PHOTOGRAPHIC SCIENCE AND ENGINEERING
Volume 5, Number 1, January-February 1961

Description of *D*-Log *E* Curves by Specifically Chosen Parameters

B. E. BAYER, J. L. SIMONDS, AND F. C. WILLIAMS, *Research Laboratories, Eastman Kodak Company, Rochester, N.Y.*

Electronic computing machines offer to the sensitometrist not only speed, precision, and economy, but also new freedom in choosing parameters for describing sensitometric effects. Many properties of primary interest which are not easy to measure by graphical methods can be readily specified for digital analysis, and their amounts can be uniquely determined even though several properties vary simultaneously.

Photographic sensitometry has, since its beginning, relied heavily upon graphical methods. Density vs. log exposure curves are commonly plotted and drawn by hand, after which values of speed, contrast, and fog are measured from them. In recent years, however, high-speed digital computers have made it practical to compute these same values directly from the digital density data. A significant improvement in time, cost, and precision can result in large-scale operations where automatic data-recording is practical.

But if digital methods are to surpass graphical methods in factors other than time, cost, and precision, they must not be restricted to imitating graphical techniques. Because high-speed data-processing is a new tool with new capabilities, photographic sensitometry may profitably exploit it to do new things.

It is therefore sensible to reconsider thoroughly the problem of defining a set of abbreviated descriptions of the significant features of sensitometric curves in terms of what may be called sensitometric parameters. Most investigators may agree that, even through the choice of specific parameters will depend upon the immediate problem, the following three properties are quite general requirements.

1. The parameters should be few in number.
2. They should retain the significant information of the original density data.
3. They should convey this information in clear, convenient, and useful terms.

Photographic scientists have not been satisfied with the degree to which the traditional parameters of speed, contrast, and fog have satisfied the second requirement—that all the important information of the original *D*-log *E* data be retained in the parameter values. This requirement can be met only if one can construct the entire *D*-log *E* curve on the basis of the information contained in the parametric data, and generally this cannot be done with confidence.

One way to satisfy the requirement, in a way adaptable to automatic computation, is to postulate that a particular mathematical model will satisfactorily simulate each *D*-log *E* curve at hand, provided its parameters are properly chosen. As an illustration of this approach, consider the function:

$$y = \frac{1}{1 + e^{-x}}, \tag{1}$$

illustrated in Fig. 1. The curve of this model has an S-shape similar to that of a typical *D*-log *E* curve. If one now adds three parameters to form

$$y = \frac{k}{1 + e^{-a(x - x_0)}}, \tag{2}$$

one has the equation for the curve known in the field

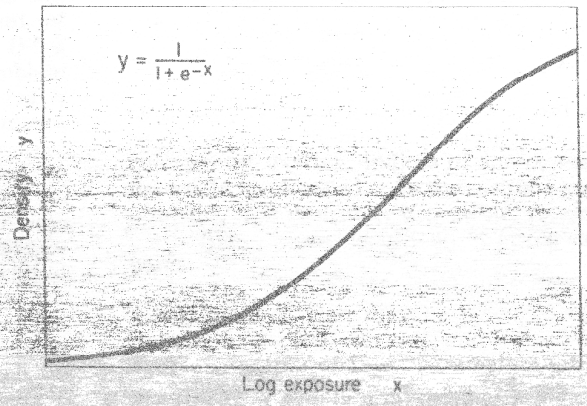

Fig. 1. Plot of the logistic curve of growth, showing its similarity in appearance to a *D*-log *E* curve.

Communication No. 2103 from the Kodak Research Laboratories, presented in part at the Annual Conference, Rochester, N.Y., 8 October 1958. Received 5 August 1960.

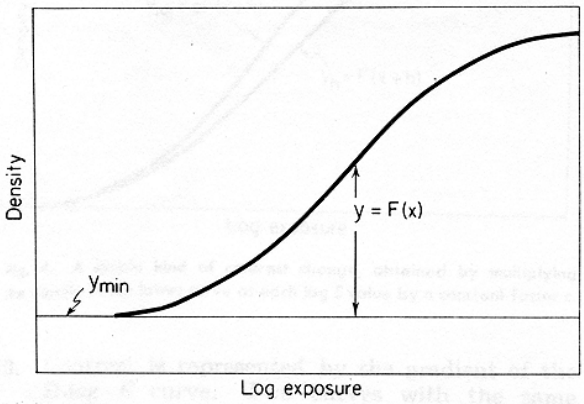

Fig. 2. *D*-log *E* curve for an experimental control. The reference curve, $y = F(x)$, is obtained by subtracting the minimum density, y_{min}.

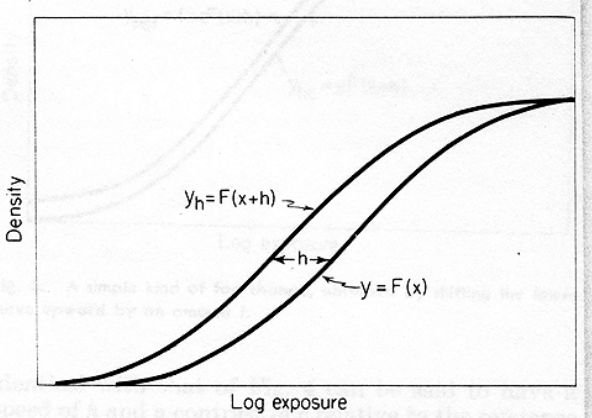

Fig. 3. A simple kind of speed change, obtained by shifting a reference curve, $y = F(x)$, to the left by a log *E* increment *h*.

of statistics as the logistic curve of growth. Now suppose that Eq. (2) be fitted to each of a group of *D*-log *E* curves, associating y with density above fog and x with log exposure. If the fit is good, then different values of k, a, and x_0 will correspond to different *D*-log *E* curves; and, since one can construct the original curves from knowing only the values of k, a, and x_0, these parameter values contain the same basic information as the original set of densities. If these parameters are also convenient and useful as sensitometric indices, then they comprise a set which satisfies the requirements listed before. Unfortunately, however, simple empirical models such as this one often fail to fit *D*-log *E* curves met in practice. Or even if a good fit is found, the parameters of the well-fitting model may not measure directly the kinds of changes in which the investigator is primarily interested.

It should be emphasized that many models can be found which, by simultaneous adjustment of three or four parameters, will satisfactorily fit *D*-log *E* curves encountered in practice. The real problem is to find a model which will not only fit *D*-log *E* curves, but also describe the curves in simple meaningful terms, such as speed, contrast, or fog.

But even as computers are not restricted to imitating graphical methods, neither are they restricted to standard functional forms. Whereas reason dictates that the *results* of an analysis be simple and concise, there is no requirement that the actual data-handling be so. Today, automatic computers can perform complicated numerical operations repeatedly and rapidly. This facility allows the experimenter broad limits within which to choose practical definitions for a set of sensitometric parameters.

This paper describes a simple alternative to using a prescribed mathematical function as a model for a *D*-log *E* curve. By a device which will be explained, the method of simulating a *D*-log *E* curve by a mathematical equation is changed from a specialized technique to a flexible way of exploiting the capabilities of data processors to achieve an analysis specifically tailored to meet specific requirements.

An Illustration of the Method

Let us suppose that sensitometric tests have been made of a large number of experimental films and that their speed, contrast, and fog values are to be compared with those of a control film. Each experimental film is represented by a test strip from which are made measurements of densities, z_i, resulting from a series of exposures, x_i. The data for such a strip might be tabulated as follows:

Log Exposure	Density
x_1	z_1
x_2	z_2
x_3	z_3
.	.
.	.
.	.
x_k	z_k

Ordinarily, a *D*-log *E* curve for each test strip would be drawn; and by means of straight-edge, or special transparent overlay, indices of speed, contrast, and fog would be obtained. In the case of a great number of strips, it is desirable, if possible, to employ automatic data-processing instead of graphical measurement.

For the purpose of automatic data-processing, a new point of view may be taken in regard to the definition of speed, contrast, and fog. The following simple concepts of these parameters are suggested:

1. Speed is simply a measure of required exposure level. Two curves alike in shape but displaced along the log-exposure axis differ in speed, and only in speed.
2. Fog is an underlying density upon which the useful part of the *D*-log *E* relationship is superimposed. Two curves alike except for a small displacement along the density axis differ in fog, and only in fog.

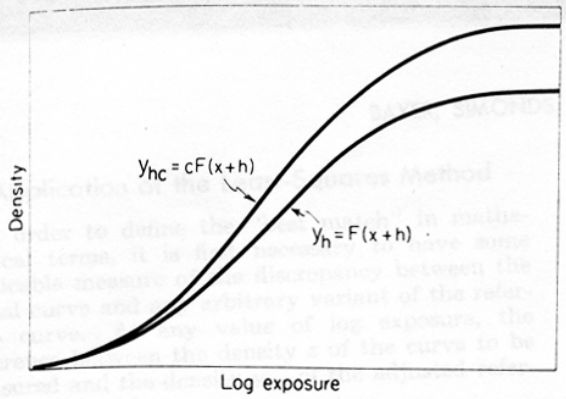

Fig. 4. A simple kind of contrast change, obtained by multiplying the density of the lower curve at each log E value by a constant factor c.

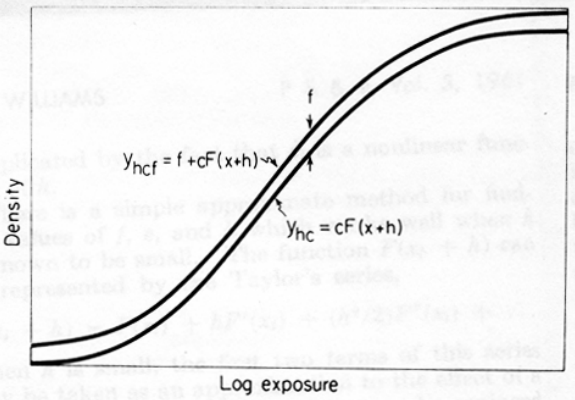

Fig. 5. A simple kind of fog change, obtained by shifting the lower curve upward by an amount f.

3. Contrast is represented by the gradient of the D-log E curve. Two curves with the same minimum density, in which the gradient of one at each value of log exposure is a constant multiple of the gradient of the other at the same log exposure, differ in contrast, and only in contrast.

In order to provide a common reference for the parameter values, each of the experimental films may be compared with an experimental standard or control, whose D-log E curve is given in Fig. 2. The densities of this film may be represented by the equation:

$$\text{Density} = y_{\min} + F(x), \tag{3}$$

where $y_{\min}$ is the density of an unexposed, processed portion of the film and x is log exposure. In the remaining discussions, reference will be made to $y = F(x)$ as the "reference curve."

Effecting the three simple speed, contrast, and fog changes just described in the reference curve will produce an infinite variety of curves. Within this domain of possible curve shapes, a reasonable match can be expected for each of the curves of the experimental films.

Shifting the reference curve, $y = F(x)$ from Fig. 2, to the left by an amount h will produce the curve of Fig. 3. It has a density, at log exposure x, of

$$y_h = F(x + h). \tag{4}$$

A curve that is like the one of Fig. 3 can be said to have a speed of h relative to the reference curve.

In a similar way, multiplying the curve of Fig. 3 by a constant multiplier, c, will yield a curve which differs from the reference curve not only in speed but also in contrast. This treatment yields the curve of Fig. 4, which has the equation

$$y_{hc} = cF(x + h). \tag{5}$$

At each value of log exposure, the gradient of the curve of Fig. 4 is a constant multiplier, c, of the gradient of the curve of Fig. 3. A curve nearly identical with that of Fig. 4 can be said to have a speed of h and a contrast of c relative to the reference curve.

Adding to the curve of Fig. 4 a constant density increment, f, will yield the curve

$$y_{hcf} = f + cF(x + h), \tag{6}$$

shown in Fig. 5. A curve nearly identical with Fig. 5 can be said to have a speed of h, a contrast of c, and a fog of f relative to the curve of Fig. 4.

Now let us consider the more realistic case in which a particular curve is to be compared with the reference curve. Of the infinite variety of curves which may be produced by varying f, c, and h in Eq. (4), a unique set of variations must be found which will produce an optimum match with the sample curve to be evaluated. If this can be done, the corresponding best values of f, c, and h can be used as measures of fog, contrast, and speed.

These indices have the following properties as sensitometric parameters:

1. They are few in number.
2. Provided only that the variant of the reference curve closely simulates the sample curve to be measured, the values of h, c, and f contain as much basic information as the original densities of the sample.
3. The parameters are defined according to the stated concepts of speed, contrast, and fog, and therefore convey their information in clear, convenient, and useful terms.

The foregoing defines a set of sensitometric indices. There are three problems to solve before the system described can be employed as a working method. (1) some reasonable criterion for the best match, in mathematical form, is required; (2) a numerical method for finding the best values of f, c, and h by automatic computation is needed; and (3) some way to treat the case in which the actual and fitted curves differ excessively must be found.

An Application of the Least-Squares Method

In order to define the "best match" in mathematical terms, it is first necessary to have some applicable measure of the discrepancy between the actual curve and any arbitrary variant of the reference curve. At any value of log exposure, the difference between the density z of the curve to be measured and the density y_{hcf} of the adjusted reference curve will be

$$z - y_{hcf} = z - [f + cF(x + h)]. \quad (7)$$

Densities $z_1, z_2, \ldots z_k$ are measured at a finite number of values of log exposure $x_1, x_2, \ldots x_k$. It is clear that any measure of the discrepancy between the two curves will be some function of the differences:

$$\begin{aligned} e_1 &= z_1 - f - cF(x_1 + h) \\ e_2 &= z_2 - f - cF(x_2 + h) \\ &\cdots\cdots \\ e_k &= z_k - f - cF(x_k + h). \end{aligned} \quad (8)$$

Any one of several possible functions might be used. For example, one might define as the best set of values, h, c, and f, that which makes as small as possible the largest of $e_1, e_2, \ldots, e_k$. This criterion is simple in principle. Unfortunately, it can be awkward in the face of random experimental errors in the data, and difficult to program for automatic computation.

The criterion of least squares fits the chosen mathematical model to the empirical D-log E curve with what is usually an acceptable degree of approximation; it affords a useful test of goodness of fit; and, fortunately, it is simple to apply. This criterion employs as a measure of the discrepancy between the actual and the fitted curve the sum of squares of the e_i.

$$SS = \sum_{i=1}^{k} e_i^2. \quad (9)$$

The best match is that for which SS is a minimum. This requirement that SS be made as small as possible leads to a logical treatment of data which contain random measurement errors. The fits obtained usually appear reasonable. Furthermore, the method is easily adapted to digital methods.

Numerical Methods for Computing Parameter Values

The problem is to fit the equations,

$$\hat{z}_i = f + cF(x_i + h) \text{ and } i = 1, 2, \ldots, k, \quad (10)$$

to a set of densities $z_1, z_2, \ldots, z_k$, choosing f, c, and h so that the sum of squares of residuals,

$$SS = \sum_{i=1}^{k} (z_i - \hat{z}_i)^2, \quad (11)$$

is made as small as possible. In Eqs. (10) and (11), $\hat{z}_i$ is the value of y_{hcf} at x_i. The general solution is complicated by the fact that $\hat{z}_i$ is a nonlinear function of h.

There is a simple approximate method for finding values of f, c, and h which works well when h is known to be small. The function $F(x_i + h)$ can be represented by the Taylor's series,

$$F(x_i + h) = F(x_i) + hF'(x_i) + (h^2/2)F''(x_i) + \ldots$$

When h is small, the first two terms of this series may be taken as an approximation to the effect of a horizontal shift. Thus, Eq. (10) may be replaced by the approximation,

$$\hat{z} \cong f + cF(x_i) + chF'(x_i). \quad (12)$$

Here, $F(x_i)$ is simply the value of the reference curve at log exposure x_i, and $F'(x_i)$ is the corresponding gradient.

Equation (12) is an ordinary regression equation, linear in f, c, and the product, ch. Methods for fitting such equations are well documented. Best values of f, c, and ch are given by simple linear equations of the form,

$$\begin{aligned} f &= w_{f1}z_1 + w_{f2}z_2 + \ldots\ldots + w_{fk}z_k; \\ c &= w_{c1}z_1 + w_{c2}z_2 + \ldots\ldots + w_{ck}z_k; \\ ch &= w_{ch1}z_1 + w_{ch2}z_2 + \ldots\ldots + w_{chk}z_k. \end{aligned} \quad (13)$$

The coefficients w_{ji} depend only upon the densities of the reference curve and on the particular set of values, $x_1, x_2, \ldots x_k$ chosen. Equations (13) are therefore simple to use to evaluate any number of sets of data relative to the same reference curve, for which the multipliers $w_{f2} \ldots w_{chk}$ can be obtained in an advance calculation. They are easy to program for a digital computer because they involve only multiplication and addition.

When h is large, Eq. (12) may be found inadequate. One must then consider a solution which does not depend on such a simple approximation as Eq. (12). A description of one method of analysis in this more complex situation is given in the Appendix to this paper.

Degree of Fit

The value of the system just defined depends partly on whether the fitted equation provides a satisfactory approximation to the measured densities it represents. The sum of squares SS is an indicator of the degree of fit, inasmuch as no deviation can exceed the square root of SS, and no deviation will usually exceed, say, one half of this value. Of course, a look at the individual differences $(z_i - \hat{z}_i)$ is better than a look at their sum of squares.

Experience with the equation,

$$y_{hcf} = f + cF(x + h), \quad (14)$$

has shown it to be adequate for most curves of black-and-white negative materials, for which it is common practice not to measure densities in the region of maximum density. When the set of parameters chosen does not provide a good fit, then the range

of densities over which the fit is made may be important. Thus, when shoulder-region densities are included in the data, a different treatment may be desirable. When the fit is poorer than desired, the solution has usually been the addition of one more parameter. The nature of this additional parameter, when not suggested by *a priori* knowledge of the group of curves, is best determined from an examination of the discrepancy for a sample of curves. Either a subjective appraisal or a numerical analysis of the density deviations will usually disclose a common pattern in the discrepancies, and on this the definition of a new parameter may be based.

It is possible to choose a greater number of parameters than can be accurately estimated. If three parameters will provide a good fit, then the addition of a fourth parameter may improve the fit slightly, but will probably lead to poorer estimation of the individual parameter values.

The Method in General

The foregoing is intended as an example of a general method which has these features:

1. The investigator chooses an actual *D*-log *E* relationship $F(x)$ against which he wants to compare the density data from each of his test strips. His choice might be an experimental standard or control.
2. The investigator then chooses and names a small number of ways in which this reference curve may be altered to produce other curves which he can easily visualize. He associates the amount of these changes with parameter values $b_1, b_2, \ldots b_n$.
3. After choosing in this way a small number of sensitometric parameters $b_1, b_2, \ldots\ldots b_n$, he then employs digital methods to find that combination of simultaneous changes in the reference curve which produces a curve $f(x; b_1, b_2, \ldots\ldots b_n)$ in best agreement with each set of density data he wishes to evaluate.
4. He finally interprets the amounts of $b_1, b_2, \ldots b_n$ as a specification, in his own terms, of the density–exposure relationship of each test strip.

The parameter definitions may be tailored to the needs of the individual experimenter.

APPENDIX. A More Accurate Method for Computing *h*, *c*, and *f*

In the section on numerical methods for computing parameter values, we indicated a simple approximate method for fitting the equation,

$$y_{hcf} = f + cF(x + h), \tag{1a}$$

to a set of measured densities. That method was said to yield good approximations to the least-squares values of h, c, and f when h is small. This Appendix describes a method which can be used to yield accurate values within a wide range of the parameter h. The method is easily adapted to a computing machine, such as the IBM 705 Data Processor.

In deriving this method, the following notation will be used:

1. The measured densities of the sample will be called: $z_1, z_2, \ldots, z_k$.
2. The corresponding log exposure values will be called: $x_1, x_2, \ldots, x_k$.
3. Substituting these k values of log exposure into Eq. (1a) will yield k values of y_{hcf}; these are functions of h, c, and f and will be called: $\hat{z}_1, \hat{z}_2, \ldots, \hat{z}_k$.
4. Similarly substituting the k values of log exposure into $F(x + h)$ will yield k functions of h; these will be called: $u_1, u_2, \ldots, u_k$.
5. The first partial derivatives of $F(x + h)$ with respect to h are also functions of h; these will be called: $u_1', u_2', \ldots, u_k'$.

Using the above notation, we can write Eq. (1a) for each value of log exposure, as follows:

$$\begin{aligned} \hat{z}_1 &= f + cF(x_1 + h) = f + cu_1. \\ \hat{z}_2 &= f + cF(x_2 + h) = f + cu_2. \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \hat{z}_k &= f + cF(x_k + h) = f + cu_k. \end{aligned} \tag{2a}$$

The problem is to find those values of h, c, and f which minimize.

$$SS = \sum_{i=1}^{k} (z_i - \hat{z}_i)^2 = \sum_{i=1}^{k} (z_i - f - cu_i)^2. \tag{3a}$$

Given a set of densities, $z_1, z_2, \ldots, z_k$, treat SS as a function of h, c, and f and find its minimum. At the minimum, the first partial derivatives of SS with respect to h, c, and f are simultaneously zero. This fact permits writing three simultaneous equations which the least-squares values of h, c, and f must satisfy:

$$\partial SS/\partial f = -2\sum_i (z_i - f - cu_i) = 0 \tag{4a}$$

$$\partial SS/\partial c = -2\sum_i (z_i - f - cu_i)(u_i) = 0 \tag{5a}$$

$$\partial SS/\partial h = -2c\sum_i (z_i - f - cu_i)(u_i') = 0 \tag{6a}$$

The parameter h enters into these equations in a complex way, but the first two equations can be solved for f and c in terms of h, and these values of f and c can be substituted into the third equation. This results in an equation for f in terms of h, an equation for c in terms of h, and a third equation, not involving f or c, which h must satisfy:

$$f = \Sigma z_i W_i, \tag{7a}$$

$$c = \Sigma z_i X_i, \tag{8a}$$

$$O = \Sigma z_i Y_i, \tag{9a}$$

where, in (7a), (8a), and (9a), the quantities W_i, X_i, and Y_i are defined by the following series of relationships:

$$
\begin{aligned}
u_i &= F(x_i + h) & u_i' &= \partial F(x_i + h)/\partial h \\
\bar{u} &= (1/k)\ \Sigma u_i & \bar{u}' &= (1/k)\ \Sigma u_i' \\
v_i &= u_i - \bar{u} & v_i' &= u_i' - \bar{u}' \\
U &= \Sigma v_i^2 & U' &= \Sigma v_i v_i' \\
X_i &= (v_i/U) & Y_i &= v_i' - (U'/U)v_i \\
W_i &= (1/k) - (\bar{u}X_i) & & \qquad (10a)
\end{aligned}
$$

Equations (7a), (8a), and (9a) are the basis for finding the least-squares values of f, c, and h. The quantities W_i, X_i, and Y_i for $i = 1, 2, \ldots, k$ are each functions of h, which do not vary with the measured densities z_i or the values of the parameters c and h. The least-squares value of h must satisfy Eq. (9a). When substituted into Eqs. (7a) and (8a), it will yield the least-squares values of f and c.

Thus far, the necessary relationships have been established which the parameters f, c, and h must bear to the measured densities $z_1, z_2, \ldots, z_k$, but a practical method for their use in an automatic computer has not been indicated. In particular, it has not been indicated how the functions of h represented by W, X, and Y may be simulated for digital computations. A practical way is to approximate each of these functions of h by a polynomial of degree r:

$$W_i = \alpha_{i0} + \alpha_{i1}h + \alpha_{i2}h^2 + \ldots + \alpha_{ir}h^r \qquad (11a)$$

$$X_i = \beta_{i0} + \beta_{i1}h + \beta_{i2}h^2 + \ldots + \beta_{ir}h^r \qquad (12a)$$

$$Y_i = \gamma_{i0} + \gamma_{i1}h + \gamma_{i2}h^2 + \ldots + \gamma_{ir}h^r \qquad (13a)$$

Eqs. (7a), (8a), and (9a) can be written in a form which involves only multiplication and addition:

$$f = A_0 + A_1h + A_2h^2 + \ldots + A_rh^r \qquad (14a)$$

$$c = B_0 + B_1h + B_2h^2 + \ldots + B_rh^r \qquad (15a)$$

$$O = C_0 + C_1h + C_2h^2 + \ldots + C_rh^r \qquad (16a)$$

where

$$A_j = \sum_i \alpha_{ij}z_i \qquad i = 1, 2, \ldots, r \qquad (17a)$$

$$B_j = \sum_i \beta_{ij}z_i \qquad (18a)$$

$$C_j = \sum_i \gamma_{ij}z_i \qquad (19a)$$

The values of α_{ij}, β_{ij}, and γ_{ij} may be chosen as constants for a particular reference curve and set of log E values $x_1, x_2, \ldots, x_k$. A practical way of choosing these values is to (1) select several values of h in the region of interest; (2) obtain corresponding values of u_i and u_i' from measurements of the reference curve; (3) evaluate the W_i, X_i, and Y_i using Eqs. (10a); and (4) fit Eqs. (11a), (12a), and (13a) by ordinary curve-fitting methods. This work need be done only once for each reference curve.

Least-squares values of f, c, and h for any number of sets of measured densities $z_1, z_2, \ldots, z_k$ may then be obtained by the following procedure:

1. Given $z_1, z_2, \ldots, z_k$, values of A_j, B_j, and C_j for $j = 1, 2, \ldots, r$ are obtained from Eqs. (17a), (18a), and (19a).
2. Using an appropriate root-solving method, a value of h is found to satisfy the polynomial (16a).
3. Values of f and c are obtained by substituting this least-squares value of h into Eqs. (14a) and (15a).

This method of evaluating f, c, and h has these advantages over other possible methods:

1. The equations are easily programmed for an automatic computer.
2. The only approximations involved are in simulating W, X, and Y by polynomials. These polynomials may be taken of sufficient degree to assure adequate accuracy over the range of interest.
3. Except for one calculation of the root of a polynomial, computations are direct and involve only multiplication and addition.