

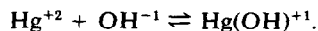
# USE OF THE NORMAL DISTRIBUTION IN THE STUDY OF SIMPLE STEPWISE EQUILIBRIA

SAMUEL M. GRAFF AND SELMAN A. BERGER

City University of New York  
John Jay College of Criminal Justice  
445 West 59 Street  
New York, New York 10019, USA

Communicated by X. J. R. Avula

**Abstract**—The normal distribution is used to estimate the pH range for the existence of metal ion concentrations in hydrolysis reactions. The standard deviation of the metal ion concentration is expressed as a function of the formation constant for the reaction. An example is presented using data for the equilibrium reaction



## 1. INTRODUCTION

The application of mathematics to provide models of physical phenomena has become of increased significance in the past decade. Conversely, the study of the physical world has enhanced the development of mathematical science as well. This paper presents a model for a chemical process which extends and enriches the classical approach employed in studying a broad class of similar chemical reactions.

The particular phenomenon to be discussed is simple stepwise equilibrium. In its simplest form such a reaction involves the addition of a ligand  $L$  to a metal  $M$  as symbolized by



Typically, the ligand  $L$  may be the hydroxide ion  $\text{OH}^{-1}$  and then Eq. (1) has the form



where  $n$  stands for the positive charge of the free metal ion. In order to simplify the exposition, we will confine our attention to the hydrolysis process of Eq. (2).

The classical chemistry theory postulates the existence of an equilibrium constant of formation  $K_f$  such that

$$K_f = \frac{[\text{MOH}]}{[M][\text{OH}]}, \quad (3)$$

where the square brackets denote concentrations. This equation is the only one which specifically links the concentrations of the three species. Since the concentration of the hydroxide ion is always related to the concentration of the hydrogen ion by the relationship

$$K_w = [\text{H}^+][\text{OH}^{-}], \quad (4)$$

one frequently expresses the equilibrium equation (3) in terms of  $[H^+]$ ,

$$K_f K_w = \frac{[MOH^{n-1}][H^+]}{[M^{+n}]}, \quad (5)$$

where  $K_w$  is the equilibrium constant for the dissociation of water. The constants  $K_f$  and  $K_w$  are determined experimentally and assumed to be known.

The principal question to which this paper is addressed is whether one can predict with some degree of confidence an interval for  $[H^+]$  within which most of the hydrolysis reaction (2) occurs. Traditionally, the question is posed in terms of  $pH \equiv -\log_{10}[H^+]$ , i.e., determine an interval for pH within which most of the hydrolysis occurs. Part of the difficulty in providing an answer is that the problem posed is not entirely well defined. A reasonable definition of what is meant by "most of the hydrolysis reaction" will have to be included as part of any model for the solution to the problem.

In the past, the problem has been approached by using the relationship between the ratio  $[M^{+n}]/[MOH^{n-1}]$  and pH given by Eq. (5),

$$\frac{[M^{+n}]}{[M(OH)^{n-1}]} = \frac{1}{K_f K_w} 10^{-pH} = \frac{1}{K_f K_w} e^{-pH/\log_{10}e} \quad (6)$$

(see Fig. 1). Basically, one assumes that most of the hydrolysis is confined to an interval where the dimensionless quantity  $[M^{+n}]/[M(OH)^{n-1}]$  satisfies

$$10^{-6} \leq [M^{+n}]/[M(OH)^{n-1}] \leq 10^6 \quad (7)$$

and then uses (6) to compute the corresponding pH interval.

There are objections that can be raised concerning the use of (7) to determine the pH interval where significant hydrolysis occurs. The criterion employed doesn't incorporate any measurable parameter of the problem. Indeed, it is not obvious that the inequality used to define significant hydrolysis should be the same for two entirely different reactions. A better approach would be to formulate a condition based upon a percentage of the total area underneath the graph of (6). However, the area is not finite.

Our model endeavors to circumvent the difficulties mentioned above. If the pH is considered to be the independent variable with the concentrations  $[M^{+n}]$  and  $[M(OH)^{n-1}]$  being dependent, then Eq. (6) provides only one relationship between these quantities. Thus, there is insufficient information to deduce the explicit dependence of  $[M^{+n}]$  and  $[M(OH)^{n-1}]$  upon pH from the classical theory. The question naturally arises as to whether it is possible to hypothesize the explicit dependence of  $[M^{+n}]$  and  $[M(OH)^{n-1}]$  upon pH. Such a hypothesis is justifiable if it affords results which are consistent with experimentally observable behavior.

The exponential form of the right hand side of Eq. (6) motivates the hypothesis regarding the dependence of  $[M^{+n}]$  and  $[M(OH)^{n-1}]$  upon pH. Since the quotient of the two quantities  $[M^{+n}]$  and  $[M(OH)^{n-1}]$  must be exponential, one might assume they each have the form

$$[M^{+n}] = \begin{cases} A e^{-\alpha(x-x_0)}, & x \geq x_0 \\ A & , x < x_0 \end{cases}$$

$$[M(OH)^{n-1}] = \begin{cases} B e^{-\beta(x-x_1)}, & x \leq x_1 \\ B & , x > x_1 \end{cases}$$

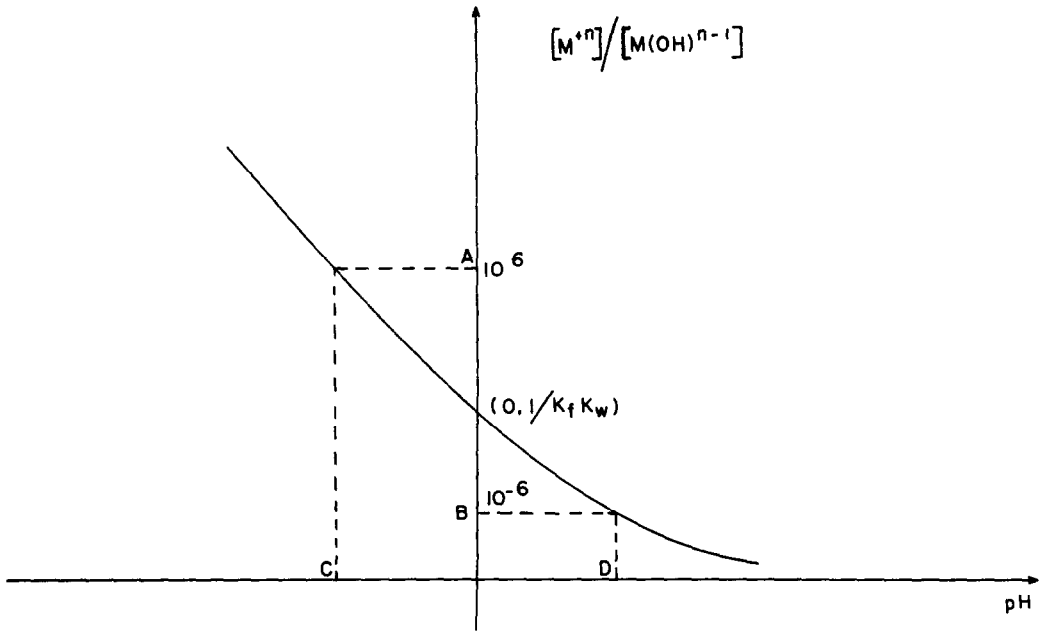


Fig. 1. Ratio of  $[M^{+n}] : [M(OH)^{n-1}]$  as a function of pH. Sector A-B, region where significant hydrolysis occurs. Sector C-D, corresponding pH interval.

with  $x_0 < x_1$  and  $x = \text{pH}$ . However, such an assumption is not adequate since for  $x_0 < x < x_1$  one has

$$\frac{[M^{+n}]}{[M(OH)^{n-1}]} = \frac{A}{B} \exp\{(\beta - \alpha)x + (\beta x_1 - \alpha x_0)\}$$

implying that the ratio would be constant if  $\alpha = \beta$ . (The fact that  $\alpha = \beta$  can occur is the subject of another paper.) The next most obvious hypothesis is to assume a dependence of the type

$$\begin{aligned} [M^{+n}] &= \begin{cases} A e^{-\alpha(x-x_0)^2}, & x \geq x_0 \\ A & , x < x_0 \end{cases} \\ [M(OH)^{n-1}] &= \begin{cases} B e^{-\beta(x-x_1)^2}, & x \leq x_1 \\ B & , x > x_1 \end{cases} \end{aligned} \quad (8)$$

with  $x_0 < x_1$  which precludes the difficulty above when  $\alpha = \beta$ . (Inherent in any such hypothesis is that no significant hydrolysis is postulated when  $x < x_0$  which only reflects the fact that a model is an idealization of reality.) Furthermore, since

$$\frac{[M^{+n}]}{[M(OH)^{n-1}]} = \frac{A}{B} \exp\{(\beta - \alpha)x^2 + 2(\alpha x_0 - \beta x_1)x + \beta x_1^2 - \alpha x_0^2\}, \quad x_0 < x < x_1,$$

one can argue that  $\alpha$  always equals  $\beta$  for then the quotient  $[M^{+n}]/[M(OH)^{n-1}]$  has the

correct behavior. Observe that the nonconstant parts of the expressions in (8) utilize the right half of a normal distribution with mean  $x_0$  and the left half of a normal distribution with mean  $x_1$ , respectively. Hence, the most elementary hypothesis for the dependence of  $[M^{+n}]$  and  $[M(OH)^{n-1}]$  upon pH is that of a normal distribution.

## 2. THE MODEL

The dependence of the metal ion  $M^{+n}$  upon pH, by hypothesis, is expressed as

$$[M^{+n}] = \begin{cases} C_{M^+} e^{-\alpha(\text{pH}-\text{pH}_0)^2}, & \text{pH} \geq \text{pH}_0 \\ C_{M^+} & , \text{pH} < \text{pH}_0 \end{cases} \quad (9)$$

where  $\text{pH}_0$  is the value at which  $[M^{+n}]$  is a maximum just before it begins to undergo significant hydrolysis. Furthermore,

$\text{pH}$  is any value of  $\text{pH} = -\log[\text{H}^+]$  on the pH axis;

$\alpha = 1/2\sigma^2$ , where  $\sigma$  is the standard deviation of the normal distribution;

$C_{M^+}$  is the total metal ion concentration in all states.

In particular, one assumes that  $[M^{+n}]$  is approximately constant up to the point  $\text{pH}_0$  which is the value of pH where the hydrolysis of  $M^{+n}$  is first considered significant. As previously discussed, the level of significance is fixed by establishing a predetermined value of the ratio  $[M^{+n}]/[M(OH)^{n-1}]$ , say  $10^6$ , as in Eq. (7). The behavior of  $[M^{+n}]$  is described by the right half of a normal distribution (Fig. 2) for larger values of pH, i.e., that part of the normal distribution to the right of the mean,  $\text{pH}_0$ .

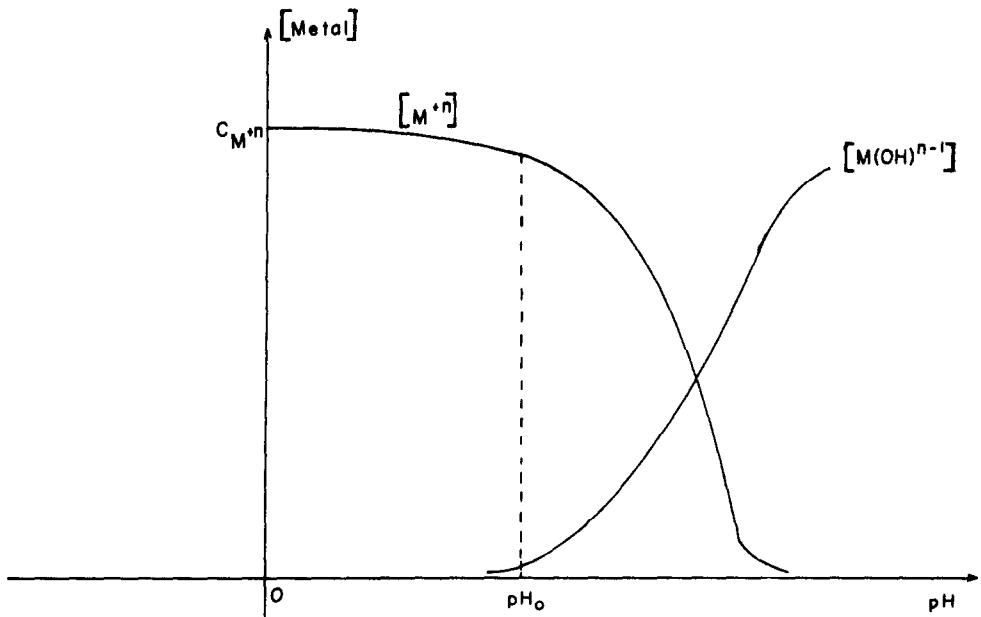


Fig. 2. The distribution curves for the concentrations of the ions  $M^{+n}$  and  $M(OH)^{n-1}$  as a function of pH.

It will be shown that the parameter  $\alpha$  and hence  $\sigma$  may be determined in terms of the equilibrium constants  $K_f$  and  $K_w$ . Therefore, one can find a confidence interval for pH where hydrolysis occurs corresponding to the predetermined significance level of the ratio  $[M^{+n}]/[M(OH)^{n-1}]$ . Note that  $\alpha$  measures the rate of decrease of the  $M^{+n}$  concentration; for large values of  $\alpha$ ,  $[M^{+n}]$  will be depleted in a small pH interval while for small values of  $\alpha$  the opposite is true.

Solving Eq. (9) for  $\alpha$  by taking natural logarithms and converting them to base 10 yields

$$\alpha = 2.3 \frac{\log C_{M^{+n}} - \log [M^{+n}]}{(\text{pH} - \text{pH}_0)^2} \quad (10)$$

or

$$\alpha = 2.3 \frac{\log C_{M^{+n}} - \log [M^{+n}]}{(-\log [H^+] + \log [H^+]_0)^2} \quad (10a)$$

There is also a normal distribution curve for the concentration of  $M(OH)^{n-1}$  and consequently a point where the two normal distribution curves intersect. At this point,  $[M^{+n}] = [M(OH)^{n-1}]$ , and therefore, from Eq. (3)

$$K_f = 1/[OH^-]. \quad (11)$$

Since  $K_w = [H^+][OH^-]$ , Eq. (11) may be rewritten as

$$[H^+] = K_f \cdot K_w. \quad (12)$$

It is important to remember that Eq. (12) represents the value of  $[H^+]$  at which the normal distribution curves for  $[M^{+n}]$  and  $[M(OH)^{n-1}]$  intersect.

The mass balance relationship

$$C_{M^{+n}} = [M^{+n}] + [M(OH)^{n-1}],$$

expresses that  $C_{M^{+n}}$  represents the total metal ion concentration in all states. The  $[M(OH)^{n-1}]$  term may be eliminated by substituting from Eq. (3),

$$C_{M^{+n}} = [M^{+n}] + K_f [M^{+n}][OH^-], \quad (13)$$

and then solving for  $[M^{+n}]$  and converting  $[OH^-]$  into  $[H^+]$  gives

$$[M^{+n}] = \frac{C_{M^{+n}}}{1 + (K_f \cdot K_w)/[H^+]}. \quad (14)$$

If we use Eq. (12) in (14), we find

$$[M^{+n}] = C_{M^{+n}}/2, \quad (15)$$

when  $[M^{+n}] = [M(OH)^{n-1}]$ . [Of course, Eq. (15) states what is obvious from the curves but an analytic proof of the assertion is desirable.] Therefore, Eqs. (12), (15), and (10a) yield

$$\alpha = 2.3 \frac{\log C_{M^{+n}} - \log(C_{M^{+n}}/2)}{(-\log K_f \cdot K_w + \log[H_0^+])^2} \quad (16)$$

$$= \frac{0.6924}{(-\log K_f \cdot K_w + \log[H_0^+])^2} \quad (16a)$$

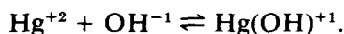
Finally,

$$\sigma^2 = \frac{1}{2\alpha} = \frac{(-\log K_f \cdot K_w + \log[H_0^+])^2}{1.3847} \quad (17)$$

or

$$\sigma = 0.8498 \left| \log \frac{[H_0^+]}{K_f \cdot K_w} \right| \quad (17a)$$

The value of  $pH_0$  which is the largest value of pH just prior to significant hydrolysis is found from Eq. (6) by assuming an upper bound for the ratio  $[M^{+n}]/[M(OH)^{n-1}]$ ; the corresponding  $\sigma$  value is then calculated using Eq. (17a). Results for various values of  $[M^{+n}]/[M(OH)^{n-1}]$  are shown in Table 1 for the equilibrium reaction [5]



The value of  $\sigma$  is important in that it gives a confidence interval for the pH range where hydrolysis occurs. As one is only interested in the right half of the distribution curve, a confidence interval of 95% will correspond to  $1.96\sigma$  which is obtained from standard statistical tables [4]. Therefore, the upper limit of the pH value which includes 95% of all the possible  $M^{+n}$  concentrations corresponding to the predetermined significance level of  $[M^{+n}]/[M(OH)^{n-1}]$  will be found by using the relationship  $pH_0 + 1.96\sigma$ .

From the tabulated results in Table 1, for a smaller ratio of  $[M^{+n}]/[M(OH)^{n-1}]$ , a lower value of  $pH_0$  will be found along with a larger value of  $\sigma$  which implies that the pH interval for statistically significant existence of  $[M^{+n}]$  will increase in length as the  $pH_0$  value decreases. Examination of Eq. (17a) indicates that  $\sigma$  varies directly with the logarithm term. For example, as  $K_f$  increases and  $[H_0^+]$  decreases,  $\sigma$  will decrease.

Table 1. Calculation of equilibrium parameters\*

$\frac{[M^{+n}]}{[M(OH)^{n-1}]}$	$pH_0^a$	$\sigma^b$	$pH_0 + 1.96\sigma^c$
$5 \times 10^3$	+0.001	0.626	1.228
$1 \times 10^4$	-0.299	2.70	4.993
$1 \times 10^5$	-1.299	3.24	5.051
$1 \times 10^6$	-2.299	3.45	4.463
$1 \times 10^7$	-3.299	3.58	3.7178

\* Data from equilibrium values for  $Hg^{+2} + OH^{-} \rightleftharpoons Hg(OH)^{+1}$ ,  $K_f = 1.995 \times 10^{10}$  [5].

<sup>a</sup> Calculated using Eq. (6).

<sup>b</sup> Calculated using Eq. (17a).

<sup>c</sup> Upper pH limit at 95% confidence level.

The traditional equilibrium expression, described by Eq. (3) or (6), is used to determine the concentrations of any one of the ions given the other two by simple substitution. The hypothesis (8) which incorporates the normal distribution provides an estimate of the pH range for the existence of the free metal ion  $M^{+n}$  in terms of the equilibrium constant  $K_f$ . The range for the existence of a particular metal ion may thus be determined within a specific limit of confidence, but not precisely, due to the asymptotic nature of the normal distribution.

In general, an identical argument applies for the equilibrium of complex formation described in Eq. (1). One hypothesizes that the concentration of the ions versus  $\log[L]$  has a form similar to (8) which yields the relationship

$$\sigma = 0.8498 | \log K_f \cdot [L_0] | ,$$

where  $[L_0]$  is the maximum concentration for the formation of the complex  $ML$ .

### REFERENCES

1. L. G. Sillen, P. W. Lange, and C. O. Gabrielson, *Problems in Physical Chemistry*, Prentice-Hall, New York (1952).
2. J. Guerra, *J. Chem. Educ.* **34**, 34 (1957).
3. L. G. Sillen, *Treatise on Analytical Chemistry*, Chap. 8, Wiley-Interscience, New York (1959).
4. N. M. Downie and R. W. Heath, *Basic Statistical Methods*, 3rd ed., Harper and Row, New York (1970).
5. J. Butler, *Ionic Equilibrium, A Mathematical Approach*, Addison-Wesley, Reading, Mass. (1964).