single variable $v = aK$. Choosing our experimental variables, and the single parameter as

$$x = \log a; y = Z; p_1 = \log K \quad (8)$$

we may use a normalized variable (2)

$$X = \log v = \log a + \log K \quad (8a)$$

The normalized curve $y(X)$, which may easily be constructed, is obtained from (7, 8, 8a):

$$y(X) = \frac{v}{1+v} = \frac{10^x}{1+10^x} \quad (8b)$$

This calculated curve $y(X)$, Fig. 2a, is moved parallel with a plot of the experimental data $y(x)$, Fig. 2b, until the best fit is obtained, Fig. 2c. Then one may read, from the value of x coinciding with $X = 0$

$$x_0 = -p_1 = -\log K \quad (8c)$$

Calculated curves of the shape (8b) have been used in several laboratories; *e. g.* Simms¹² mentions the use of a paper mould of this shape for finding acidity constants.

Comparison with alternative methods. One may find K from the same data $Z(\log a)$ by normalized graphs, using the same x and X (8, 8a), but choosing y so as to obtain a linear plot¹³

$$\left. \begin{aligned} x &= \log a; \quad X = \log a + \log K \\ y &= \log \frac{Z}{1-Z} = \log aK = X \end{aligned} \right\} \quad (9)$$

Plotting $y = \log Z/(1-Z)$ versus $x = \log a$, one would thus get a straight line of slope 1, which may be fitted to a normalized curve $y = X$. At $X = 0$ one reads $x_0 = -p_1 = -\log K$ as in (8c).

One might also use the elimination method, calculating

$$K = Za^{-1}(1-Z)^{-1} \quad (10)$$

from each experimental point, and taking the average.

In comparison with these methods, the choice of variables in (8) has some real advantages. The variables $\log a$ and Z are so close to the data primarily measured that one may mark out in the diagram how much an experimental point $y(x)$ would be changed by a certain experimental error — say by 0.3 mV in the emf in measuring a , or by 0.5 % in the analysis of A or B. So one may judge whether the deviations from the normalized curve are reasonable in view of the expected experimental accuracy.

Further, systematic analytical errors are easily recognized and corrected for since they give rise to a parallel shift along the y axis (Fig. 2d), or to a small change of scale along the y axis (Fig. 2e). With (9) or (10) such errors would give a trend in K but it would not be as easy to see the cause immediately. Treatment with the least square method, as usual, would do no good in the presence of systematic errors.

If it is found necessary to correct for small analytical errors of both kinds, one may connect these errors in Z with two parameters, one p_2 involving a shift along the y axis, and another p_3 a shift in the y scale. With

$$x = \log a; \quad y = Z; \quad p_1 = \log K \quad (8)$$

$$X = x + p_1 = \log v; \quad Y = y + p_2 = p_3 v(1+v)^{-1} \quad (11)$$

we have a system with two normalized variables and three parameters, which may be solved by means of a set of $Y(X)_{p_i}$ curves. The two other methods are not so easily adapted to correcting for analytical errors.

Z(log a) data for two complexes

If the reactants A and B form two complexes, AB and A₂B, with the overall formation constants ("complexity products") $\beta_1 = K_1$ and $\beta_2 = K_1K_2$ (K_1 and K_2 stepwise formation constants), we have

$$\left. \begin{aligned} & [\text{AB}] = \beta_1 ab; [\text{A}_2\text{B}] = \beta_2 a^2 b \\ Z = \frac{[\text{AB}] + 2[\text{A}_2\text{B}]}{[\text{B}] + [\text{AB}] + [\text{A}_2\text{B}]} &= \frac{\beta_1 a + 2\beta_2 a^2}{1 + \beta_1 a + \beta_2 a^2} \end{aligned} \right\} \quad (12)$$

Setting as before (cf. 8, 8a)

$$x = \log a; y = Z; X = x + p_1 = \log v \quad (12a)$$

we may normalize log a, leaving one parameter in the expression for y. Two possible ways are, expressing for convenience the relationship $y(X, p_2)$ by means of the auxiliary variable v in (12a)

$$p_1 = \log \beta_1; p_2 = \beta_2 \beta_1^{-2}; y = \frac{v + 2 p_2 v^2}{1 + v + p_2 v^2} \quad (13)$$

$$p_1 = \frac{1}{2} \log \beta_2 = \frac{1}{2} \log K_1 K_2; p_2 = \beta_1 \beta_2^{-\frac{1}{2}} = (K_1 K_2^{-1})^{\frac{1}{2}}; y = \frac{p_2 v + 2 v^2}{1 + p_2 v + v^2} \quad (14)$$

Fig. 3 shows a family of curves $y(X)_{p_2}$, calculated using the second approach (14). They may be regarded as sections of the three-dimensional surface $y(X, p_2)$. — The properties of functions like (12) have been discussed by Auerbach and Smolczyk¹⁴. A family of curves like Fig. 3 was given by J. Bjerrum^{3, p. 28}. Neither seems to have used them for curve-fitting. J. Bjerrum's average constant is 10^{21} , and his spreading factor is $\frac{1}{2} p_2$, in (14).

To find the required parameters, the experimental data $y(x)$ should be moved along the p_2 and X axes until the best fit is found. By comparison with the family of curves $y(X)_{p_2}$, Fig. 3, one may first test for, and if necessary correct for, systematic analytical errors, like in Figs. 2d and e. Then p_2 can be estimated from the two curves giving the nearest fit; a better value may be found by inserting curves for intermediate values of p_2 , or by using some suitable measure (say, the difference between $X(y = 1.5)$ and $X(y = 0.5)$, cf. Irving and Rossotti¹). The other parameter is found from the values of x_0 corresponding to $X = 0$.

For determining the parameters accurately, the method described by Rossotti, Rossotti and Sillén¹¹ is advantageous: a projection strip $(x)_y$ of the experimental data is moved parallel with the X and p_2 axes to give the best fit with the family of curves $p_2(X)_y$.

Alternative methods

Linear plot. We may transform (12) to the form

$$Z + \beta_1 a(Z - 1) + \beta_2 a^2(Z - 2) = 0 \quad (15)$$

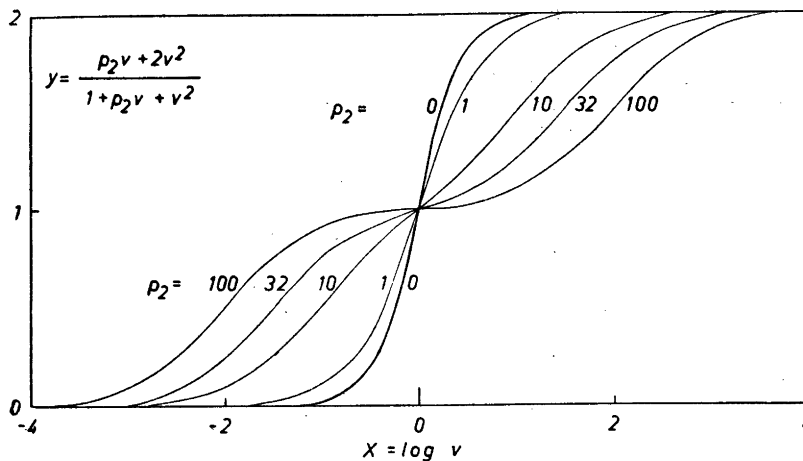


Fig. 3. Two complexes: calculated curves $y(X)_{p_2}$, (eq (12a) and (14) for various values of p_2 .

Dividing by suitable factors and rearranging, this equation may be reduced to the linear form, with X normalized

$$y = p_2x + p_1p_2 = p_2X \quad (5)$$

by any of six transformations, three of which are given here (in the other three, x and y are only interchanged)

$$a^2(2-Z)Z^{-1} = \beta_1\beta_2^{-1}a(Z-1)Z^{-1} + \beta_2^{-1} \quad (16)$$

$$Za^{-1}(Z-1)^{-1} = \beta_2a(2-Z)(Z-1)^{-1} - \beta_1 \quad (17)$$

$$Za^{-2}(2-Z)^{-1} = -\beta_1a^{-1}(Z-1)(2-Z)^{-1} + \beta_2 \quad (18)$$

For instance, with (16) we have

$$y = a^2(2-Z)Z^{-1}; \quad x = a(Z-1)Z^{-1}; \quad p_1 = \beta_1^{-1}; \quad p_2 = \beta_1\beta_2^{-1} \quad (16a)$$

Equation (16) is essentially identical with that given by Speakman¹⁵ and (17) with that of Irving and Rossotti¹, also used by Rossotti and Rossotti¹⁰.

The method given by Schwarzenbach, Willi, and Bach⁸ may be classified as an *elimination* method, using the parameters

$$p_1 = \beta_1^{-1}; \quad p_2 = \beta_2\beta_1^{-1} \quad (19)$$

and dividing (17) by β_1 to obtain

$$p_1 \frac{Z}{a(Z-1)} = p_2 \frac{a(2-Z)}{Z-1} - 1 \quad (20)$$

Each experimental point (Z, a) gives a straight line (20) in a plot (p_1, p_2) , which is conveniently constructed from the intercepts $-a(Z-1)Z^{-1}$ and $(Z-1)a^{-1}(2-Z)^{-1}$. All these lines should intersect at the point (p_1, p_2)

which gives the correct set (β_1, β_2) by means of (19). (For a dibasic acid H_2X , as discussed in Ref.⁸, with $A = H^+$ and $B = X^{2-}$, the acid dissociation constants are related to the formation products by $K_{a1} = \beta_1\beta_2^{-1}$ and $K_{a2} = \beta_1^{-1}$).

Linear versus curved plots. The search for the best straight line (5) is a special case of curve-fitting. The advantage of a linear plot is the simplicity of construction. However, considering the time that will usually be spent in getting accurate data, the little extra work of constructing a set of normalized curves is negligible.

There are several limitations to the linear plots, which may often make it worthwhile to try a curved plot.

1) The connection between the functions used and the experimental data will often be so remote in a linear plot that it may be hard to decide whether a certain deviation is compatible with the experimental accuracy. With curved plots one has a greater freedom to choose variables more *directly connected with the experimental data*. Compare, for instance, the variables $y = Z$ and $x = \log a$ in the curved plots (Figs. 2 and 3) with the complicated functions of Z and a used in the linear plots (eqns. 16, 17, 18).

2) *Systematic errors* may make a linear plot deviate from linearity; however, such deviations are not always readily explained and corrected for. A curved plot may be devised so that systematic errors to be expected (say, analytical errors) are easily recognized and corrected for. The systematic error may be treated as one or two extra parameters to be determined, as in (11).

3) In studies of chemical equilibria, the concentrations of the reacting species are often varied by several powers of ten — and should be so varied, wherever possible. Then the scale of a linear plot may become awkward: the experimental points being scarce at one end and compressed at the other. A small experimental error may cause much greater deviations in some parts than in others. In estimating the best line, it is difficult to weigh deviations in different parts of the diagram against each other, especially if the variables plotted, y and x , are also rather remote from the measured quantities.

With curved plots this difficulty is overcome by using a *logarithmic concentration scale*, which is also very convenient for normalization. Especially with emf measurements, the accuracy at different parts of a logarithmic scale is approximately the same, and this facilitates the estimation of the best curve.

4) A single linear plot can give only two parameters, whereas a curved plot will allow the simultaneous determination of *three parameters*.

5) Sometimes one may want to use the data for choosing between different hypotheses, each of which gives a different expression (1) for y . If one insists upon linear plots, the various hypotheses will in general give very different variables, and so it is hard to make a good comparison between them. Using curved plots, one has a greater freedom of choice. It is often possible to use a plot of the same variables to *test for several different hypotheses*, e.g. in ref.^{6,9}, where mechanisms with a single polynuclear complex, with series of "core + links" polynuclear complexes, and with the formation of a precipitate, could be tested on the same plot $y(x)$; in this case, $y = Z/t$, $x = t \log a + \log B$. Thus it is easier to judge which deviations are really to be considered as essential.

The scale should of course be so chosen that full advantage is made of the experimental accuracy. Thus, for a rather accurate set of data it may prove desirable to use a plot on a much larger scale than would seem necessary for another set of data, otherwise using a plot with the same normalized functions.

DETERMINING THE COEFFICIENTS IN POLYNOMIALS

Suppose that we have measured, over a certain range of the variable a , a number of experimental points (a, f) , where f is related to a by

$$f(a) = k_0 + k_1 a + k_2 a^2 \tag{21}$$

k_0, k_1 and k_2 being unknown constants to be determined.

For instance, in studies of systems with complexes A_nB , a may be the concentration of free A, and $-59.16 \log f$ may be the change in the emf of a B electrode; or $a^2 f^{-1}$ may be the distribution ratio of B_{total} between an organic solvent and an aqueous solution of A.

Problems of this kind may be solved by using normalized curves. Setting as experimental variables

$$x = \log a; y = \log f \tag{22}$$

we may use the following normalization

$$\left. \begin{aligned} X &= x + p_1 = \log v \\ Y &= y + p_2 = \log(1 + p_3 v + v^2) \end{aligned} \right\} \tag{23}$$

Solving $p_1, p_2,$ and p_3 from (21), (22), and (23) we find

$$\left. \begin{aligned} -x_0 &= p_1 = \frac{1}{2} \log k_2 - \frac{1}{2} \log k_0 \\ -y_0 &= p_2 = -\log k_0 \\ \log p_3 &= \log k_1 - \frac{1}{2} \log k_0 - \frac{1}{2} \log k_2 \end{aligned} \right\} \tag{23a}$$

Fig. 4a gives a family of curves $Y(X)_{p_3}$, calculated from (23). The curves have two common asymptotes, $Y = 0$ and $Y = 2X$. If a given set of experimental data $y(x)$, eqn. (22), is moved so as to give the best fit (Fig. 4b), the shape of the curve gives p_3 , and the coordinates (x_0, y_0) coinciding with $(X = 0, Y = 0)$ give $-p_1$ and $-p_2$.

To refine the value of p_3 one may interpolate in some way between the calculated curves, possibly inserting new curves with intermediate values of p_3 . If the accuracy admits it, one may use the vertical distance $y - y_0$ from the experimental curve to (x_0, y_0) ; at this point $X = 0$, and (23) gives

$$\text{for } X = 0; v = 1; Y = y - y_0 = \log(2 + p_3) \tag{23b}$$

The required constants are calculated using (23a)

$$\left. \begin{aligned} \log k_0 &= -p_2 \\ \log k_1 &= p_1 - p_2 + \log p_3 \\ \log k_2 &= 2p_1 - p_2 \end{aligned} \right\} \tag{23c}$$

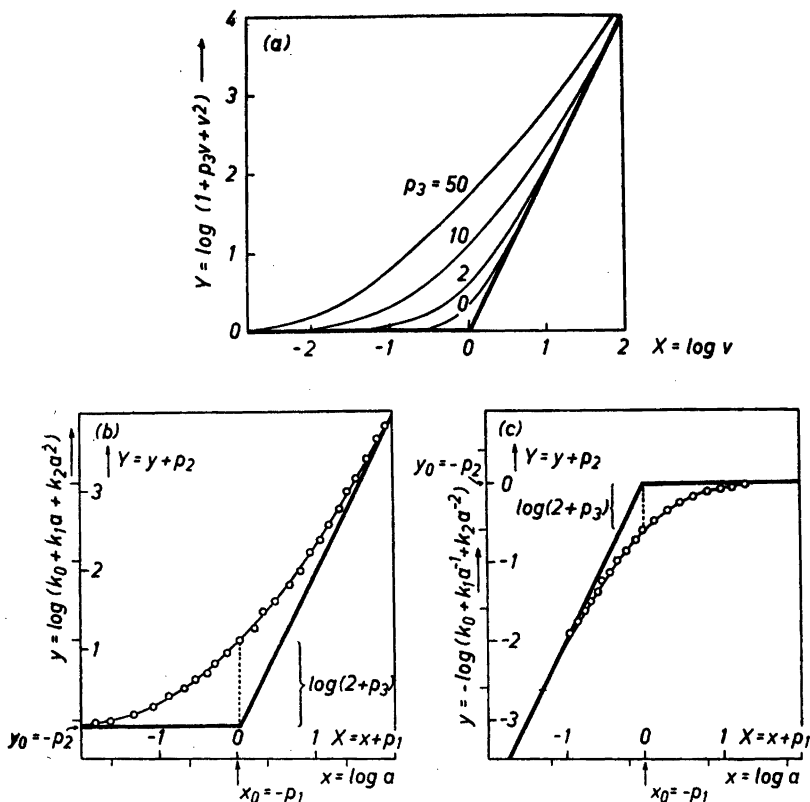


Fig. 4a. Family of curves (23), $X = \log v$, $Y = \log(1 + p_3v + v^2)$ for finding coefficients of polynomials $k_0 + k_1a + k_2a^2$.

b. Fit of experimental data.

c. Fitting inverted curve, $X = -\log v$, $Y = -\log(1 + p_3v + v^2)$, eq (26), to data $x = \log a$, $y = \log f = -\log(k_0 + k_1a^{-1} + k_2a^{-2})$.

It should be noted that the same set of curves may be used also to find the coefficients of functions of the types

$$f(a) = k_0 + k_1a^{-1} + k_2a^{-2} \quad (24)$$

$$f(a) = (k_0 + k_1a + k_2a^2)^{-1} \quad (25)$$

$$f(a) = (k_0 + k_1a^{-1} + k_2a^{-2})^{-1} \quad (26)$$

One needs only reverse the signs of the X axis for (24), of the Y axis for (25), or of both for (26). For (26) we have for instance (Fig. 4c) $X = -\log v$, $Y = -\log(1 + p_3v + v^2)$.

In practice, the reversal is simply made by rotating the paper, on which the curves have been drawn, by 180° around a vertical or horizontal axis.

Separate curves should be drawn for the values $k_0 = 0$ and $k_2 = 0$ which appear only as limiting cases in the treatment by means of (23). With $k_2 = 0$ we have

$$f(a) = k_0 + k_1 a \tag{27}$$

We may use the same coordinates

$$x = \log a; y = \log f \tag{22}$$

and normalize, using the auxiliary variable v

$$X = x + p_1 = \log v; Y = y + p_2 = \log(1 + v) \tag{27a}$$

Solving p_1 and p_2 from (22), (27), and (27a) we find

$$p_1 = \log k_1 - \log k_0; p_2 = -\log k_0 \tag{27b}$$

With $k_0 = 0$ we have correspondingly

$$f(a) = k_1 a + k_2 a^2 \tag{28}$$

Using the same set of experimental variables (22), and normalizing

$$X = x + p_1 = \log v; Y = y + p_2 = \log(v + v^2) \tag{28a}$$

we would find

$$p_1 = \log k_2 - \log k_1; p_2 = \log k_2 - 2 \log k_1 \tag{28b}$$

The curves (27a) and (28a) are shown in Fig. 5; they can of course, like those in Fig. 4a, be applied to functions with inverted values of f or a , just by reversing one coordinate axis or both.

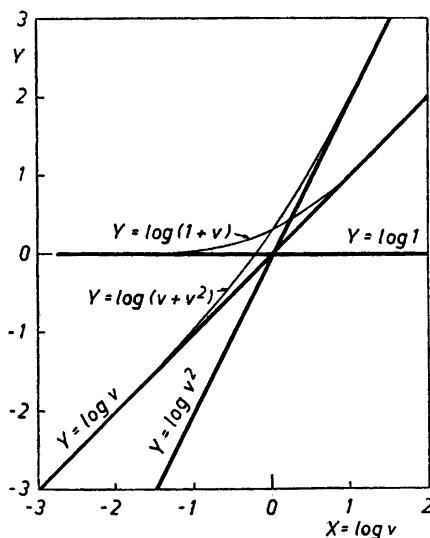


Fig. 5. Normalized curves (27a) and (28a) for finding coefficients in $f = k_0 + k_1 a$ and $f = k_1 a + k_2 a^2$.

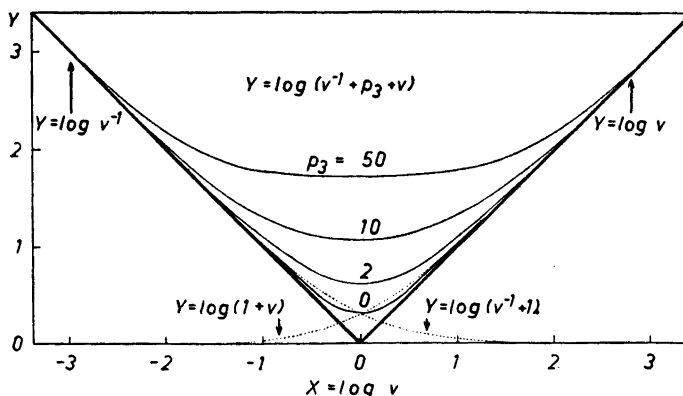


Fig. 6. Normalized curves (29) for finding coefficients of polynomial $k_0 + k_1a + k_2a^2 = f$, by comparison with $\log fa^{-1}$ ($\log a$) diagrams. Curves are also drawn for cases with one or two coefficients = 0.

Using the double-logarithmic plot (22), all three unknowns in $f = k_0 + k_1a + k_2a^2$ may be found by comparing with the curves in Figs. 4a and 5. — It may seem surprising that a curve-fitting method is advocated for problems which have usually been solved by linear plots, and would seem admirably suited to such. However, with experimental data extending over a number of powers of 10, a logarithmic plot may give a better estimate of the real accuracy (see above, p. 196).

The same approach may be used for any polynomial with only three parameters, say $f(a) = k_1a + k_3a^3 + k_4a^4$ etc. It may prove useful in some cases to use as y , not $\log f$ but for instance $\log fa^{-1}$ or $\log fa^2$. Thus the polynomial (24) may also be treated by considering $y = \log fa^2 = \log(k_2 + k_1a + k_0a^2)$, which may be treated with the curves (23).

Even the polynomial (21) $f = k_0 + k_1a + k_2a^2$ may sometimes preferably be treated using (Fig. 6)

$$\left. \begin{aligned} x &= \log a; X = \log v \\ y &= \log fa^{-1}; Y = \log(v^{-1} + p_3 + v) \end{aligned} \right\} \quad (29)$$

Equations (23 a, b, c) still hold. Thus p_3 is found from the intersection of the Y axis for $X = 0$: $Y = \log(2 + p_3)$.

SMOOTHED CURVES

It seems a good rule to select such graphical methods as allow one to retain a good feeling for the possible effect of experimental errors. Thus one would like to plot each experimental point as a separate point, with such coordinates that the influence of an experimental error can be readily estimated. In studies of complexity equilibria, this principle has been especially strongly advocated by Leden^{e.g. 16}.

Thus, it does not seem advisable to draw from the onset a smoothed curve through a set of data, and then to keep strictly to that curve. This approach is

especially dangerous if the data have only intermediate accuracy, and the points are rather scarce in some part of the diagram. On scrutinizing such "smoothed curves" one may find that the position of certain points will unduly influence the way of drawing the curve. If it so happens that from a certain "smoothed curve" one has calculated values for the unknown constants that give a fair agreement with the original smoothed curve, one is too often inclined to think that one has found the right set of parameters, without trying to imagine how wide a domain in parameter space ($p_1, p_2 \dots$) would also have given an acceptable agreement. Only if a smoothed curve gives constants that fail to reproduce it, it is emphasised that the curve must be incorrectly drawn, *i. e.* that it does not belong to the right family of curves.

Sometimes it may be unavoidable to use a smoothed curve, especially when a differentiation or integration is necessary. To avoid the dangers of smoothing the wrong way, one should use the following precautions:

1) Besides the best "smoothed curve" one should also draw two or more curves that deviate from it but would still give an acceptable agreement with the data. A comparison with constants calculated from these curves will give an indication of the real accuracy.

2) The results obtained with a smoothed curve should be compared with those obtained by other, independent methods.

METHODS FOR SEVERAL PARAMETERS

Successive approximation

The experimental range may often be divided into several parts so that one or two parameters, which give the main effect in one part of the range, give only a small correction in the other. Suppose for instance that we have an experimental quantity $y(x, p_1, p_2, p_3, p_4)$, and that p_1 and p_2 predominate in one range of x whilst p_3 and p_4 give a small correction, whereas in another range it is the other way round. Then p_1 and p_2 may be determined from the first range by one of the methods for a two-parameter problem, using approximate values for p_3 and p_4 or even neglecting them; then p_3 and p_4 are determined by some two-parameter method in the other range, assuming the first approximate values for p_1 and p_2 . Using these values for p_3 and p_4 in the first range, better approximations for p_1 and p_2 are found etc.

In this way four or more parameters may be obtained by successive approximation using two-parameter methods. This procedure is not necessarily unsatisfactory, since one need never lose the contact with the experimental data.

"Sweeping" methods

Especially for the determination of consecutive constants of complex formation, another type of method has been used, which may be described as "sweeping" for the parameters. The term will indicate that, just like with a sweeping search-light, the interest is centered on one or two constants at a time, and that the centre of interest is moved along the series of parameters.

The function $y(x, p_1, p_2, p_3, p_4 \dots)$ is first reduced to a two-parameter problem by neglecting other parameters than p_1 and p_2 . In the range where

p_1 and p_2 are predominant, a plot (usually linear) is made from which approximate values for p_1 and p_2 are found. The value for p_1 is accepted and inserted into the function y . Now the function $y(x, p_2, p_3 \dots)_p$ is considered, keeping p_1 constant, neglecting p_4 and following parameters, and having p_2 and p_3 as unknown parameters to be determined. This plot gives a value for p_2 , which is then used in a new plot with p_3 and p_4 as unknown parameters. The procedure may be speeded up if one accepts both values for p_1 and p_2 obtained from the first plot, and thus gets two parameters from each plot. It is often advantageous to use different experimental functions y and x in different parts of the range.

With such a method there is a danger that errors in the values for the first parameters p_1, p_2 etc will accumulate, making the plots for p_4 and p_5 somewhat unreliable; at any rate, the feeling of contact with the experimental data is definitely lost, and it is very hard to estimate the real accuracy.

The parameters obtained may be checked and improved:

1) by sweeping the series of parameters from both ends; an example is given in part I¹⁰. The deviations will give an indication of the accuracy.

2) by successive approximation. After a first approximate set of parameters has been obtained, the terms corresponding to the higher parameters are inserted into the first plot for finding p_1 and p_2 , which gives better values for these parameters. The process may be repeated, either always from one end, or sweeping from alternate ends of the parameter series, depending on the problem and type of data. As a rule, this simple precaution appears to have been neglected previously.

If curved plots are also considered, one would have a greater freedom of choice.

I wish to thank my friends George Biedermann, David Dyrssen, Brita Hök-Bernström, Gösta Lagerström, Francis Rossotti, Hazel Rossotti, and Jan Rydberg for pleasant cooperation and for many valuable discussions.

This paper is part of a program supported by *Atomkommittén (Swedish Atomic Energy Commission)* and *Statens Naturvetenskapliga Forskningsråd (Swedish Natural Science Research Council)*.

REFERENCES

1. Irving, H. and Rossotti, H. S. *J. Chem. Soc.* **1953** 3397.
2. Block, B. P. and McIntyre, G. H. Jr, *J. Am. Chem. Soc.* **75** (1953) 5667.
3. Bjerrum, J. *Diss.* Copenhagen 1941.
4. Dyrssen, D. and Sillén, L. G. *Acta Chem. Scand.* **7** (1953) 663.
5. Sillén, L. G. *Acta Chem. Scand.* **8** (1954) 299.
6. Sillén, L. G. *Acta Chem. Scand.* **8** (1954) 318.
7. Edsall, J. T., Felsenfeld, G., Goodman, D. S. and Gurd, F. R. N. *J. Am. Chem. Soc.* **76** (1954) 3054.
8. Schwarzenbach, G., Willi, A. and Bach, R. O. *Helv. Chim. Acta* **30** (1947) 1303.
9. Hietanen, S. and Sillén, L. G. *Acta Chem. Scand.* **8** (1954) 1607.
10. Rossotti, F. J. C. and Rossotti, H. S. *Acta Chem. Scand.* **9** (1955) 1166, part I.
11. Rossotti, F. J. C., Rossotti, H. S. and Sillén, L. G. *Acta Chem. Scand.* **10** (1956) 203, part III.
12. Simms, H. S. *J. Am. Chem. Soc.* **48** (1926) 1239.
13. McClendon, J. F. *J. Biol. Chem.* **54** (1922) 647.
14. Auerbach, F. and Smolczyk, E. *Z. physik. Chem.* **110** (1924) 65.
15. Speakman, J. C. *J. Chem. Soc.* **1940** 855.
16. Leden, I. *Diss.* Lund 1943.

Received November 3, 1955.

Acta Chem. Scand. **10** (1956) No. 2