JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematics and Mathematical Physics
Vol. 69B, No. 3, July–September 1965

A Development of the Theory of Errors With Reference to Economy of Time[1]

Mayo Dyer Hersey [2]

Prefatory Note
by
Churchill Eisenhart

This article, a 1914 revision of a 1913 talk, deals with the scientific management of scientific work. Specifically, it deals with statistical theory applied to experimentation; and eloquently illustrates the main aim of the application of statistical theory today: to determine the allocation of skills, effort, time, and money that will achieve desired overall precision and accuracy in experimentation at lowest possible cost under existing limitations of equipment, materials, and personnel. Until now this article has remained unpublished, at first because of the disruptions of World War I, and later because the author felt that most of his colleagues were familiar with its content from his various oral presentations. It is published here at this time, not only for its historical interest, but also in the hope that its timeless instructions on the efficient planning and execution of measurement programs may be useful to experimenters in many parts of the world.

An abstract of the original paper is given in the BAAS *Report*, Birmingham, 1913:399–400; but does not contain the essential formulae, nor, of course, the details of their illustrative application to particular problems. The formulae of the present paper, and the illustrative examples of their application, were discussed by Mr. Hersey in a lecture on "The Theory of Errors of Physical Measurements" at Harvard University on December 6, 1915; and again, more fully, in a series of conferences that he conducted at the National Bureau of Standards in July 1920, as part of the work of the Aeronautic Instruments Section, of which he was Chief. The formulae of the present paper were also included in a list of fourteen theorems relative to the errors of physical measurements presented by Mr. Hersey in an Informal Communication at the 779th meeting of the Philosophical Society of Washington (D.C.), held at the Cosmos Club on November 25, 1916. The summary of this Informal Communication (*Journal of the Washington Academy of Sciences* **7**:23 (1917)) contains the remark, "The manuscript notes, such as they are, are available to any individual interested." This remark led me, early in May 1965, to address a letter of inquiry to Mr. Hersey, with whom I had been corresponding on the history of the formula given as equation (10) in the present paper. The author responded by forwarding a photocopy of his 1914 typescript and of the long abstract that he had sent to the British Association in advance of the 1913 meeting.

The present paper strictly follows the 1914 revision, except for the first two paragraphs, which were the opening paragraphs of the 1913 long abstract. The *Abstract* that precedes the text of the paper consists of a rearrangement (and slight abridgment) of the abstract published in the 1913 BAAS *Report*, to conform to the organization of the present paper, plus an opening sentence taken from the body of the 1913 long abstract. I have added the section numbers and section headings, in keeping with the present style of this Journal; and also a few additional footnotes, which are identified by my initials.

Abstract

Methods of economizing time can be considered with reference to the design and disposition of apparatus; or with reference to the experimental observations; or with reference to the computation of the result.

In connection with the problem of *designing* (or adjusting) apparatus so as to secure the most favorable result in a limited time, a criterion for "best magnitudes," previously proposed, is here further considered, and illustrated by an application to the interferometer.

Investigation of economy of time in taking the *observations* themselves leads to two distinct problems: first, that of the division of time amongst the components of an indirect measurement; second, that of the best grouping of observations in determining any one quantity.

[1] Revision, dated June 22, 1914, of a paper presented at a meeting of Section A (Mathematics and Physics) of the British Association for the Advancement of Science, in Birmingham, England, September 1913, and not published heretofore.

[2] Physicist, National Bureau of Standards, 1910–1922, 1926–1931; now Visiting Professor of Engineering (Research), Brown University, Providence, Rhode Island, 02912.

The solution of the first problem comes out in terms of three data—namely, the relative precision of, and the relative time consumed in, a single observation on the respective components; together with the derivatives expressing the sensitiveness of the result with respect to the several components. Of these data the first two are postulated, while the third is implicitly contained in the equation which defines the measurement in question. The solution is independent of the existence of constant errors.

The second problem consists in establishing the most profitable compromise between the extremes of (1) repeating a large number of readings under the same conditions (or on the same sample), in order to diminish the effect of *observational* errors; or (2) resting content with a lower precision on each determination, in order to cut down *systematic* errors by making numerous independent determinations (or by trying many different samples). The most economical number of observations to make in any one group before stopping to change conditions (or to set up a new sample) in preparation for a new group is directly expressible in terms of two postulated data: first, the ratio of the average observational error to the average systematic error anticipated; and, second, the ratio of the time required in preparing for a new group to the time used in a single observation. This result is independent of the total time available.

The first problem is illustrated by the division of time in a gravity determination by Kater's pendulum; the second, by the determination of the heat of combustion of coal from a series of samples. A combination of the two problems may also arise. The solution is equally straightforward.

Finally, in regard to *computation*, the availability of an automatic device for linear least-squares adjustment makes it now desirable to have some means of throwing an assumed relation into linear form without disturbing the relative weights of the observations. A general formula for doing this is here proposed, and applied to the determination of thermal expansion coefficients.

Throughout, the object of the paper is to establish certain *general* principles governing the accuracy attainable in physical measurements, independently of the *particular* apparatus or process in question.

1. Introduction

Industrial laboratories, which have always been subject to the dictates of economy of time, have recently been giving greater attention to questions of accuracy. On the other hand, research laboratories such as the Bureau of Standards which have always been subject to the dictates of accuracy, are nowadays forced to give continually greater attention to questions of economy of time. These two facts unite in lending interest to any systematic investigation of methods for attaining a given accuracy in the least time.

Particular expedients for saving time will continue to be discovered by individual investigators in mechanical, thermal, electrical, and optical measurements, chemical analysis, agricultural experiments and other problems separately considered. A different avenue of approach is contemplated in the present paper however, the object of which is to set forth once and for all some of the *general* principles governing the accuracy attainable in physical measurements, independently of the *particular* apparatus or process in question.

The rules gathered together in this paper are intended to serve as a guide in standardizing the routine of technical physical measurements. They were developed in the belief that it is not illogical to apply "scientific management" to scientific work, and that in analyzing questions of accuracy which arise in physical measurements, certain refinements borrowed from the astronomer or geodesist may profitably be combined with something of the spirit of the efficiency engineer.

We shall primarily be concerned with problems in maxima and minima, arising from the necessity for a compromise between precision and time-economy.

For the most part, our reasoning consists simply in unfolding the consequences of introducing this new variable *time* into the commonly accepted formulas of the theory of errors. The resulting equations deal, of course, with idealized or *limiting* cases. They are not intended to supplant the use of personal judgment, but on the contrary to afford it better tools with which to work. And physicists need hardly be reminded that insofar as such mathematical criteria are regarded as tools, they are to be regarded as keen-edged tools and used with discretion.

The practical usefulness of accuracy analysis hinges on clearly distinguishing between accuracy, precision, and reliability. The sense in which these terms are here used may be fixed in mind by the formula

$$A = R \pm P \tag{1}$$

in which A denotes the accuracy (i.e., departure from the truth) of a measurement whose reliability (i.e., constant, systematic, or *concealed* error) is R, and whose precision (i.e., accidental error, observational error, or *deviation*) is P. Of course, R and P are essentially different, for while P is governed by the laws of chance, R depends on the laws of physics.

Our attention will now be directed to questions of precision alone. It is possible by means of eq (1) to explicitly take account of concealed errors, but that will not be done in this paper.

In addition to this limitation, three assumptions which underlie the following reasoning may be stated at the start.

First, we assume that the precision measure of the mean of a series of observations is to be inversely proportional to the square root of the number of observations. Thus if p be the precision of each of n observations, the precision of their mean becomes

$$P = p/\sqrt{n} \cdot \tag{2}$$

Strictly the root-mean-square-deviation, but, practically, the average deviation, may be taken as the precision measure in any case.

Second, we assume that the resulting error in any quantity due to the simultaneous existence of indifferently + or − errors E_1, E_2 . . . will be

$$E=\sqrt{E_1^2+E_2^2+\ldots} \tag{3}$$

Third, we assume that if on an average each observation requires the time t, the total time consumed in a series of n observations will be

$$\theta=nt. \tag{4}$$

Let us now consider some of the questions which arise in the work of the *designer*, the *observer*, and the *computer* respectively.

2. Apparatus Design to Minimize Measurement Time Required for Results of Prescribed Precision

The problem of *designing* apparatus so as to secure the most favorable result in a limited time involves the choice of "best magnitudes" for the components x . . . in terms of which, and by means of some relation

$$y=f(x\ldots) \tag{5}$$

the quantity y under investigation is defined. If in general the function for which we seek a minimum is $\psi\Delta y$; if the experimental error Δx varies throughout the available range of x in proportion to some function φ; and if F be written for f' while primes denote first derivatives by x; then the best magnitudes are found to be in accordance with a certain criterion which, in the case of a single component, x, reduces to the equation

$$\frac{\psi'}{\psi}+\frac{\varphi'}{\varphi}+\frac{F'}{F}=0. \tag{6}$$

This criterion for best magnitudes was presented in a more general form in an earlier paper,[3] but has not previously been examined with reference to economy of time. If we inquire how many times longer it would take to secure a prescribed precision in y, if x were *not* adjusted to its best magnitude x_0, we find that, calling θ the total time required, and using the subscript 0 for values corresponding to the best magnitude, it can be shown that the answer to the foregoing question is given by the equation

$$\frac{\theta}{\theta_0}=\left(\frac{\psi\,\varphi\,F}{\psi_0\varphi_0F_0}\right)^2. \tag{7}$$

Demonstration. The error Δy in y due to an error Δx in x is by eq (5) $\Delta y=(\partial f/\partial x)\Delta x$, or $\Delta y=F\Delta x$ by definition of F. Also by definition of φ we may write $\Delta x=\varphi P$, in which P denotes the error in x when x has some standard value, and φ is a function of x. But from eqs (2) and (4), $P=p/\sqrt{n}=p\sqrt{t/\theta}$, so that $\Delta y=F\Delta x=F\varphi P=\frac{F\varphi}{\sqrt{\theta}}\cdot p\sqrt{t}$. Now if it is not the absolute error Δy, but in general $\psi\Delta y$ that we wish to minimize by our choice of best magnitude for x, then it is this same quantity $\psi\Delta y$ which in the present problem must be kept constant. But

$$\psi\cdot\Delta y=\psi\cdot\frac{F\varphi}{\sqrt{\theta}}\cdot p\sqrt{t}$$

so that

$$\psi_0\cdot\Delta y_0=\psi_0\cdot\frac{F_0\varphi_0}{\sqrt{\theta_0}}\cdot p\sqrt{t},$$

p and t being independent of x. Therefore, to maintain the same precision under adverse conditions, θ must be taken enough larger than θ_0 to keep $\psi\Delta y$ down to the size of $\psi_0\Delta y_0$. Equating the expressions for these last two quantities, and dropping the common factor $p\sqrt{t}$ gives

$$\frac{\psi F\varphi}{\sqrt{\theta}}=\frac{\psi_0F_0\varphi_0}{\sqrt{\theta_0}}$$

from which

$$\theta/\theta_0=(\psi\varphi F/\psi_0\varphi_0F_0)^2,\quad \text{Q.E.D}$$

As an illustration of eq (7), we may further consider the Fabry-Perot interferometer problem. In treating this problem in the previously mentioned paper,[4] the best magnitude for the order of interference was found to be 22,000. In other words, in that particular case, the front surface of the silvered plate, and the back surface of the transparent plate should be $\frac{1}{2}$ of 22,000 wavelengths or about 5 mm apart. Let us now inquire what would be the effect on the time needed to attain the same degree of precision if the plates were set either 1 mm or 25 mm apart. Expressing ψ and ψ_0, φ and φ_0, F and F_0 as functions of the order of interference x, and then replacing x first by the round number 4,000, corresponding to the distance 1 mm, and second by the value 100,000, corresponding to the distance 25 mm, it is found that in the first case

[3] J. Wash. Acad. Sci. **I**, 187 (1911).

[4] J. Wash. Acad. Sci. **I**, 187 (1911).

it would take approximately $\frac{\theta}{\theta_0}=8$, and in the second case $\frac{\theta}{\theta_0}=$ about 5 times as long to secure a fixed degree of precision as it would if x had been given the best magnitude 22,000.[5]

Solution. In this problem, it was required to determine an unknown wavelength λ in terms of the reference wavelength λ_r and the corresponding order of interference N_r by observing the order of interference N. For this purpose, the equation of the interferometer may be written $\lambda=\lambda_r N_r/N$, and we may consider $\lambda_r N_r$ without error. By order of interference is meant the integral and fractional number of wavelengths of the radiation in question in the optical path of the interfering rays. That is, $N=2D/\lambda$ where D is the inside distance between the plates. It might at first sight be supposed that N should be chosen very large in order that a given error in N should have the smallest possible fractional influence on the result. But from the viewpoint of time economy, this result would be wrong. It ignores the increasing diffuseness of fringes with large path-difference. Now the φ term in the criterion was designed to cover just such cases as this. Suppose that for several adjustments of the plates, the average deviation of fringe settings were found to increase with order of interference in the following manner:

$$\text{avg. dev. of } N=\frac{1}{100}\,(1+2\cdot 10^{-9}N^2).$$

We then take for the function φ, $1+2\cdot 10^{-9}N^2$. Changing over to our generalized notation by the substitutions

$$y\equiv\frac{\lambda}{\lambda_r N_r},\quad x\equiv N,$$

it is seen that, in measuring a fixed wavelength λ, a minimum fractional error is desired in y, so that $\psi=1/y$. But $y=1/x$. Consequently, the functions needed in eq (7) are $\psi=x$, $\varphi=1+2\cdot 10^{-9}x^2$, and $F=-(1/x^2)$. In the previous paper, it was shown that by differentiating these expressions, and substituting in the criterion, here eq (6), and solving for x, the best magnitude turned out to be $x_0=22{,}000$. Consequently, to solve the problem now before us, we need only substitute this value, together with the above expressions for Ψ, φ, and F, into eq (7). This gives

$$\frac{\theta}{\theta_0}=\left[\frac{22{,}000}{x}\cdot\frac{1+2\cdot 10^{-9}x^2}{1+2\cdot 10^{-9}\,(22{,}000)^2}\right]^2.$$

Now $x=N=2D/\lambda$ so that for wavelengths of $\frac{1}{2}$ micron, $x=4{,}000$ when $D=1$ mm and $x=100{,}000$ when $D=25$ mm. In the first case then

$$\frac{\theta}{\theta_0}=\left[\frac{22{,}000}{4{,}000}\cdot\frac{1+2\cdot 10^{-9}\,(4{,}000)^2}{1+2\cdot 10^{-9}\,(22{,}000)^2}\right]^2=8.3$$

while in the second case $\theta/\theta_0=5.5$.

3. Division of Measurement Time Among Component Quantities, and Grouping of Measurements of a Single Quantity, to Secure Results of Maximum Precision

The investigation of economy of time in taking the *observations* themselves leads to two distinct problems: first, that of the division of time among the components of an indirect measurement; second, that of the best grouping of observations in determining any one quantity.

The character of these two problems may be illustrated by considering a particular example of each. To illustrate the first, let it be required to determine the most economical division of time between length and period measurements in a gravity determination by Kater's pendulum. Having given the form which eq (1) takes in this case — namely $g=\pi^2L/T^2$ —, the question to be answered is: — What proportion of the total available time should be devoted to repeated measurements of the knife-edge distance L, and what proportion to continued observations of the period T?

As an instance of the second problem, let it be required to determine the most economical grouping of observations in a measurement of the heat of combustion of coal from a series of samples. Is it better to take many observations on a few samples or a few observations on many samples? [6]

3.1. Division of Time Among Components of an Indirect Measurement

The solution of the first problem comes out in terms of three data: namely, the precision p_i and the time t_i of a single observation of any component x_i, together with the derivative F_i, representing $\partial f(x\ .\ .\ .)/\partial x_i$ or the sensitiveness of the result with respect to the com-

[5] A number of small changes have been made in this paragraph and in the following "Solution" at Mr. Hersey's suggestion, in the interest of clarity. In this connection he has commented:

"Strictly, x should be 22000/5 = 4400 in the first case and 22000 x 5 = 110000 in the second, leading to θ/θ_0 results of 7.3 and 5.0. I had used round numbers to go along with the roughly stated distance of about 5 mm." (C.E.)

[6] Before reading further, the reader may wish to jot down his personal answers to the following quiz that Mr. Hersey gave to his conference participants on July 9, 1920, for comparison with the corresponding results yielded by the formula to be given shortly:

If you had at your disposal a certain amount of time and no more in which to determine the best representative value for some physical property of a substance by observations on a series of samples, and if you knew it would require 100 times as long to set up each new sample as it would to take each additional observation on a given sample, *how many observations* would you take on *each sample* before changing to a new sample, assuming

1° That you had reason to expect the average deviation of the different samples (due to lack of homogeneity of the substance) to be of the same order of magnitude as the average deviation of the observations on any one sample (due to accidental errors in the measuring apparatus)?

2° That you had reason to expect the deviations of the observations to be 10 times as great as the deviations of the samples?

Should the foregoing data appear to you either physically or numerically insufficient, kindly add whatever specifications you think suitable in order to make the problem a practical one, and then solve that.

Please give the two results numerically, . . . and, if your answer is based on any ground other than a conscientious guess, say what. (C.E.)

ponent in question. In general, the highest precision will be attained in a fixed time, or conversely a given precision will be secured in the least time, when the time assigned to any component x_i is given by the formula

$$\theta_i \propto p_i F_i \sqrt{t_i} \tag{8}$$

or in other words when the division of time is made in the proportion

$$\theta_1 : \theta_2 : \ldots \theta_i = p_1 F_1 \sqrt{t_1} : p_2 F_2 \sqrt{t_2} : \ldots p_i F_i \sqrt{t_i} \tag{9}$$

Demonstration. The error in the result of an indirect measurement y due to an error p_i in any one component x_i is equal to the rate of change of the result with respect to this component, times the error in the component, or, in the notation of eq (5), $(\partial f(x \ldots)/\partial x_i) \cdot P_i$. The square of the error in the result due to the simultaneous errors p_i in the several components x_i is therefore by eq (3) $\Sigma(F_i P_i)^2$ in which F_i has been written for $\partial f(x \ldots)/\partial x_i$. By eqs (2) and (4), $P_i = p_i \sqrt{t_i/\theta_i}$. Therefore,

$$(F_i P_i)^2 = \frac{p_i^2 F_i^2 t_i}{\theta_i} = \frac{C_i^2}{\theta_i}$$

in which C_i denotes $p_i F_i \sqrt{t_i}$. The square of the resultant error in y is consequently $\Sigma(C_i^2/\theta_i)$ while the total time available is $\Sigma\theta_i$. We have therefore to discover what relation among the several θ's will make a minimum of $\Sigma(C_i^2/\theta_i)$ while $\Sigma\theta_i$ remains constant. Imagining a family of equilateral hyperbolas constructed with θ_i as abscissas and C_i^2/θ_i as ordinates, the problem reduces to that of finding the locus of all points, one on each curve, satisfying the condition that the sum of their ordinates shall be a minimum, while the sum of their abscissas remains fixed. The locus proves to be a straight line through the origin defined by the relation $(\theta_i/C_i) = \text{constant}$. Hence, finally, $\theta_i \propto C_i \propto p_i F_i \sqrt{t_i}$, Q.E.D.

Our rule for the division of time does not require the assumption that concealed errors do not exist, and if the concealed errors which do exist are constant, then eq (9) leads not only to the condition for highest precision, but also to the condition for highest accuracy. This statement will be physically evident to most persons, though it could readily be proved by the use of eq (1).

Now to apply eq (9) to the pendulum example, we may postulate that a length measurement taking an hour's time can be repeated with an average fractional deviation $\Delta L/L$, which is one-half as large as the deviation $\Delta T/T$ in a period determination say of four hours' duration. If we further assume the pendulum in question is a second's pendulum, these data are sufficient to show that one-ninth of the observer's time should be spent in the knife-edge measurements and the remaining eight-ninths in swinging the pendulum.

Solution. Let y stand for g/π^2; x_1, for L; and x_2 for T. The characteristic equation of type (5) defining the measurement in question then becomes $y = x_1/x_2^2$. The elements of the problem which have been given numerically are: the period, $x_2 = 1$ sec; the length, $x_1 = y\,x_2^2 = 980(1)^2/\pi^2 =$ approx. 100 cm; the relative time of a single determination $t_1/t_2 = 1/4$; and the relative fractional precision measures $(p_1/x_1)/(p_2/x_2) = 1/2$, so that $p_1/p_2 = x_2/2x_1 = 50$ cm/sec. From the characteristic equation

$$F_1 = (\partial(x_1)/\partial x_1)/x_2^2 = 1/x_2^2 = 1 \text{ sec}^{-2}$$

and

$$F_2 = \frac{-2}{x^2}\frac{(x_1)}{x_2^2} = \frac{-2x_1}{x_2^3} = \text{numerically } 200 \frac{\text{cm}}{\text{sec}^3}.$$

For two components eq (9) becomes

$$\frac{\theta_1}{\theta_2} = \frac{p_1}{p_2} \cdot \frac{F_1}{F_2} \cdot \sqrt{\frac{t_1}{t_2}}.$$

In this problem $p_1/p_2 = 50$ cm/sec, $F_1/F_2 = (1/200)$ sec/cm, $\sqrt{t_1/t_2} = 1/2$. Hence $\theta_1/\theta_2 = 50 \times 1/200 \times 1/2 = 1/8$, from which $\theta_1/(\theta_1 + \theta_2) = 1/9$ and $\theta_2/(\theta_1 + \theta_2) = 8/9$.

3.2. Best Grouping of Observations in Determining Any One Quantity

Turning to the problem of best-grouping, it is seen that the solution consists in establishing the most profitable compromise between the extremes of, first, repeating a large number of readings under the same conditions (or on the same sample) in order to diminish the effect of *observational* errors; or, second, endeavoring to cut down *systematic* errors by numerous independent determinations (or by trying many different samples). The conception of sampling is introduced to fix the ideas, but our reasoning is applicable to any change in physical conditions analogous to the alternation of samples. Now the most economical number of observations to make in any one group before stopping to change conditions (or to set up a new sample) in preparation for a new group, is directly expressible in terms of two postulated data. These are, first, the ratio, σ, of the average observational error to the average systematic error anticipated; and, second, the ratio, τ, of the time required in preparing for a new group to the time used in a single observation. The most economical number of observations in a group is in fact

$$n_0 = \sigma\sqrt{\tau} \tag{10}$$

to which of course may be added any allowance that we wish to make as a check against mistakes. It is important to note that the solution is independent of

the total time available. Extra time is better spent in setting up new groups.[7]

Demonstration: Imagine first that an indefinitely great number of observations were made on each sample. The mean result on each sample would then differ from the mean of a great number of samples by exactly the reliability of the sample, r_1. Dropping subscripts to denote mean values, $r/\sqrt{N}$ would then be the precision of the mean of N samples. But in the actual case of only n observations on a sample, the result on each sample differs from the final mean by $r_1 \pm P_1$ in which P_1 is the precision of the result on one sample. By (2) and (3) the mean value of this deviation can be written $\sqrt{r^2+(p^2/n)}$ in which p is the precision of a single observation. In an actual case, instead of $r/\sqrt{N}$ we then write $\sqrt{r^2+(p^2/n)}/\sqrt{N}$ for the precision of the final result. It is this quantity which is to be made a minimum by properly choosing n and N, both of which cannot be simultaneously increased owing to the fact that the total available time $\theta = N(T+nt)$ remains fixed. This latter equation is simply the appropriate interpretation of eq (4) with T for the time used in preparing a new sample. Eliminating N from the expression for final precision by reference to the θ equation, the former reduces to

$$\sqrt{\frac{r^2 t}{\theta}\left(1+\frac{\sigma^2}{n}\right)(\tau+n)}$$

in which σ has been written for p/r and τ for T/t. While the factor r^2t/θ may be unknown it is assumed to be constant so the condition sought for is simply

$$\frac{\partial}{\partial n}\left[\left(1+\frac{\sigma^2}{n}\right)(\tau+n)\right]=0,$$

together with a + second derivative. The value of n which satisfies this condition is found to be $n_0 = \sigma\sqrt{\tau}$. Q.E.D.

In order to apply eq (10) to the heat of combustion problem, let us postulate that the average deviation of the results on any one sample (whether due to observational errors or to imperfect mixing before subdividing the sample for check determinations) is of the same order of magnitude as the average deviation of the different samples, while the time required to secure and prepare a new sample is approximately the same as that for repeating a determination on any one sample. In this event, a single test of each sample is sufficient. The precision of the final result can be indefinitely improved by now collecting a greater number of samples, while by increasing the number of tests on each sample, the ultimate precision would actually be made worse. If, however, the preparation of a sample were to consume tenfold as much time as the repetition of a test, then it would be well to make three or four tests on each sample.

Solution. In the first case the precision and the reliability of an observation, though numerically unknown, are assumed equal, so that $\sigma = 1$. Similarly $\tau = 1$. Substituting in eq (10) gives at once $n_0 = 1\sqrt{1} = 1$. In the second case again $\sigma = 1$, but τ is assumed = 10. Hence (10) gives $n_0 = 1\sqrt{10} = 3+$.

A combination of the two problems will arise when the components of an indirect measurement are not determined by a continuous series of observations. But in this case too an equally straightforward solution can be deduced.

4. Economizing Computing Time in Determining Constants of Empirical Equations

Lastly, and from the viewpoint of the computer, a question to be considered here is that of most conveniently determining the constants of empirical equations. An obvious means of economizing time is in the further use of graphical methods and particularly in the rectification of the plot by some change of variable which will throw the assumed relation into linear form. Almost any relation can be transformed into a straight line from which two unknown constants may be found by inspection, while three or more can be determined by successive approximations. Why, then, are graphical methods so commonly neglected? Possibly on account of an unjustifiable faith in least squares. Possibly, also, on account of the straight line method frequently proving illusory in practice by giving undue influence to the observations at one end of the range. This embarassing fact is familiar to those who have used logarithmic coordinate paper. Having been obliged on this ground to abandon the straight line, the computer may very naturally turn to the least squares adjustment in preference to the drawing in of a curve, because the latter involves a greater element of personal judgment than a straight line.

A method has been found which, it is hoped, will remedy this objection. The proposed method has not been put to the test of practical use, and, therefore, no certain claims can be made for it, but in principle, at least, it restores to the straight-line transformation all the advantages commonly attributed to it. This consists in reweighting all the points plotted in

[7] The above formula for "the most economical number of observations in a group" was employed, and attributed to Mr. Hersey, as follows, in the section on "Desirable Number of Shots Per Filling" of appendix I (on the "Joint Comparative Tests of Spark Gaps for Testing Transformer Oils," carried out in the spring of 1917) to the Report of ASTM Committee D-9 on Electrical Insulating Materials, ASTM *Proceedings* **21**: 404(1921): —

The question of how many shots to take from a single filling is, therefore, to be determined entirely by the relative length of time required to make a shot and to refill the testing vessel, and by the possibility of contaminating a single filling of oil while pouring it from the container to the testing cup. It may be shown that if t = time required to make one shot and T = time required to make a fresh filling, and if p = probable error of a single shot from the mean of all shots on one filling, and P = probable error of mean of one filling from mean of all fillings, then the time of the observer is most economically spent if the number of shots per filling is $n = \left(\frac{T}{t}\frac{p^2}{P^2}\right)$.*

*This relation was deduced by M. D. Hersey in some as yet unpublished work on the theory of errors.

Dr. Francis B. Sillsbee (NBS, 1911–1959; now Consultant, NBS), who prepared appendix I on behalf of the Committee, tells me that this formula was actually used in the correct and (nowadays) more familiar form $n = \frac{p}{P}\sqrt{\frac{T}{t}}$, which the editor apparently squared to avoid radicals; and the left-hand side was unfortunately printed unsquared, as "n". (C.E.)

the X, Y diagram in proportion to the factor

$$r=\left[\frac{dy/dx}{dX/dx}\right]^2 \tag{11}$$

the assumed relation

$$y=f(a, b, x) \tag{12}$$

having been transformed into the straight line

$$Y=A+BX \tag{13}$$

by the substitutions

$$X \equiv \text{some function of } x \text{ and } y$$

$$Y \equiv \text{some function of } y \text{ and } x$$

thus leading to the evaluation of a and b in terms of the intercept A and slope B. The reweighting of an observation in this ratio r consists merely in treating it as if that observation had been repeated r times during the experiment.

Demonstration. Equation (11) assumes that the curve (12) and the line (13) satisfy the conventional least squares condition. Denote by W the weight in the X, Y diagram of an observation of deviation ΔY, whose weight and deviation in the original X, Y diagram were w and Δy. Physically, we are interested only in minimizing the sum of the (weighted) squares of the original deviations Δy; but we can accomplish this by minimizing the sum of the weighted squares of the fictitious deviations ΔY if the new weights W are so chosen as to satisfy the condition $\Sigma w(\Delta y)^2 = \Sigma W(\Delta Y)^2$ or $W/w=(\Delta y)^2/\Delta Y$. But if the deviations be so small that the slope dy/dx is sensibly the same at y and $y+\Delta y$, the deviations can be treated as infinitesimals, writing

$$\frac{W}{w}=\left[\frac{dy/dx}{dY/dx}\right]^2.$$

Now by eq (13) $dY=BdX$, so that if r be written for $B^2 \cdot W/w$, we get for the reweight factor, expressible as a function of x, $r=\left[\frac{dy/dx}{dX/dx}\right]^2 \cdot$ Q.E.D.

As an application of the reweight factor, consider the determination of the thermal expansion coefficients a and b in the equation

$$L=L_0(1+at+bt^2)$$

in which the directly observed quantities are the length increment $L-L_0$ and the temperature t. We assume that accidental errors in t are negligible compared with those in $L-L_0$. This assumption is implicitly contained in the conventional method of least squares adjustment, and it underlies eq (11). Reverting now to our standard notation by setting

$$\frac{L-L_0}{L_0} \equiv y, \text{ and } t \equiv x,$$

we have in the x, y plane a nonlinear equation

$$y=ax+bx^2$$

of the type (12), which can be transformed into the linear equation (13) in the X, Y plane by the substitutions

$$Y \equiv \frac{y}{x}$$

$$X \equiv x$$

while the four constants satisfy the relations $A=a$, and $B=b$. In other words, by plotting out values of y/x as ordinates against x as abscissas, we now get a straight line, the Y-intercept and slope of which are respectively identical with a and b.

Unfortunately, however, if the deviations in the original x, y diagram were of the same order of magnitude all along the curve, then in the rectified X, Y diagram, (provided the specimen is one for which both a and b are +), the deviations will be very much larger at low temperatures. This gives the low-temperature observations a disproportionate influence on the final result. The reweight factor is intended to compensate for this effect. If the specimen has twice the expansivity at high temperature that it has at low, then it can be shown that we should reweight the high temperature observations relatively to the low-temperature observations in the ratio 4 to 1.

Solution. Since $X=x$, $dX/dx=1$, and, therefore, by (11) the reweights are proportional to $(dy/dx)^2$, the square of the slope of the original x, y curve (12). The slope of this curve is the thermal expansivity. If the expansivity is twice as great at the high temperature end of the range as it is at the low, the reweights increase fourfold in the same interval. They may conveniently be applied by dividing the whole range into four zones in which the observations are respectively assigned one, two, three, and four times their original weights.

Having rectified and reweighted his data, the computer is now at liberty to evaluate the constants a and b by any of the following methods:

First, by the numerical but *linear* least-squares computation;

Second, by the use of an automatic device for linear least squares adjustment, a rubberband model of this being exhibited on the lecture table [at the 718th meeting of the Philosophical Society of Washington held at the Cosmos Club on January 18, 1913];[8]

[8] And described briefly in the Proceedings of the meeting (J. Wash. Acad. Sci. **III**, 296, (1913)), as follows:

Mr. M. D. Hersey presented a paper on *A mechanical model of the least square adjustment*. The apparatus exhibited consisted of a sheet of coordinate paper mounted on a board for the plotting of points, a light aluminum rod, and a supply of rubber elastics and push pins. The values under discussion were plotted by the push pins, allowance being made for the unstretched lengths of the elastics by which the bar was suspended. The model was used to show the mechanical adjustment of the tests of a mercury barometer; the results were compared with those by the usual solution of normal equations. The speaker discussed methods for weighting different observations, the determination of the probable error by the model, and the application of the principle for solution of case involving several unknowns. (C.E.)

Third, by the simple drawing of a straight line, which in technical physical measurements is usually sufficient.

5. Summary

In conclusion, the principal working rules which have been formulated above may be briefly catalogued as follows:

I. From the viewpoint of the designer:

(a) The amount of time which can theoretically be saved by selecting the "best magnitudes" for the optional elements of an apparatus, instead of designing it or assembling it haphazard, is given by eq (7).

II. From the viewpoint of the observer:

(b) The best division of time among the components of an indirect measurement is that given by eq (9).

(c) The most economical grouping of observations in order so far as possible to eliminate systematic errors is given by eq (10).

III. From the viewpoint of the computer:

(d) The constants of practically any empirical equation can be found (two at a time) by the simple inspection of a straight line if the observations be reweighted according to eq (11).

The commonly accepted formulas of the theory of errors which underlie these rules have already been stated. Different assumptions would lead to different rules. The rules can properly be applied only to facts which fit the assumptions approximately. But the use of the theory of errors in establishing technique is a different order of affair from its use in evaluating results, and one in which a rougher degree of approximation is sufficient.

Particular expedients for saving time will continually be discovered by investigators in mechanical, thermal, optical, and electrical measurements individually and specifically considered. This paper is not intended as a substitute for such investigations, but rather as a foundation or a starting point for them, in establishing, to begin with, certain relations which are independent of the particular nature of the experiment.

(Paper 69B3–144)