The Spectrophotometric Determination of Dissociation
Constants of Dibasic Acids. Methods Using a Minimum
Amount of Data*

by

Burton J. Thamer

UNIVERSITY OF CALIFORNIA

LOS ALAMOS SCIENTIFIC LABORATORY

LOS ALAMOS, NEW MEXICO

Abstract

Two spectrophotometric methods are given for determining dissociation constants of dibasic acids. Each method is applicable for any degree of overlapping of the dissociation constants. Neither method requires a direct measurement of the extinction coefficient of any of the individual species that are in equilibrium with each other. A minimum amount of data is required in either method.

281-2

The ionization equilibria of a light-absorbing acid frequently may be investigated by measurements of the optical density versus pH at a wave length such that the different ionic forms of the acid have markedly different absorption coefficients. The concentration of the acid under investigation and the length of the absorption cell are kept constant for this purpose. If, in addition, the temperature and ionic strength are maintained constant the equilibria of a dibasic acid H_2A may be represented by the equations

$$k_1 = \frac{a[HA]}{[H_2A]} \text{ and } k_2 = \frac{a[A]}{[HA]} \tag{1}$$

where $\underline{a}$ represents the measured hydrogen ion activity and brackets indicate concentrations.[1] The thermodynamic dissociation constants can be estimated from the measured dissociation constants k_1 and k_2 if the value of the ionic strength is fairly low. As previously indicated the optical density D after correction for absorption by the solvent and any buffer that may be present may be represented by

$$D = \frac{Lc[E_1 + \frac{k_1}{a} E_2 + \frac{k_1 k_2}{a^2} E_3]}{1 + \frac{k_1}{a} + \frac{k_1 k_2}{a^2}} \tag{2}$$

where L is the length of the absorption cell, c is the total concentration of the acid in all its forms, and E_1, E_2 and E_3 are the molar extinction coefficients of the forms H_2A, HA and A, respectively.[2]

If k_1 and k_2 are greatly different in magnitude they can be determined separately[3] since then only one stage of ionization is appreciably operative at a given pH. If the dissociation constants do not overlap too much, for example if $\frac{k_1}{k_2}$ is greater than 50, they could be determined separately without serious error

----

[1] Charges are not attached to the symbols H_2A, HA and A in the equilibria

$$H_2A \rightleftharpoons H^+ + HA$$

and

$$HA \rightleftharpoons H^+ + A$$
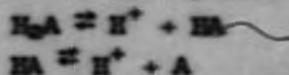
because the "dibasic acid" H_2A might be a conjugate acid of an ampholyte, for example.

[2] B. J. Thamer and A. F. Voigt, This Journal, 56, 225 (1952).

[3] D. H. Rosenblatt, ibid., 58, 40, equation (c) (1954).

by the method of "indirect colorimetry," which, however, requires equipment not ordinarily available in the laboratory.[4] The separate determination of k_1 and k_2 can be accomplished by a new spectrophotometric method by Rosenblatt but the method is subject to some inaccuracy if k_1 and k_2 are strongly overlapping.[5] A method by Thamer and Voigt is designed specifically for the case in which the dissociation constants do overlap and $\frac{k_1}{k_2}$ is less than 1000.[7] Their treatment requires that a plot of optical density versus pH have a maximum or a minimum in the region of the curve where the intermediate species HA predominates. The latter requirement can frequently be met, but one must devote at least a few points to the determination of the optical density and pH at the maximum (or minimum) of the curve. It is assumed in their method that the optical densities $D_1 = LcE_1$ and $D_3 = LcE_3$ can be measured directly by obtaining the pure forms H_2A and A, respectively. This may not be possible and D_1 or D_3 may have to be estimated from the available data at the lowest or highest pH, respectively.

Equation 2 is valid for each solution that is measured at the given wave length and under the experimental conditions mentioned in the first paragraph. Making the substitutions $LcE_1 = D_1$ and $LcE_3 = D_3$, and rearranging terms, equation 2 can be brought into the following form for each i-th designated solution of a series of solutions:

$$a_i D_i (k_1) + D_i (k_1 k_2) - (D_3 k_1 k_2) - a_i (LcE_2 k_1) - a_i^2 (D_1) = -a_i^2 D_i \quad (3)$$

In the general case of a dibasic acid each of the five quantities enclosed in parentheses in equation 3 is an unknown constant and all quantities outside of parentheses are measurable. As has been suggested by Rosenblatt the five unknowns could be determined by the simultaneous solution of five equations of the type of equation 3 thus requiring measurements at only five different values of pH.[6] Two methods of implementing this attack on the problem will be derived. Each of these methods (1) will be rigorously applicable for any value of $\frac{k_1}{k_2}$ that may be encountered, (2) will not require direct measurement of $D_1 = LcE_1$ for pure H_2A or $D_3 = LcE_3$ for pure A, and (5) will not require that a plot of optical density

----

[4] L. Sacconi, This Journal, 54, 829 (1950).
[5] Ibid., 58, 41, equation e (1954).
[6] Ibid., 58, 42 (1954).

281-4

- 2 -

against pH have a maximum or a minimum corresponding to the intermediate species HA. Furthermore a minimum amount of data will be required.

In either method the hydrogen ion activity m_i and a measure of the optical density D_i must be known for each of five solutions. The five measured solutions will be referred to as α, β, γ, δ and ϵ. The individual significance of each solution can best be described by reference to Fig. 1 which contains data previously reported by Thamer and Voigt for isophthalic acid.[2] (The fact that there is a maximum in the curve is immaterial for the present discussion.) The labeled solid points have been chosen for solutions α, β, γ, δ and ϵ. Solution α, in effect, serves for a determination of $D_1 = Lc\epsilon_1$ in equation 3 and, as such, should have as low a pH as possible in order to have a high concentration of H_2A.[7] Solution γ should be located at a pH near which a maximum concentration of the intermediate species HA exists in order for it to best fulfill its function of allowing an evaluation of the quantity $Lc\epsilon_2$ in equation 3. Solution ϵ corresponds to the determination of D_3 and it should have as much as possible of the species A present, i.e., be at a high pH. Solution β, in conjunction with points α and γ, is the most important point in the evaluation of k_1. For best accuracy its optical density should lie about midway between the optical densities of α and γ. Similarly solution δ, in conjunction with γ and ϵ, is the important point in evaluating k_2. Its optical density preferably should lie midway between the optical densities of γ and ϵ.

The first method of handling the problem might be referred to as a "complete" solution because it allows the direct calculation of k_1 and k_2 without making any simplifying assumptions regarding the values of D_1, $Lc\epsilon_2$ or D_3. In the derivation advantage may be taken of the fact that only <u>differences</u> in optical densities need to be employed. Thus the basic equation, equation 2, may be written in the form

----

[7] In the example given, a pH of approximately 1.6 was the lowest obtainable in order not to exceed an ionic strength of 0.030 that was maintained for each solution.

281-5

- 3 -

$$D_1 - D_{\hat{d}} = \frac{(D_1 - D_\alpha) + \frac{k_1}{a_1}(LcE_2 - D_\alpha) + \frac{k_1 k_2}{a_1^2}(D_3 - D_\alpha)}{1 + \frac{k_1}{a_1} + \frac{k_1 k_2}{a_1^2}}$$

or

$$a_1(D_1 - D_\alpha)[k_1] + (D_1 - D_\alpha)[k_1 k_2] - [k_1 k_2(D_3 - D_\alpha)] - a_1[k_1(LcE_2 - D_\alpha)] - a_1^2[D_1 - D_\alpha] =$$

$$-a_1^2(D_1 - D_\alpha)$$

where $D_1 = LcE_1$, $D_3 = LcE_3$ and i represents the i-th solution. When i denotes solution α, i.e., $i = \alpha$, the latter equation reduces to

$$-[k_1 k_2(D_3 - D_\alpha)] - a_\alpha[k_1(LcE_2 - D_\alpha)] - a_\alpha^2[D_1 - D_\alpha] = 0.$$

Upon combining the latter two equations one obtains

$$a_1(D_1 - D_\alpha)[k_1] + (D_1 - D_\alpha)[k_1 k_2] + (a_\alpha - a_1)[k_1(LcE_2 - D_\alpha)] + (a_\alpha^2 - a_1^2)[D_1 - D_\alpha] =$$

$$-a_1^2(D_1 - D_\alpha). \tag{4}$$

The four unknown quantities enclosed in brackets in equation 4 may be determined by the simultaneous solution of the four equations resulting when i equals successively β, γ, δ and ϵ. Solving for the first two unknown quantities one obtains

$$k_1 = \frac{\Delta_1}{\Delta}$$

and

$$k_1 k_2 = \frac{\Delta_2}{\Delta}$$

giving

$$k_2 = \frac{\Delta_2}{\Delta_1}$$

281-6

- 4 -

where

$$\Delta = \begin{vmatrix} a_\beta(D_\beta-D_\alpha) & (D_\beta-D_\alpha) & (a_\alpha-a_\beta) & (a_\alpha^2-a_\beta^2) \\ a_\gamma(D_\gamma-D_\alpha) & (D_\gamma-D_\alpha) & (a_\alpha-a_\gamma) & (a_\alpha^2-a_\gamma^2) \\ a_\delta(D_\delta-D_\alpha) & (D_\delta-D_\alpha) & (a_\alpha-a_\delta) & (a_\alpha^2-a_\delta^2) \\ a_\epsilon(D_\epsilon-D_\alpha) & (D_\epsilon-D_\alpha) & (a_\alpha-a_\epsilon) & (a_\alpha^2-a_\epsilon^2) \end{vmatrix}$$

$$\Delta_1 = \begin{vmatrix} -a_\beta^2(D_\beta-D_\alpha) & (D_\beta-D_\alpha) & (a_\alpha-a_\beta) & (a_\alpha^2-a_\beta^2) \\ -a_\gamma^2(D_\gamma-D_\alpha) & (D_\gamma-D_\alpha) & (a_\alpha-a_\gamma) & (a_\alpha^2-a_\gamma^2) \\ -a_\delta^2(D_\delta-D_\alpha) & (D_\delta-D_\alpha) & (a_\alpha-a_\delta) & (a_\alpha^2-a_\delta^2) \\ -a_\epsilon^2(D_\epsilon-D_\alpha) & (D_\epsilon-D_\alpha) & (a_\alpha-a_\epsilon) & (a_\alpha^2-a_\epsilon^2) \end{vmatrix}$$

$$\Delta_2 = \begin{vmatrix} a_\beta(D_\beta-D_\alpha) & -a_\beta^2(D_\beta-D_\alpha) & (a_\alpha-a_\beta) & (a_\alpha^2-a_\beta^2) \\ a_\gamma(D_\gamma-D_\alpha) & -a_\gamma^2(D_\gamma-D_\alpha) & (a_\alpha-a_\gamma) & (a_\alpha^2-a_\gamma^2) \\ a_\delta(D_\delta-D_\alpha) & -a_\delta^2(D_\delta-D_\alpha) & (a_\alpha-a_\delta) & (a_\alpha^2-a_\delta^2) \\ a_\epsilon(D_\epsilon-D_\alpha) & -a_\epsilon^2(D_\epsilon-D_\alpha) & (a_\alpha-a_\epsilon) & (a_\alpha^2-a_\epsilon^2) \end{vmatrix}$$

Upon simplification these equations give solutions for k_1 and k_2 in the form

$$k_1 = \frac{[-\lambda_1(a_\beta+a_\gamma)+\lambda_2(a_\beta+a_\delta)-\lambda_3(a_\beta+a_\epsilon)-\lambda_4(a_\gamma+a_\delta)+\lambda_5(a_\gamma+a_\epsilon)-\lambda_6(a_\delta+a_\epsilon)]}{[\lambda_1-\lambda_2+\lambda_3+\lambda_4-\lambda_5+\lambda_6]} \quad (5)$$

and

$$k_2 = \frac{[\lambda_1 a_\beta a_\gamma-\lambda_2 a_\beta a_\delta+\lambda_3 a_\beta a_\epsilon+\lambda_4 a_\gamma a_\delta-\lambda_5 a_\gamma a_\epsilon+\lambda_6 a_\delta a_\epsilon]}{[-\lambda_1(a_\beta+a_\gamma)+\lambda_2(a_\beta+a_\delta)-\lambda_3(a_\beta+a_\epsilon)-\lambda_4(a_\gamma+a_\delta)+\lambda_5(a_\gamma+a_\epsilon)-\lambda_6(a_\delta+a_\epsilon)]} \quad (6)$$

where

$$\lambda_1 = (D_\beta-D_\alpha)(D_\gamma-D_\alpha)(a_\alpha-a_\delta)(a_\alpha-a_\epsilon)(a_\beta-a_\gamma)(a_\delta-a_\epsilon)$$

$$\lambda_2 = (D_\beta-D_\alpha)(D_\delta-D_\alpha)(a_\alpha-a_\gamma)(a_\alpha-a_\epsilon)(a_\beta-a_\delta)(a_\gamma-a_\epsilon)$$

$$\lambda_3 = (D_\beta-D_\alpha)(D_\epsilon-D_\alpha)(a_\alpha-a_\gamma)(a_\alpha-a_\delta)(a_\beta-a_\epsilon)(a_\gamma-a_\delta)$$

$$\lambda_4 = (D_\gamma-D_\alpha)(D_\delta-D_\alpha)(a_\alpha-a_\beta)(a_\alpha-a_\epsilon)(a_\gamma-a_\delta)(a_\beta-a_\epsilon)$$

$$\lambda_5 = (D_\gamma-D_\alpha)(D_\epsilon-D_\alpha)(a_\alpha-a_\beta)(a_\alpha-a_\delta)(a_\gamma-a_\epsilon)(a_\beta-a_\delta)$$

$$\lambda_6 = (D_\delta-D_\alpha)(D_\epsilon-D_\alpha)(a_\alpha-a_\beta)(a_\alpha-a_\gamma)(a_\delta-a_\epsilon)(a_\beta-a_\gamma)$$

281-7

The points referred to in Fig. 1 have been tabulated in Table I. The values of k_1 and k_2 calculated

Table I

Spectrophotometric Data for Isophthalic Acid at 2460.0 $\overset{o}{A}$[2]

i	$(pH)_i$	a_i	D_i	$D_i - D_\alpha$
α	1.58	2.63×10^{-2}	1.298	0.000
β	3.12	7.59×10^{-4}	1.375	0.077
γ	3.75	1.78×10^{-4}	1.451	0.153
δ	4.55	2.82×10^{-5}	1.346	0.048
ϵ	7.90	1.26×10^{-8}	1.149	-0.149

from equations 5 and 6 are $k_1 = 2.51 \times 10^{-4}$ and $k_2 = 3.96 \times 10^{-5}$. These figures agree with $k_1 = (2.67 \pm 0.28) \times 10^{-4}$ and $k_2 = (4.00 \pm 0.42) \times 10^{-5}$ previously calculated by Thamer and Voigt from all of the data of Fig. 1.[2]

The second method involves successive approximations which converge rapidly to the final solution. In this method, as will be shown shortly, successive approximations are made of the value of $D_1 - LcE_1$ from the point α and of the value of $D_3 - LcE_3$ from the point ϵ. In each cycle of the successive approximations k_1 and k_2 are calculated as if the approximations to D_1 and D_3 were the actual values of D_1 and D_3. A rearrangement of equation 3 with D_1 and D_3 assumed to be known gives the equation

$$D_1 k_1 + \frac{1}{a_i}(D_1 - D_3)k_1 k_2 - LcE_2 k_1 = a_i(D_1 - D_i) \tag{3'}$$

where k_1, k_2 and LcE_2 are unknown quantities. Points β, γ and δ provide the necessary three equations of the form of equation 3'. The solution of the three equations gives

$$k_1 = \frac{\begin{vmatrix} a_\beta(D_1 - D_\beta) & \frac{1}{a_\beta}(D_\beta - D_3) & -1 \\ a_\gamma(D_1 - D_\gamma) & \frac{1}{a_\gamma}(D_\gamma - D_3) & -1 \\ a_\delta(D_1 - D_\delta) & \frac{1}{a_\delta}(D_\delta - D_3) & -1 \end{vmatrix}}{\begin{vmatrix} D_\beta & \frac{1}{a_\beta}(D_\beta - D_3) & -1 \\ D_\gamma & \frac{1}{a_\gamma}(D_\gamma - D_3) & -1 \\ D_\delta & \frac{1}{a_\delta}(D_\delta - D_3) & -1 \end{vmatrix}}$$

- 6 -

$$k_1 = \frac{\left[a_\beta(D_\beta-D_1)-a_\gamma(D_\gamma-D_1)\right]\left[\frac{1}{a_\delta}(D_\delta-D_3)-\frac{1}{a_\beta}(D_\beta-D_3)\right] - \left[a_\beta(D_\beta-D_1)-a_\delta(D_\delta-D_1)\right]\left[\frac{1}{a_\gamma}(D_\gamma-D_3)-\frac{1}{a_\beta}(D_\beta-D_3)\right]}{(D_\gamma-D_\beta)\left[\frac{1}{a_\delta}(D_\delta-D_3)-\frac{1}{a_\beta}(D_\beta-D_3)\right] - (D_\delta-D_\beta)\left[\frac{1}{a_\gamma}(D_\gamma-D_3)-\frac{1}{a_\beta}(D_\beta-D_3)\right]} \tag{7}$$

$$k_2 = \frac{\begin{vmatrix} D_\beta & a_\beta(D_1-D_\beta) & -1 \\ D_\gamma & a_\gamma(D_1-D_\gamma) & -1 \\ D_\delta & a_\delta(D_1-D_\delta) & -1 \end{vmatrix}}{\begin{vmatrix} a_\beta(D_1-D_\beta) & \frac{1}{a_\beta}(D_\beta-D_3) & -1 \\ a_\gamma(D_1-D_\gamma) & \frac{1}{a_\gamma}(D_\gamma-D_3) & -1 \\ a_\delta(D_1-D_\delta) & \frac{1}{a_\delta}(D_\delta-D_3) & -1 \end{vmatrix}}$$

$$k_2 = \frac{(D_\gamma-D_\beta)\left[a_\beta(D_\beta-D_1)-a_\delta(D_\delta-D_1)\right] - (D_\delta-D_\beta)\left[a_\beta(D_\beta-D_1)-a_\gamma(D_\gamma-D_1)\right]}{\left[\frac{1}{a_\delta}(D_\delta-D_3)-\frac{1}{a_\beta}(D_\beta-D_3)\right]\left[a_\beta(D_\beta-D_1)-a_\gamma(D_\gamma-D_1)\right] - \left[\frac{1}{a_\gamma}(D_\gamma-D_3)-\frac{1}{a_\beta}(D_\beta-D_3)\right]\left[a_\beta(D_\beta-D_1)-a_\delta(D_\delta-D_1)\right]} \tag{8}$$

Values of k_1 and k_2 are calculated from equations 7 and 8. These values are used in the following equations (9, 10 and 11) which are rearranged forms of equation 2. A value for LcE_2 is calculated from point γ and the equation

$$LcE_2 = D_\gamma + \frac{a_\gamma}{k_1}(D_\gamma-D_1) + \frac{k_2}{a_\gamma}(D_\gamma-D_3) \tag{9}$$

The values obtained thus far are then used in the following two equations to obtain better approximations to D_1 and D_3. Point α is used for a recalculation of the value of D_1:

$$D_1 = D_\alpha - \frac{k_1}{a_\alpha}(LcE_2-D_\alpha) + \frac{k_1 k_2}{a_\alpha^2}(D_\alpha-D_3) \tag{10}$$

Point ϵ is used for the recalculation of the value of D_3:

$$D_3 = D_\epsilon - \frac{a_\epsilon}{k_2}(LcE_2-D_\epsilon) + \frac{a_\epsilon^2}{k_1 k_2}(D_\epsilon-D_1) \tag{11}$$

281-9

- 7 -

The new values of D_1 and D_3 are then used in the next cycle of calculations using equations 7, 8, 9, 10 and 11 in succession. The calculations would be continued until the new values for D_1 and D_3 agreed within the limits of experimental error with their respective preceding values. For the very first cycle of calculation it might be assumed that $D_1 \doteq D_\alpha$ and $D_3 \doteq D_\epsilon$. The application of this method to the data of Table I is summarized in Table II.

Table II

The Calculation by Successive Approximations
of k_1 and k_2 for Isophthalic Acid from the Data of Table I.

	First Cycle	Second Cycle
Assumed value of D_1	1.298	1.295
Assumed value of D_3	1.149	1.149
Calculated value of k_1	2.54790×10^{-4}	2.50441×10^{-4}
Calculated value of k_2	4.06920×10^{-5}	3.97290×10^{-5}
Calculated value of LcE_2	1.656,052	1.629,281
Calculated value of $D_1(-LcE_1)$	$1.294,984 \doteq 1.295$	$1.294,848 \doteq 1.295$
Calculated value of $D_3 (-LcE_3)$	$1.148,849 \doteq 1.149$	$1.148,848 \doteq 1.149$

Since the respective calculated values of D_1 and D_3 in the first two cycles agree with each other within experimental error ($\pm$ 0.001 optical density units at best), the values of k_1 and k_2 calculated in the second cycle can be taken as final values. These values of $k_1 = 2.50 \times 10^{-4}$ and $k_2 = 3.97 \times 10^{-5}$ differ by only one in the third place from the values obtained by the "complete" solution. The difference is probably due to the rounding of the assumed values of D_1 and D_3 in Table II and, in any event, the difference is negligible compared with the actual experimental error. The values of the second cycle in Table II have been used in equation 2 to calculate the solid curve plotted in Fig. I. The agreement between the plotted curve and the experimental points provides some experimental support for the validity and usefulness of the two methods presented here.

Although the two methods required about the same amount of calculating in the example that was given, the "complete" method sometimes will be quicker because the points D_α and/or D_ϵ may not lie so close to the respective values of D_1 and/or D_3 and more than two cycles then might be required in the above method of successive approximations. Disadvantages of the "complete" method are the slightly greater complexity of the terms involved in it and the fact that

281-10

- 8 -

arithmetical errors might be harder to detect than they would be in the method of successive approximations. In the interests of accuracy it would be desirable in either method to determine each of the five points at least in duplicate. Also the measurement of a few trial points might be helpful in finding a favorable pH at which to measure solution γ. Only optical density differences such as D_1-D_α need to be measured. Consequently only four spectrophotometric quantities, e.g., $D_\beta-D_\alpha$, $D_\gamma-D_\alpha$, $D_\delta-D_\alpha$ and $D_\epsilon-D_\alpha$, require measurement and in measuring them increased accuracy could be obtained by using the technique of differential spectrophotometry.[a]

Each of the above methods provides a solution to the problems posed by overlapping values of k_1 and k_2 as does also the method of Thamer and Voigt.[2] A possible disadvantage of the methods presented here is the fact that the amount of data employed does not itself permit a verification that equations 1 and footnote 1 represent the actual equilibria of the system. Agreement between the plotted curve and additional points as in Fig. I would be necessary to show that such were the equilibria over the whole pH range. The advantage of the present methods lies in the fact that they would require little experimental work and little calculating. In the event that the optical densities were measured for each solution at more than one wave length they could be combined for calculating purposes into a composite C by means of the relation

$$C = \sum_j g_j D_j$$

where j represents each wave length at which measurements were made and each g_j is an arbitrarily chosen weighting factor.[2] Each resulting C-value could be used in place of the D-value at the same pH in each of the foregoing equations. The weighting factors would be chosen to give maximum values for

$$\frac{|C_\alpha-C_\gamma|}{\sum_j|g_j|} \qquad \text{and} \qquad \frac{|C_\gamma-C_\epsilon|}{\sum_j|g_j|}.$$

Instead of making the calculation at each wave length at which data were available and averaging the results, one could by this means utilize all of the data in a less time-consuming manner by making one calculation with composite values.

----
[a]See, for example, R. Bastian, R. Weberling and F. Palilla, Anal. Chem. 22, 160 (1950).
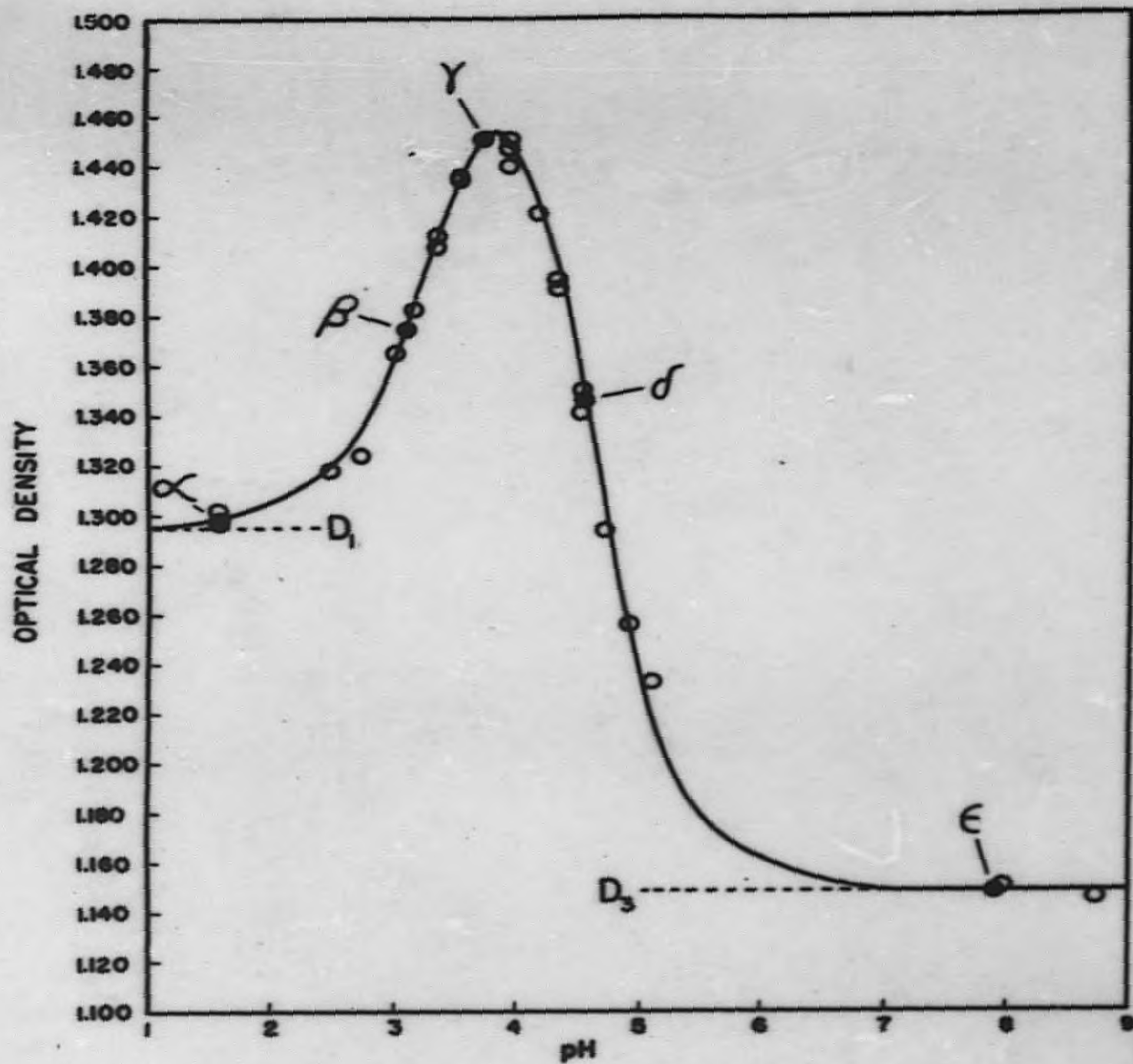
281-11

Fig. 1—Spectrophotometric data on isophthalic acid at 2460.0 Å: solid circles, experimental points used in the calculations of this paper; open circles, other experimental points; solid curve, calculated from Eq. 2 using the results of Table 2.

281-12