[CONTRIBUTION FROM THE SCHOOL OF CHEMISTRY AND PHYSICS, THE PENNSYLVANIA STATE COLLEGE]

The Calculation of Formation Constants for Systems Involving Polydentate Ligands[1]

BY B. P. BLOCK AND G. H. MCINTYRE, JR.

RECEIVED JULY 10, 1953

Bjerrum's formation function has been solved in general terms for the individual formation constants for the cases $N = 1, 2$ and 3. The values for the constants calculated using the resulting expressions agree with the values obtained by the methods currently in use, but generally the calculations are much shorter. For the most efficient use of these equations data must be taken in certain optimum $\bar{n}$ ranges.

Sullivan and Hindman[2] have recently compared the three major methods proposed for the determination of formation constants and have shown that all three give the same results within experimental limitations. They point out that it is possible to solve the N linear equations in $k_1, k_2, \ldots, k_N$ by determinants and thus to avoid the use of approximation methods. For the past few years work in this Laboratory has centered on the determination of formation constants using Bjerrum's formation function.[3] Since most workers in this field are now using Bjerrum's methods practically verbatim, we wish to report our experiences with the direct solution of the formation function expressions because we have found that direct solution appreciably lessens the labor involved in many of the calculations.

The generalized formation function developed by Bjerrum is

$$\bar{n} = \left(\overset{N}{\underset{1}{\Sigma}} nK_n[\mathrm{A}]^n\right) \Big/ \left(1 + \overset{N}{\underset{1}{\Sigma}} K_n[\mathrm{A}]^n\right)^3$$

To make the notation in the present development more compact several new symbols, which are defined in Table I, are used. Some of Bjerrum's symbols with their definitions are also listed in Table I. The formation function may be rearranged to $\bar{n} = \overset{N}{\underset{1}{\Sigma}}(n - \bar{n})[\mathrm{A}]^n K_n$, which becomes $\bar{n} = \overset{N}{\underset{1}{\Sigma}} J_n K_n$ when J_n is introduced. The present treatment will only be concerned with values of 1, 2 and 3 for N, since these are the usual cases for coördination compounds containing polydentate ligands. It is possible to extend this treatment to higher values of N, but the expressions become quite cumbersome. To set up a system of equations to solve for the N individual constants, sets of $\bar{n}$ and J_n values at N points are required. The potentiometric determination of [A] and $\bar{n}$ is the most widely used method for determining these quantities at present. If the ligand is a weak base or a weak acid for which the dissociation constant or dissociation constants are known, [A] and $\bar{n}$ can be calculated for any given solution from a knowledge of $[\mathrm{H}^+]$ and the original concentrations of the substances in the solution. Other experimental methods may also be used to determine [A] and $\bar{n}$; spectrophotometric measurements, in particular, may be used in favorable cases. A knowledge of [A] and $\bar{n}$ permits the direct calculation of J_n, L_n and M_{np} since they are fundamentally defined in terms of the former.

TABLE I

DEFINITIONS OF SYMBOLS USED

M	general symbol for a metallic ion
A	general symbol for a ligand
[ ]	molar concn. of species indicated in the brackets
n	actual number of ligands bound to a given metallic ion
$\bar{n}$	av. number of ligands bound per metallic ion in a given soln.
N	max. coördination number of metallic ion for ligand in question
k_n	$= [\mathrm{MA}_n]/[\mathrm{MA}_{n-1}][\mathrm{A}]$, stepwise formation constant
K_n	the product $k_1k_2k_3, \ldots, k_n$
J_n	$= (n - \bar{n})[\mathrm{A}]^n$
J_n'	$= (n - \bar{n}')([\mathrm{A}]')^n$
J_n^*	$= (n - \bar{n}^*)([\mathrm{A}]^*)^n$
L_n'	$= \bar{n}J_n' - \bar{n}'J_n$
L_n^*	$= \bar{n}J_n^* - \bar{n}^*J_n$
$L_n'^*$	$= \bar{n}'J_n^* - \bar{n}^*J_n'$
M_{np}'	$= J_nJ_p' - J_n'J_p$
M_{np}^*	$= J_nJ_p^* - J_n^*J_p$
$M_{np}'^*$	$= J_n'J_p^* - J_n^*J_p'$

The general expressions for the individual formation constants are listed in Table II. They are obtained by solving the appropriate N sets of $\bar{n}$ equations by algebraic methods. For the cases under consideration the $\bar{n}$ equations are $\bar{n} = J_1k_1$ for $N = 1$, $\bar{n} = J_1k_1 + J_2k_1k_2$ and $\bar{n}' = J_1'k_1 + J_2'k_1k_2$ for $N = 2$, and $\bar{n} = J_1k_1 + J_2k_1k_2 + J_3k_1k_2k_3$, $\bar{n}' = J_1'k_1 + J_2'k_1k_2 + J_3'k_1k_2k_3$, and $\bar{n}^* = J_1^*k_1 + J_2^*k_1k_2 +$

TABLE II

EQUATIONS FOR FORMATION CONSTANTS

N	Equation
1	$k_1 = \bar{n}/J_1$
2	$k_1 = (\bar{n}J_2' - \bar{n}'J_2)/(J_1J_2' - J_1'J_2$
	$k_2 = (\bar{n}'J_1 - \bar{n}J_1')(\bar{n}J_2' - \bar{n}'J_2)$
3	$k_1 = (L_3^*M_{23}' - L_3'M_{23}^*)/(M_{13}^*M_{23}' - M_{13}'M_{23}^*)$
	$= (L_3'M_{23}'^* - L_3'^*M_{23}')/(M_{13}'M_{23}'^* - M_{13}'^*M_{23}')$
	$= (L_3^*M_{23}'^* - L_3'^*M_{23}^*)/(M_{13}^*M_{23}'^* - M_{13}'^*M_{23}'^*)$
	$k_2 = (L_3^*M_{13}' - L_3'M_{13}^*)/(L_3'M_{23}^* - L_3^*M_{23}')$
	$= (L_3'M_{13}'^* - L_3'^*M_{13}')/(L_3'^*M_{23}' - L_3'M_{23}'^*)$
	$= (L_1'L_3^* - L_1^*L_3')/(L_2^*L_3' - L_2'L_3^*)$
	$= (L_1'^*L_3' - L_1'L_3'^*)/(L_2'L_3'^* - L_2'^*L_3')$
	$= (L_1{}^*L_3^* - L_1^*L_3'^*)/(L_2^*L_3'^* - L_2'^*L_3^*)$
	$k_3 = (L_1^*L_2' - L_1'L_2^*)/(L_1'L_3^* - L_1^*L_3')$
	$= (L_1'L_2'^* - L_1'^*L_2')/(L_1'^*L_3' - L_1'L_3'^*)$
	$= (L_1^*L_2'^* - L_1'^*L_2^*)/(L_1'^*L_3^* - L_1^*L_3'^*)$

---

[1] Presented in part by B. P. Block at the Meeting in Miniature of the Philadelphia Section of the American Chemical Society, January 29, 1953.

[2] J. C. Sullivan and J. C. Hindman, THIS JOURNAL, **74**, 6091 (1952).

[3] J. Bjerrum, "Metal Ammine Formation in Aqueous Solution," P. Haase and Son, Copenhagen, 1941, p. 21.

$J_3^* k_1 k_2 k_3$ for $N = 3$. The primed and starred symbols represent the requisite second or third set of [A] and $\bar{n}$ values, respectively, for that particular value of N. There are alternative expressions for the three formation constants when the value of N is 3 because the set of three $\bar{n}$ equations may be solved for the constants by different methods of elimination. Although there are six possible expressions for k_1 and k_3 as well as for k_2, only three are given for each of the former two because in both cases the other three expressions are much more complicated. They are of the form

$$k_1 = \frac{(L_2^* M_{13}' - L_2' M_{13}^* + L_3' M_{12}^* - L_3^* M_{12}')}{2(M_{12}^* M_{13}' - M_{12}' M_{13}^*)}$$

$$\pm \frac{\sqrt{(L_2' M_{13}^* - L_2^* M_{13}' + L_3^* M_{12}' - L_3' M_{12}^*)^2 - 4(M_{12}^* M_{13}' - M_{12}' M_{13}^*)(L_2^* L_3' - L_2' L_3^*)}}{2(M_{12}^* M_{13}' - M_{12}' M_{13}^*)}$$

So far it has not been necessary to resort to the use of such equations to evaluate k_1 and k_3 because in each case one of the other expressions yielded a satisfactory calculation for the desired constant.

To test these equations, data previously evaluated by Bjerrum's methods have been used. The results are compiled in Table III. In every case the corresponding values are in good agreement. All the calculated values are averages of calculations from several sets of $\bar{n}$, [A] values. The indicated range is the 95% confidence interval calculated from the individual values for each constant by the method indicated by Youden.[4]

TABLE III

COMPARISON OF METHODS OF CALCULATION

Temp., °C.	M	A	N	n	log k_n Reported	Calcd.
0	Cu^{++}	bapa^e	1	1	15.82^a	15.81 ± 0.07
30	Co^{++}	bapa	1	1	6.63^a	6.26 ± .08
30	Cu^{++}	en^e	2	1	10.36^b	10.35 ± .03
				2	8.92^b	8.94 ± .05
30	Ag^+	NH_3	2	1	3.20^c	3.18 ± .05
				2	3.83^c	3.84 ± .05
30	Ni^{++}	en	3	1	7.66^d	7.68 ± .08
				2	6.40^d	6.39 ± .10
				3	4.55^d	4.56 ± .01

[a] G. B. Hares, Doctoral Dissertation, The Pennsylvania State College, 1952. [b] Constants evaluated by Bjerrum's methods from data in the Doctoral Dissertation of G. H. M., The Pennsylvania State College, 1953. [c] Ref. 3, p. 134. [d] Ref. 3, p. 214. [e] $(H_2NCH_2CH_2CH_2)_2NH$ is represented by bapa, $H_2NCH_2CH_2NH_2$ by en.

In the course of these calculations it became apparent that certain combinations of [A], $\bar{n}$ values give better results than others in the calculation of the constants for the cases when N is 2 or 3. The results for the various calculations are related to the ranges for the $\bar{n}$ values in Table IV. In general one $\bar{n}$ value must be in the range $(n - 1)$ to (n) in order that a good value for k_n result. If the $\bar{n}$ values are within ±0.2 of an integer, the calculations frequently do not give good results. Values considered not usable are, for the most part, values which cannot be calculated with sufficient accuracy because some term in the calculation is a difference between two almost equal numbers. Values outside the 95% confidence interval also are not considered usable. It is apparently not necessary to restrict the values of $\bar{n}$ to the ranges mentioned above for systems such as Ag^+–NH_3, where $k_1 \cong k_2$. All the average values for constants calculated for this paper include only those combinations containing an $\bar{n}$ in the proper range for the constant being calculated.

Even after all values calculated for combinations containing no $\bar{n}$ in the proper range were eliminated there were still some uncalculable values for k_1, k_2 and k_3 in the Ni^{++}–$H_2NCH_2CH_2NH_2$ system. These stem from the use of one particular equation for each constant, as is shown in Table V. The '–'* equations, on the other hand, gave satisfactory results for all calculations but two. Further experience in this Laboratory with other systems where the value of N is 3 has produced the same results. This means that the order in which the $\bar{n}$ values are taken is important. The results reported here are based on the order $\bar{n} < \bar{n}' < \bar{n}^*$.

TABLE IV

EFFECT OF $\bar{n}$ RANGE

System	Number $\bar{n}$, [A] sets	$\bar{n}$ range^a	Constant	G^b	A^b	N^b
Cu^{++}–en, 30°	6	1 in 0–1	k_1	12	..	..
		0 in 0–1		..	3	..
		1 in 1–2	k_2	12	..	..
		0 in 1–2		2	1	..
		1 in 0–1 1 in 1–2	k_1, k_2	9	..	..
Ag^+–NH, 30°, 2N.NH_4NO_3	7	1 in 0–1	k_1	18	..	..
		0 in 0–1		3	..	..
		1 in 1–2	k_2	15	..	..
		0 in 1–2		6	..	..
		1 in 0–1 1 in 1–2	k_1, k_2	12	..	..
Ni^{++}–en, 30°, 1 N KCl	7	1 in 0–1	k_1	61	18	14
		0 in 0–1		4	5	3
		1 in 1–2	k_2	87	23	40
		0 in 1–2		22	10	28
		1 in 2–3	k_3	41	15	19
		0 in 2–3		7	9	14
		1 in 0–1	k_1	22	5	9
		1 in 1–2	k_2	48	3	21
		1 in 2–3	k_3	24	4	8

[a] The figures in this column indicate either that none or else a minimum of one of the $\bar{n}$ values used for the calculation was in the range listed. [b] In column G is listed the number of individual calculations giving values of the constant having the requisite number of significant figures and falling within the 95% confidence interval. In column A is listed the number of individual calculations giving values with an insufficient number of significant figures. In column N is listed the number of individual calculations which did not give a value for the constant because a difference between two numbers was approximately zero.

Calculations were made for the various systems using the equations for incorrect values of N. These are summarized in Table VI. In as far as possible $\bar{n}$ values were kept in the proper ranges for these calculations. The values calculated for a given

---

(4) W. J. Youden, "Statistical Units of Measurement," National Bureau of Standards Report 1520, March 26, 1952.

TABLE V

EFFECT OF EQUATION USED FOR $N = 3$

Equation[a]	Constant	G	A	N
'_*	k_1	..	3	9
	k_2[b]	..	2	10
	k_2[b]	12	..	..
	k_3	12	..	..
'_'*	k_1	10	2	..
	k_2	12	..	..
	k_2	12	..	..
	k_3	12	..	..
_'	k_1	12	..	..
	k_2	12	..	..
	k_2	..	1	11
	k_3	..	4	8

[a] The symbols in this column refer to the two kinds of symbols involved in the expression for k_n in Table II. [b] In each case the first line for k_2 corresponds to the first expression for k_2 in Table II using the two kinds of symbols indicated; the second corresponds to the other expression which uses the same kinds of symbols.

system are generally acceptable as long as the equations are for the correct or a larger N. If the equations used are for N less than the correct N, the values for the constants vary, especially as $\bar{n} \rightarrow n$. Spurious values are calculated for k_n if $n > N$. For example, in the Cu^{++}–$H_2NCH_2CH_2NH_2$ system a value of 4.56 is calculated for log k_3. If this value were correct, $\bar{n}$ would be approximately 2.5 when log [en] was −4.56. However, at this concentration $\bar{n}$ is actually found to be about 2.

TABLE VI

EFFECT OF USING INCORRECT N

System	N True	N Used	log k_1	log k_2	log k_3
Cu^{++}–bapa	1	1	15.81		
	1	2	15.80	13.39	
	1	3	15.81	Neg.	Neg.
Cu^{++}–en	2	1	10.36[a]		
	2	2	10.35	8.94	
	2	3	10.30	8.93	6.05
Ag^+–NH_3	2	1	3.69[a]		
	2	2	3.18	3.84	
	2	3	3.19	3.84	0.90
Ni^{++}–en	3	1	7.73[a]		
	3	2	7.66[a]	6.50[b]	
	3	3	7.68	6.39	4.56

[a] Calculated with $\bar{n} < 0.5$; for greater values of $\bar{n}$ the values calculated for k_1 are not constant. [b] No $\bar{n} > 2$ was used.

The calculations for the Ag^+–NH_3 system are outlined in Table VII. For calculations such as these to be meaningful it is necessary to know N and be certain that the only reactions involving M and A are those of the type $MA_{n-1} + A \rightleftarrows MA_n$. The same knowledge is required to use Bjerrum's convergence function. The value of N is established by determining $\bar{n}$ in solutions containing a large excess of A, since $\bar{n} \rightarrow N$ as [A] increases. In some cases there are other reactions occurring, as Schwarzenbach and co-workers[5] have shown. For these cases it is sometimes possible to set up 3 or fewer equations which may be solved directly for the constants involved; in other cases recourse to methods of approximation may be required.

TABLE VII

SAMPLE CALCULATION

Ag^+–NH_3 system[a], 2 N NH_4NO_3, 30°

$\bar{n}$	$[NH_3]$	J_1	J_2	log k_1	log k_2
0.495	1.53×10^{-4}	7.72×10^{-5}	3.52×10^{-8}		
1.477	5.86×10^{-4}	-2.80×10^{-4}	1.79×10^{-7}	3.19	3.83

[a] Values for $\bar{n}$ and [A] taken from ref. 3, p. 132.

The use of the equations in Table II offers several advantages over the use of the approximation and convergence methods of Bjerrum. There is less labor involved in the conversion of original data to formation constants, except when the approximate values are satisfactory. Even in the latter case it is frequently harder to obtain the approximate values than it is to calculate the constants directly when N is 1 or 2, and it is always harder when N is 3. The determination of the constants in systems that do not come to equilibrium rapidly is made more convenient because fewer solutions are required. In principle the N constants can be calculated from N sets of [A], $\bar{n}$ values where the N values of $\bar{n}$ fall one each in approximately the middle of the ranges 0 to 1, 1 to 2, . . ., $N - 1$ to N. If more than N sets of data are used for independent calculations of the formation constants, a good estimate of the reliability of the values can be made; whereas it is much more difficult to estimate the reliability of constants obtained from a smooth formation curve.

Acknowledgment.—The authors are greatly indebted to Dr. W. C. Fernelius for many helpful suggestions and to the Office of Naval Research for financial aid under contract No. N6 onr 26913.

STATE COLLEGE, PENNA.

---

[5] G. Schwarzenbach *et al.*, *Helv. Chim. Acta*, **32**, 1543 (1949); **33**, 947 (1950); **34**, 1492 (1951); **35**, 2333, 2337 (1952).