REVIEWS OF MODERN PHYSICS VOLUME 19, NUMBER 4 OCTOBER, 1947

# Least-Squares' Fitting of Data by Means of Polynomials

RAYMOND T. BIRGE
*University of California, Berkeley, California*

## Mathematical Appendix

J. W. WEINBERG*
*University of California, Berkeley, California*

### TABLE OF CONTENTS



### LIST OF TABLES



### LIST OF MODEL FORMS



---

* Now at the University of Minnesota, Minneapolis, Minnesota.

## LIST OF IMPORTANT SYMBOLS

(with page reference of first important mention)

| *Page* | *Symbol* | *Explanation* |
|---|---|---|
| 302, 313 | $v_j(\epsilon)$ | Residual (=obs.-calc. value $=y-u$) for a polynomial of degree $j$ |
| 303 | $n$ | Total number of observations |
| 303 | $t$ | Degree of polynomial. Used especially in connection with the series of orthogonal polynomials ($t=0$ to $j$) whose sum gives the desired least-squares' polynomial of degree $j$ (see footnote 14) |
| 304 | $\epsilon$ | Special abscissa, varying from $-q$ to $+q$, by unit interval. Hence, $n=2q+1$ |
| 304 | $e$ | Special abscissa, varying from $-qh$ to $+qh$, by interval $h$. Hence, $e=\epsilon h$ |
| 304 | $x_0$ or $m$ | Middle observation $x$, for which $\epsilon$ or $e=0$. Hence, $x-x_0=e=\epsilon h$ |
| 304 | $\delta^t y_\epsilon$ | Central difference of order $t$, of observations $y_\epsilon$ |
| 304 | $\mu\delta^t y_0$ | $\equiv \frac{1}{2}(\delta^t y_{-\frac{1}{2}}+\delta^t y_{\frac{1}{2}})$ |
| 304 | $u_j(\epsilon)$ or $u_\epsilon$ | Calculated ordinate of $j$th degree polynomial, for $\epsilon$ scale of abscissa |
| 304 | $y(\epsilon)$ or $y_\epsilon$ | Observed ordinate, for $\epsilon$ scale of abscissa |
| 306 | $T_t$ | Orthogonal polynomial of degree $t$, a function of $n$ and $\epsilon$ only. See Table III for explicit forms $T_0$ to $T_{10}$. Evaluated by Eq. (15) |
| 306 | $a_t$ or $a_{tt}$ | Coefficient of $\epsilon^t$ in polynomial of degree $t$; also the coefficient of $T_t$ in Eq. (5). Evaluated by Eq. (7) |
| 307 | $R_{kt}$ | Coefficient of $\epsilon^k$ in expression for $T_t$ (Table III). It is a function of $n$ only. $R_{tt}=1$ |
| 308 | $M_t$ | $=\Sigma_\epsilon T_t^2$, the statistical weight of $a_t$; see Section C5. Evaluated by Eq. (16) |
| 308 | $a_{kj}$ | Coefficient of $\epsilon^k$ in the power series polynomial of degree $j$, Eq. (9). Evaluated by Eq. (10) |
| 309 | $P_B$ | $=F_B\cdot T_t=2t!T_t/(t!)^2$, Eq. (11). It is the so-called "pair-factor" of the Birge and Shea method, Eq. (14) |
| 309 | $K_t$ | The "denominator" in the Birge and Shea Eq. (14). Defined by Eq. (13) and evaluated by Eq. (16) or Eq. (48) |
| 309 | $r$ | $=\lvert\epsilon\rvert$, designating a "pair" of observations of equal positive and negative $\epsilon$ value. The symbol $r$ is used also for probable error in general, and $r'$ for probable error of a quantity of unit weight, Eqs. (33), (34) |
| 309 | $P_M$ | $=F_M\cdot T_t$, the "pair-factor" of Milne method, Eq. (17) |
| 310 | $P_J$ | $=F_J\cdot T_t$, the "pair-factor" of Jordan method, Eq. (18) |
| 310 | $P_T$ | $=F_T\cdot T_t$, the "pair-factor" of Tchebycheff method, Eq. (19) |
| 310 | $a_t'$ | $=\Sigma\, y\cdot P/\Sigma\, P^2$, Eq. (20), a general symbol for any modified $a_t$ |
| 311, 318 | $f$ | Special integer factor, a function of $n$ and $t$, which may be removed from $P_B''$ of Eq. (15) to get value of $V_t$. See footnote 22. The same factor $f$ may be removed from $Q_t'$ of Eq. (51) to get $W_t$ of Eq. (55), and from $K_t'$ of Eq. (53) to get $L_t$ of Eq. (56). Numerical values of $f$ in Table VI. See footnote 29 |
| 311, 312 | $V_t$ | $=S_{tt}T_t$=minimum integral values of the set of "pair-factors" for a specified value of $n$ and of $t$. Numerical values listed in Table XIII |
| 311 | $b_t$ | $=a_t/S_{tt}$=coefficient of $V_t$ in the $b$-$V$ system, Eq. (22) |
| 311 | $N_t$ | $=\Sigma\, V_t^2$=statistical weight of $b_t$ in Eq. (23). Numerical values of $N_t$ in Table XIII |
| 311 | $K_t^*$ | $=N_t/S_{tt}$="denominator" actually tabulated by Birge and Shea, corresponding to the tabulated "pair-factors" $V_t$, by Eq. (25) |
| 312 | $S_{kt}$ | Coefficient of $\epsilon^k$ in expressions for $V_t$, corresponding to Table III for $T_t$, and used in Eq. (27). It is a $f(n)$ only. Numerical values of $S_{kt}$ in Table XII |
| 311 | $S_{tt}$ | Special $S_{kt}$ factor, connecting the $a$-$T$ and $b$-$V$ systems (see Table V). Numerical values of $S_{tt}$ in Table XII. It is denoted $\lambda_t$ by Fisher |
| 313, 327 | $r'$ | Probable error of hypothetical quantity of unit weight, given by Eq. (34) or Eq. (78) |
| 313 | $s$ | Number of undetermined coefficients in a function. See Eq. (34). Also used for the running parameter in a summation |
| 313 | $p_j(\epsilon)$ | Weight of the polynomial $u_j$ of degree $j$, at specified $\epsilon$, evaluated by Eq. (35) or Eq. (36) |
| 314 | $p_{kj}$ | Weight of $a_{kj}$, evaluated by Eq. (37) or Eq. (38) |
| 314 | $a_{kj}'$ | Coefficient of $\epsilon^k$ of power series in $e(=\epsilon h)$ of degree $j$ |
| 314 | $a_{kj}''$ | Coefficient of $x^k$ in power series in $x(=e+x_0)$ of degree $j$, evaluated by Eq. (40) or Eq. (42). Explicit expressions for $a_{kj}''$ in Table IV |
| 314 | $p_{kj}''$ | Weight of $a_{kj}''$, evaluated by Eq. (41) or Eq. (43) |
| 314 | $G_{kt}$ | Functions of $m(=x_0)$, $h$ and $R_{kt}$, needed for the evaluation of $a_{kj}''$ in Eq. (40). Explicit expressions for $G_{kt}$ may be obtained from Table IV, by replacing $S_{kt}$ by $R_{kt}$ |
| 314 | $H_{kt}$ | Functions of $m$, $h$, and $S_{kt}$, needed for the evaluation of $a_{kj}''$ in Eq. (42). Explicit expressions for $H_{kt}$ in Table IV |
| 316 | $2\nu$ | $=2q-t=n-t-1$. A set of finite differences $\delta^t y_\epsilon$ runs from $\epsilon=-\nu$ to $+\nu$, just |

| | | |
|---|---|---|
| | | as $\epsilon$ runs from $-q$ to $+q$ for the observations $y_\epsilon$ (which correspond to $t=0$ difference) |
| 316 | $Q_t$ | Weighting factor of $\delta^t y_\epsilon$, Eqs. (47) and (49), for use with "denominator" $K_t$ of Eq. (48) |
| 317 | $Q_t'$ | Smallest integer value of weighting factor, for a given value of $t$, in Eq. (52). Evaluated by Eq. (51), see footnote 30 |
| 317 | $K_t'$ | "Denominator" in Eq. (52), corresponding to weighting factor $Q_t'$. Evaluated by Eq. (53) |
| 317 | $W_t$ | Smallest integer weighting factor of $\delta^t y_\epsilon$, for a given value of $n$ and $t$, in Eqs. (54), (55). Values listed in Table XIV |
| 317 | $L_t$ | Smallest integer "denominator" to use with $W_t$, Eqs. (54), (56). Values listed in Table XIV |
| 318 | $Q_t''$ | Weighting factor of $\Delta^t y_x$, Eqs. (57), (58), to use with "denominator" $K_t'$ of Eq. (53) |
| 335 | $m_t$ | $=\sum x^t y=$power moment of the observations |

## A. INTRODUCTION

A VERY common problem in physical science is that of the representation of a set of experimental data by means of a smooth curve. For many reasons it is frequently desirable to use for this purpose an analytic function, and to obtain such a function in explicit, numerical form. Of all the functions thus used in science, the most common is undoubtedly the *polynomial* (often denoted a rational integral function). The truth of this last remark is evident from the fact that the arithmetic average of a set of $n$ unweighted observations (or the weighted average of a set of weighted observations) represents the least-squares' fitting of the observations by means of a polynomial of zero degree. Similarly the commonly found, linear relation between $x$ and $y$ is, of course, a polynomial of the first degree.

Methods for obtaining the least-squares' solutions of zero- and first-degree polynomials are well known and in frequent use. The *general* method for obtaining the least-squares' solution of a polynomial of any degree, by means of determinants, is also rather well known. But the numerical labor involved in the case of polynomials of the third or higher degree is so great that such a solution is rarely attempted.

If the observations are *equally spaced* along the abscissa axis, and are of *equal weight*, the situation is very different.[1] Such a regularity permits an enormous simplification of the necessary process, and the least-squares' fitting of a polynomial of the fifth or even higher degree can actually be carried out with reasonable ease and accuracy. One might naturally anticipate such a situation, and hence it is not surprising that a number of important papers have been published on the subject. A partial list of such material, alphabetic by author, is given in Section H.

The history of the matter is, however, a most curious one. To the writer it illustrates a more or less inevitable result of the great expansion and specialization of science that has been evolving now for a number of generations. No scientist has the time to read even a small fraction of the papers that might actually be of service to him in his own investigations. What is still more unfortunate is that increased specialization of subject has been accompanied by increased specialization of symbols and nomenclature, so that it is often difficult to follow intelligently papers in fields other than one's own.

In the case of least-squares' fitting by means of polynomials, most of the fundamental work has been done by mathematicians. Some has been done by statisticians who, incidentally, carried out such investigations because they desired to make practical use of the results. But, in general, mathematicians do not make actual numerical use of their equations, and often leave them in a form not convenient for numerical use. On the other hand, the physical scientists, who could use the results with great profit, all too often are quite unaware of the existence of such material, and have difficulty in perceiving its true significance, even when they do happen to notice it in the literature.

The history of the present subject is indicated briefly at scattered points in the paper, and at the end (Section G) certain alternative processes

[1] We assume here that there is no experimental error in $x$, or at least that such error is negligibly small in comparison with that in $y$. When both coordinates are subject to significantly large errors, the corresponding least-squares' solution becomes extremely complicated in the case of any function more complex than a polynomial of the first degree. See W. E. Deming, Phil. Mag. **11**, 146 (1931).

are discussed. But in order not to confuse the reader who desires to know merely what the pertinent results are, and how they may be used, the main body of the present paper is devoted primarily to the presentation of such explicit information, with as little interruption as possible in the form of historical remarks.

The particular methods of calculation presented here are believed by the writer to be the most advantageous for use by physical scientists. It must, however, be understood that there are many possible alternative methods and those who have proposed such alternative methods may have quite diverse opinions on the matter. Certain of these alternative methods, as just noted, will be considered near the end of the paper, when they can be compared more conveniently with the methods advocated here.

The writer's previous work in this field consists of two papers, one published in 1919 on the least-squares' fitting of a second-degree polynomial,[2] and the other, written jointly with Dr. J. D. Shea, in 1927, on the least-squares' fitting of a polynomial of any degree.[3] The writer admits with regret that when these previous investigations were carried out, he was totally unaware of the earlier work in the field. Since then numerous papers along this general line have been published and from such papers one may collect an almost complete set of important references. But it should be pointed out that, in spite of the fact that many dozens of papers on this subject are now in print, a new paper still appears occasionally, written by a person who, like the writer twenty years ago, is obviously in ignorance of all previous work.

Partly as a result of the information given in certain papers, and partly as a result of recent new investigations by Dr. Weinberg and the writer,[4] it is now possible to modify and extend the material of the Birge and Shea paper in a substantial way. In particular, Dr. Weinberg and I have been able to get a simple method for evaluating the probable errors of all quantities of interest. No general method for doing this has heretofore been presented in the literature. Thus the purposes of the present paper are two-fold—(1) to reintroduce into print the main facts and tables of the Birge and Shea method[3] since the paper itself has long been out of print, and (2) to present in detail, with certain necessary new tables, important modifications and extensions of the Birge and Shea method.

The detailed description of the modified Birge and Shea method is presented in Section C, and a brief discussion of an alternative method, which yields the desired results explicitly in terms of finite differences rather than in terms of observations, appears in Section D. Sections B and E contain details on two relatively elementary problems that appear, however, with great frequency in scientific work.

In Section F a specific problem is worked out in complete numerical detail in order to illustrate the methods presented in Sections C and D. A summary of these details constitutes Section F8. The main tables (XII, XIII, XIV) needed in the work constitute Section I, and the remaining tables, all of them brief, are scattered throughout the paper. It is therefore hoped that the reader who is not interested in the theoretical aspects of the subject will be able, from Section F alone, to understand the various calculations that are necessary in order to obtain the desired results.

As an illustration of the character and value of the information that can be obtained with truly remarkable rapidity, by means of the methods about to be presented, we consider the variation of the electrical equivalent of heat with temperature. This variation, which is merely that of the specific heat of water with temperature, was measured with great precision by Jaeger and Steinwehr[5] in 1921, and their experimental data were analyzed by the writer[6] in 1929 in connection with the determination of the most probable value of the electrical equivalent of heat. Additional information can now be obtained from their experimental material, and the details are as follows.

---

[2] R. T. Birge, Phys. Rev. **13**, 360 (1919).

[3] R. T. Birge and J. D. Shea, Univ. of Cal. Pub. in Math. **2**, 67–118 (1927). These results were first presented to the American Physical Society in 1924 (Phys. Rev. **24**, 206A, (1924)).

[4] R. T. Birge and J. W. Weinberg, Phys. Rev. **68**, 106A (1945); also J. W. Weinberg, Phys. Rev. **62**, 304A (1942). The theoretical contributions of Dr. Weinberg are summarized in Section J of the present paper.

Jaeger and Steinwehr measured 67 values of the electrical equivalent $J'$, at mean temperatures ranging from 4.75°C to 49.60°C. Since the various temperatures are spaced at unequal intervals, the writer in 1929 first collected the data into 19 points, spaced at 2.5°C intervals, from 5° to 50°C. This process, denoted by astronomers as the "formation of normal places," was carried out with the utmost care, and the resulting 19 points, all of equal weight, can be taken as a reasonable reproduction of the original data. Jaeger and Steinwehr fitted their data to a second-degree polynomial and then interpolated on this curve, at $t=15$°C, to obtain $J'_{15}=4.18420$ int. joules. But as the writer has shown,[6] at least a *fourth*-degree polynomial is required for the proper representation of their data, and with such a function one obtains $J'_{15}=4.18327$ int. joules. As already noted, it was not possible in 1929 to calculate in any simple way the probable errors of these results.

The best criterion of the fit of any function to the experimental data is furnished by the magnitude of $\sum v^2$, i.e., the sum of the squares of the residuals of the several (unweighted) points. By the use of the methods to be described one can obtain with comparative rapidity the $\sum v^2$ for a least-squares' polynomial of any degree, and in the process one obtains *simultaneously* the value of $\sum v^2$ for the least-squares' polynomials of all *lower* degrees. Beginning then with the zero-degree solution and going up to the fifth degree the respective values of $\sum v^2$ for the 19 points just described, in terms of $10^{-4}$ joule as the unit, are (0) 69,812.841, (1) 36,908.491, (2) 1464.027, (3) 15.510, (4) 1.051, (5) 0.084.

Thus the value of $\sum v^2$ is reduced from 1464.027 to 1.051 by the use of a fourth-degree polynomial, in place of one of only the second degree. A further improvement can apparently be made by the use of the fifth-degree solution, but the improvement is actually illusory because of the uncertainty introduced in connection with the reduction of the observations from 67 to 19 points. This uncertainty turns out to be of the same order of magnitude as the probable error of the fourth-degree solution.

[5] W. Jaeger and H. v. Steinwehr, Ann. d. Physik 64, 305 (1921).

[6] R. T. Birge, Rev. Mod. Phys. 1, 1 (1929).

We next calculate the value of the function, at $t=15$°C, for the fifth-degree solution. Such a single value can be obtained directly, independent of any other value and, as in the case of the evaluation of $\sum v^2$, its determination furnishes *simultaneously* the corresponding value for all least-squares' polynomials of lower degree. The experimental point used here for 15°C, is $41832.70\times 10^{-4}$ int. joules, and this value is used in calculating the respective residuals ($v$). Finally, we can easily evaluate the *probable error of the function*, at $t=15$°C, for each of the least-squares' polynomials. All of these results, together with the change in the calculated value with changing degree of the polynomial, are given in Table I.

Thus the actual change in the calculated value of $J_{15}'$, in passing from a second-degree to a fourth-degree polynomial, is $10.78\times 10^{-4}$ joule. This change is not only 5.3 times the probable error of the second-degree solution, but is 124 times that of the fourth-degree solution. Hence one sees clearly the unreliability of the value of $J_{15}'$ based on a second-degree least-squares' fitting of the data, as carried out by Jaeger and Steinwehr, and the very great improvement resulting from the use of a fourth-degree solution. One is, of course, considering here only the purely *accidental* errors of the experiment, as shown by the "scatter" of the points from a smooth curve. It is only such accidental errors that can be revealed and measured by any purely mathematical treatment of a single set of data.

TABLE I. Summary of treatment of data used in determining $J'$.

| Degree of function | $J'_{15}$ (calc.) ($10^{-4}$ joule) | $v$ ($10^{-4}$ joule) | Probable error ($10^{-4}$ joule) |
|---|---|---|---|
| 0 | 41,813.62 | +19.08 | ±9.64 |
| | +37.99 | | |
| 1 | 41,851.61 | −18.91 | ±9.23 |
| | −8.08 | | |
| 2 | 41,843.53 | −10.83 | ±2.02 |
| | −9.89 | | |
| 3 | 41,833.64 | −0.94 | ±0.262 |
| | −0.894 | | |
| 4 | 41,832.746 | −0.046 | ±0.087 |
| | +0.010 | | |
| 5 | 41,832.756 | −0.056 | ±0.021 |

As one further illustration of the use of the material about to be presented, let us consider the determination of the acceleration of gravity by means of an Atwood's machine. In this case, the second difference of points equally spaced in time should be constant. One then naturally first calculates a difference table, in order to determine whether the second differences are, in fact, constant, except for statistical variations. Let us suppose that one thus obtains 10 second differences. Then a *properly weighted average* of these ten differences gives the value of $2!a_2$, where $a_2$ is the coefficient of $x^2$ in the least-squares' solution of the data by means of a second-degree polynomial. From Table XIV of this paper, one finds that the respective weights $W_2$ for the 10 second differences (or 12 observations of position, i.e., $n=12$, $t=2$ of Table XIV), are 55, 135, 216, 280, 315, 315, 280, 216, 135, and 55. The sum of these weights (denoted $L_2$) is 2002, as also given in the table. The series of weights is always symmetrical about the center, and hence only the last half of the values appears in the table.

The table gives also the proper set of weights for any number of observations ($n$) up to 30, for the highest differences not only of the second-degree polynomial ($t=2$), but also of $t=1$, 3, 4, and 5. If, as has been done only too often in practice, one takes merely the arithmetic average of the second differences, one thereby *cancels out* automatically all but the first two and last two observations of position. So far as I know, a table such as Table XIV of this paper has not been published previously. With its use, the *least-squares'* value of the quantity sought (i.e., the coefficient $a_t$ of $x^t$ in a polynomial of degree $t$), can be obtained almost as rapidly as a value that admittedly fails to make any use of the major portion of the available experimental data.

The probable error of any coefficient $a_t$ may also be evaluated, but as shown in detail in Sections D and F7, the calculation of the probable error is always a more involved process than the calculation of the coefficient itself. The complexity arises primarily from the fact that the probable error in any coefficient, whether of the highest degree term of the polynomial or of any other term, involves the value of $a_t$ not only for $t=j$ (the degree of the polynomial) but for *all* values of $t$ from 0 to $j$.

## B. EXACT FITTING OF DATA

The main purpose of this paper is to discuss the *least-squares'* fitting of data by means of polynomials. A special case to which the suggested procedures still apply is that in which the number of observations just equals the number of undetermined coefficients (i.e., $t+1$ observations fitted to a polynomial of degree $t$). In that case each observed point is exactly fitted by the calculated function, all residuals are zero, and the question of probable error does not enter. If the main purpose of setting up such a polynomial is to use it for interpolation, then the best method is undoubtedly to employ one of the well-known interpolation formulas. Such formulas are merely polynomials in factorial form. If one wishes, finally, to express the polynomial in power-series form, a simple transformation from the factorial form makes this possible.

Before passing to the general case of least-squares' fitting, I accordingly first give a method of setting up a polynomial in power-series form which exactly satisfies all of the observed points. As usual, we are here considering only points equally spaced along the abscissa axis. If they are unequally spaced, one is forced to use one of the entirely general methods for the solution of a set of simultaneous linear equations.

Every process discussed in this paper will be illustrated by a fully worked-out numerical example. For this purpose I have chosen a set of seven equally spaced observations, for which the fourth differences are nearly constant. Hence the set is well represented by a fourth-degree polynomial. The numerical values have been deliberately chosen in such a way that all of the resulting coefficients and other desired quantities are given by terminating decimals. Thus *exact* numerical results may be obtained and an exact comparison of various procedures becomes possible.

In order to illustrate the fitting of $t+1$ points by means of a polynomial of degree $t$, we choose merely the first five of the seven points of our standard set and fit these points to a fourth-degree polynomial. The data and necessary differences are given in Table II.

For the case of least-squares' fitting, we always introduce a new abscissa, $\epsilon$, which varies from $-q$ to $+q$ where $n=2q+1=$ number of observations. Thus $\epsilon=(x-x_0)/h=e/h$, where $h$ is the interval, in terms of $x$, between successive observations, and $x_0$ is the middle observation, corresponding to $\epsilon=0$. In Table IV $x_0$ is replaced, for convenience, by $m$. If $n$ is even, $x_0$ (or $m$) is the value of $x$ half-way between the middle *pair* of observations, and the actual observations then correspond to $\epsilon=\pm\frac{1}{2}$, $\pm\frac{3}{2}$, etc.

On the other hand, in the case of *interpolation* formulas $\epsilon$ is so defined as always to have *integral* values for the actual observations. Then, in the case of an *even* number of observations, *either one* of the central pair of observations is taken as $\epsilon=0$ and the symmetry of the limiting plus and minus values of $\epsilon$ is lost. There are a great variety of interpolation formulas in common use, but I advocate the use of *central* difference formulas. A central difference of order $t$ is denoted by $\delta^t y$. The attached subscript indicates the value of $\epsilon$, as just defined for such formulas. All differences lying on the same horizontal line have the same subscript; $\mu\delta^t y_0$ represents $\frac{1}{2}(\delta^t y_{-\frac{1}{2}}+\delta^t y_{\frac{1}{2}})$.

The Newton-Stirling central difference formula uses only quantities lying on the horizontal line $\epsilon=0$. These quantities, which involve the arithmetic averages $\mu\delta^t y_0$ for $t=$ odd integer,[7] are listed in Table II (i.e., 8.61, 8.925, 4.83, 25.095, and 49.35). To obtain any polynomial up to the eighth degree in power-series form from such a set of central differences we use the following equation.[8]

$$u_\epsilon=u_0+\epsilon\left\{\mu\delta u_0-\frac{\mu\delta^3u_0}{6}+\frac{\mu\delta^5u_0}{30}-\frac{\mu\delta^7u_0}{140}\right\}$$
$$+\epsilon^2\left\{\frac{\delta^2u_0}{2}-\frac{\delta^4u_0}{24}+\frac{\delta^6u_0}{180}-\frac{\delta^8u_0}{1120}\right\}$$
$$+\epsilon^3\left\{\frac{\mu\delta^3u_0}{6}-\frac{\mu\delta^5u_0}{24}+\frac{7\mu\delta^7u_0}{720}\right\}$$
$$+\epsilon^4\left\{\frac{\delta^4u_0}{24}-\frac{\delta^6u_0}{144}+\frac{7\delta^8u_0}{5760}\right\}$$
$$+\epsilon^5\left\{\frac{\mu\delta^5u_0}{120}-\frac{\mu\delta^7u_0}{360}\right\}+\epsilon^6\left\{\frac{\delta^6u_0}{720}-\frac{\delta^8u_0}{2880}\right\}$$
$$+\epsilon^7\left\{\frac{\mu\delta^7u_0}{5040}\right\}+\epsilon^8\left\{\frac{\delta^8u_0}{40320}\right\}. \quad (1)$$

For the set of five observations given, we have, accordingly,

TABLE II. Set of data and differences for illustrative example.

| $x$ | $\epsilon$ | $y$ | $\delta y$ | $\delta^2 y$ | $\delta^3 y$ | $\delta^4 y$ |
|---|---|---|---|---|---|---|
| 0 | $-2$ | 0 | | | | |
| | | | 2.10 | | | |
| 5 | $-1$ | 2.10 | | 4.41 | | |
| | | | 6.51 | | 0.42 | |
| 10 | 0 | 8.61 | (8.925) | 4.83 | (25.095) | 49.35 |
| | | | 11.34 | | 49.77 | |
| 15 | $+1$ | 19.95 | | 54.60 | | |
| | | | 65.94 | | | |
| 20 | $+2$ | 85.89 | | | | |

[7] The Newton-Stirling formula can be used for the case of either an odd or an even number of observations. In the latter case and for a polynomial of degree $j$, where $j$ is odd, the single available difference, $\delta^j y$, is used for the $\mu\delta^j y_0$ demanded by the formula since the difference of order $j$ is constant by assumption.

[8] It is customary, especially among statisticians, to use $u$ for the *calculated* value of the ordinate and $y$ for the *observed* value. These convenient symbols are adopted throughout the present paper. But in the case of an interpolation formula, based on a table of differences, the calculated function passes exactly through each given point ($y_0$, $y_1$, etc.), and hence there is no distinction between such values and the corresponding calculated values ($u_0$, $u_1$, etc.). For that reason it is standard practice to write $u$ in place of $y$ in a table of differences as well as in all interpolation formulas and in equations derived from them, such as Eq. (1). Since, however, we shall later (Table VII) use this same table of differences, extended to seven observations, for a least-squares' solution, we here retain $y$ in the table, even though it is replaced by $u$ in Eq. (1), which is in standard form [see Whittaker and Robinson, *Reduction of Observations* (Blackie and Son, London, England), p. 65].

As just indicated, the value of the abscissa is here shown by a *subscript*, not only for the observations $y$ (as is customary) but also for the calculated function $u$. But elsewhere in the present paper we write $u_j(\epsilon)$ for a polynomial in $\epsilon$, of degree $j$. In other words, in the case of $u$ as well as many other symbols, we reserve the subscript for the degree of the polynomial or some quantity associated with it, and we write the abscissa, when needed, in parenthesis (see also footnote 14).

$$u_\epsilon = 8.61 + \epsilon\left\{8.925 - \frac{25.095}{6}\right\} + \epsilon^2\left\{\frac{4.83}{2} - \frac{49.35}{24}\right\} + \epsilon^3\left\{\frac{25.095}{6}\right\} + \epsilon^4\left\{\frac{49.35}{24}\right\},$$

or

$$u_\epsilon = 8.61 + 4.7425\epsilon + 0.35875\epsilon^2 + 4.1825\epsilon^3 + 2.05625\epsilon^4. \quad (2)$$

To change from a $f(\epsilon)$ to a $f(e)$, where $e = h\epsilon$, we divide the coefficient of $\epsilon^i$ by $h^i = 5^i$, and substitute $e$ for $\epsilon$, thus getting

$$u_e = 8.61 + 0.9485e + 0.01435e^2 + 0.03346e^3 + 0.00329e^4.$$

To obtain the final equation in $x$, we perform a Horner shift (synthetic division) of amount $-x_0 = -10$. In order to make all directions explicit, we give this shift in full.

| | | | | |
|---|---|---|---|---|
| 0.00329 | 0.03346<br>−0.03290 | 0.01435<br>−0.00560 | 0.9485<br>−0.0875 | 8.61<br>−8.61 |
| 0.00329 | +0.00056<br>−0.03290 | +0.00875<br>+0.32340 | +0.8610<br>−3.3215 | 0 |
| 0.00329 | −0.03234<br>−0.03290 | +0.33215<br>+0.65240 | −2.4605 | |
| 0.00329 | −0.06524<br>−0.03290 | +0.98455 | | |
| 0.00329 | −0.09814 | | | |

Hence,

$$u_x = -2.4605x + 0.98455x^2 - 0.09814x^3 + 0.00329x^4. \quad (3)$$

I believe that the foregoing process is the most rapid and convenient one for setting up a polynomial in power-series form which exactly satisfies a given set of equally spaced data.

## C. LEAST-SQUARES' POLYNOMIAL FITTING OF DATA IN TERMS OF OBSERVATIONS

### C1. Introduction. Various Forms of Solution

As noted in the Introduction, numerous papers have been published on the least-squares' fitting of equally spaced data. Many different forms of solution are possible, and what constitutes the most advantageous form depends in a major way on the use to be made of the information. Thus the final objectives of the statistician and of the physical scientist are often quite different, and for just this reason the procedures advocated by statisticians may often not be the most advantageous for physical scientists. The primary purpose of the Birge and Shea paper[3] was to present a method for obtaining, by least squares, a polynomial in *power-series form*. I believe that the method presented there still remains the best for this particular purpose.

On the other hand, if one desires merely to *smooth* the data (i.e., to obtain calculated or *adjusted* values of each point) and has no interest in the actual function that furnishes such values, then a certain modification of the Birge and Shea method is undoubtedly more convenient. Furthermore, if one desires the probable error of any one of the quantities evaluated (coefficients, calculated points, etc.) then a further modification of the Birge and Shea method is desirable. The method about to be presented represents what now appears to the writer the one most advantageous for the general use of physical scientists. Its relation to the original Birge and Shea method will be indicated as we proceed.

The new method involves in part the use of certain results already existing in the literature at the time the Birge and Shea work was done but of which we were then quite unaware, as already noted. In fact, the entire subject of *orthogonal polynomials*, of which so much use is made in this paper, goes back to Tchebycheff[9] who in 1859 applied such polynomials to the particular problem of concern here. His derivations were, however, very involved, and his results have since been reproduced in a far simpler manner by J. W. Weinberg. In fact, Weinberg has independently derived all the necessary equations by direct algebraic methods, including some results that are quite new. An outline of these results and of the method of derivation is given by Weinberg in an appendix (Section J) to the present paper. Accordingly, I shall in general merely quote such results and refer to the appendix for their derivation.

Needless to say, the present article would never have been written, had it not been for the invaluable aid that I have received from Dr. Weinberg. As a result of his own work, as well

[9] P. L. Tchebycheff (1854 to 1875). See *Oeuvres* (1899), Vol. **1**, pp. 203–230, 381–384, 473–498, 541–560, 701–702; Vol. **2**, pp. 219–242.

as his ability to read and interpret correctly some of the involved mathematical papers that exist in the field, it now seems possible to present all of the necessary material in a fairly simple and straightforward manner. As already noted, it is most unfortunate that so much valuable mathematical work lies virtually unnoticed in the literature just because of the difficulty found by most potential users of the material in understanding and properly evaluating the practical possibilities of the work. In fact, the relation of the physical scientist to the mathematician is much like that of the engineer to the physicist. To discover a result is one thing, to formulate it in such a way as to make it of obvious practical value constitutes quite a different, but equally important problem.

The least-squares' solution of a polynomial may be given in ($A$) power-series form, or ($B$) factorial form. The solution may involve the explicit use of (1) power moments of the observations, (2) factorial moments of the observations, (3) the observations themselves, and (4) the various finite differences of the observations. Finally, the solution may be in ($\alpha$) orthogonal polynomial form, ($\beta$) non-orthogonal polynomial form. The method presented by Birge and Shea,[3] and its various modifications as given here, are then properly labeled $A3\alpha$. The $A4\alpha$ method also is treated briefly in Section D. The method given by Kerawala[10] is $A3\beta$, that by H. T. Davis[11] is $A1\beta$, that advocated by Sasuly[12] is $B2\alpha$, etc. It is obvious that numerous additional forms of solution are possible although some would not be very logical. Thus if the polynomial is expressed in factorial form, it is logical to use factorial moments rather than power moments, etc.

Of the preceding sets of alternatives, the last ($\alpha$ *versus* $\beta$) is by far the most important. There are very great advantages resulting from the use of an orthogonal form of solution, and just this fact makes the methods given by Kerawala and Davis of relatively little value. There are also serious objections to Davis' method on the score of accuracy, as will be explained in Section G.

## C2. Solution as Sum of Orthogonal Polynomials

Two functions $T_j$ and $T_k$ are orthogonal to each other over a specified interval of $x$ if they satisfy the relations

$$\sum_x T_j T_k = 0 \ (k \neq j), \quad \sum_x [T_j]^2 = M_j. \qquad (4)$$

The role of the orthogonality property in the present problem is shown more clearly by the following treatment, which follows that of p. 34 of the reference in footnote 12.

We wish to express the least-squares' polynomial of degree $j$ as the *sum of a series of orthogonal polynomials* of degree $t=0$ to $j$. Then by merely adding or subtracting terms of such a sum, one can pass directly from the least-squares' solution of any degree to the corresponding solution of any other desired degree. This last possibility constitutes the fundamental advantage of the orthogonal form. Thus let $T_t$ represent an orthogonal polynomial of degree $t$, satisfying Eq. (4), and let us multiply it by $a_t$. The $a_t$ is to be a function of the ordinates $y$ but *not* of the abscissae ($x$ or $\epsilon$), and $T_t$ is to be a function of the abscissae ($x$ or $\epsilon$) but *not* of the ordinates $y$. It may be shown that for the particular form of $T_t$ about to be listed (see Table III), $a_t$ is actually the coefficient of $\epsilon^t$ in the least-squares' polynomial of degree $t$, and hence this symbol[13] is appropriate.

Now we know that the least-squares' polynomial $a_0T_0$ of degree zero is merely the arithmetic average $\sum y/n$ of the observations. This polynomial may also be designated[14] $u_0(\epsilon)$. It is thus evident that $a_0 = \sum y/n$ and $T_0 = 1$. Simi-

[10] S. M. Kerawala, Indian J. Phys. **15**, 241 (1941).

[11] H. T. Davis, *Tables of the Higher Mathematical Functions* (Principia Press, Bloomington, Indiana, 1935), II, 307–385.

[12] Max Sasuly, *Trend Analysis of Statistics* (The Brookings Institution, Washington, D. C., 1934).

[13] Sasuly, footnote 12, uses $K_t$ for this quantity ($a_t$). The $T_t$ is his symbol, which in turn is denoted $\xi_t$ by R. A. Fisher, *Statistical Methods for Research Workers* (Oliver and Boyd, Edinburgh, 1946), 10th edition, p. 147, and by his various co-workers and followers. Note that $a_tT_t$ is also an orthogonal polynomial if $T_t$ is, since $a_t$ is independent of $x$ (or $\epsilon$).

[14] As noted in footnote 8, the *degree* of the polynomial will normally be indicated by a subscript followed by the abscissa scale in parentheses. We shall use either $t$ or $j$ to denote the degree of a polynomial. As just mentioned, a polynomial of degree $j$ may be expressed as a sum of *orthogonal* polynomials of degree $t$, where $t$ varies from 0 to $j$. *In general* we shall use $t$ for the degree of an *orthogonal* polynomial—like $T_t$ or $a_tT_t$—and $j$ for the degree of the final polynomial to which the observational data are fitted, as in Eq. (5).

larly, the first-degree least-squares' polynomial for the same set of data is designated $u_1(\epsilon)$. But the *difference* of a polynomial of degree unity and one of degree zero is necessarily a polynomial of degree unity. The difference, in this case, is just the *desired orthogonal polynomial* of degree unity, and hence denoted $a_1T_1$. Thus

$$u_1(\epsilon)-u_0(\epsilon)\equiv a_1T_1,$$

or

$$u_1(\epsilon)=a_1T_1+u_0(\epsilon)=a_1T_1+a_0T_0.$$

Similarly,

$$u_2(\epsilon)-u_1(\epsilon)\equiv a_2T_2,$$

or

$$u_2(\epsilon)=a_2T_2+u_1(\epsilon)=a_2T_2+a_1T_1+a_0T_0,$$

and, in general,

$$u_j(\epsilon)=\sum_{t=0}^{j}a_tT_t. \tag{5}$$

Equation (5), for the least-squares' solution of a polynomial of degree $j$, was first given in 1859 by Tchebycheff[9] who also derived explicit expressions for $T_t$, or rather for a quantity $P_T$ proportional to our $T_t$ (see Eq. 19), up to $t=5$. Birge and Shea[3] derived explicit expressions for another quantity $P_B$ proportional to $T_t$ (see Eq. (11)), also up to $t=5$, by a completely different method and in entire ignorance of Tchebycheff's earlier work. Miss Allan[15] some years later derived a different form of general expression for $T_t$ and by means of it worked out the explicit forms for $T_t$ up to $t=10$. Weinberg has now found a much simpler method of obtaining such explicit forms by means of a modification of Tchebycheff's recursion formula for $T_t$.

This modification consists of a recursion formula for $R_{kt}$, where $R_{kt}$, a function of $n$ only, is

TABLE III. Expressions for $T_t$, orthogonal polynomials with unit leading coefficients ($V_t$ may be obtained from $T_t$)*

| | | |
|---|---|---|
| $T_0(\epsilon)=1$ | | |
| $T_1(\epsilon)=\epsilon$ | | |
| $T_2(\epsilon)=\epsilon^2+R_{02}$ | where | $R_{02}=-(n^2-1)/12$ |
| $T_3(\epsilon)=\epsilon^3+R_{13}\epsilon$ | where | $R_{13}=-(3n^2-7)/20$ |
| $T_4(\epsilon)=\epsilon^4+R_{24}\epsilon^2+R_{04}$ | where | $R_{24}=-(3n^2-13)/14$ |
| | | $R_{04}=+3(n^2-1)(n^2-9)/560$ |
| $T_5(\epsilon)=\epsilon^5+R_{35}\epsilon^3+R_{15}\epsilon$ | where | $R_{35}=-5(n^2-7)/18$ |
| | | $R_{15}=+(15n^4-230n^2+407)/1008$ |
| $T_6(\epsilon)=\epsilon^6+R_{46}\epsilon^4+R_{26}\epsilon^2+R_{06}$ | where | $R_{46}=-5(3n^2-31)/44$ |
| | | $R_{26}=+(5n^4-110n^2+329)/176$ |
| | | $R_{06}=-5(n^2-1)(n^2-9)(n^2-25)/14784$ |
| $T_7(\epsilon)=\epsilon^7+R_{57}\epsilon^5+R_{37}\epsilon^3+R_{17}\epsilon$ | where | $R_{57}=-7(3n^2-43)/52$ |
| | | $R_{37}=+7(15n^4-450n^2+2051)/2288$ |
| | | $R_{17}=-(35n^6-1645n^4+17297n^2-27207)/27456$ |
| $T_8(\epsilon)=\epsilon^8+R_{68}\epsilon^6+R_{48}\epsilon^4+R_{28}\epsilon^2+R_{08}$ | where | $R_{68}=-7(n^2-19)/15$ |
| | | $R_{48}=+7(3n^4-118n^2+763)/312$ |
| | | $R_{28}=-(105n^6-6405n^4+91679n^2-231491)/34320$ |
| | | $R_{08}=+7(n^2-1)(n^2-9)(n^2-25)(n^2-49)/329472$ |
| $T_9(\epsilon)=\epsilon^9+R_{79}\epsilon^7+R_{59}\epsilon^5+R_{39}\epsilon^3+R_{19}\epsilon$ | where | $R_{79}=-3(3n^2-73)/17$ |
| | | $R_{59}=+21(3n^4-150n^2+1307)/680$ |
| | | $R_{39}=-(21n^6-1617n^4+30387n^2-112951)/3536$ |
| | | $R_{19}=+3(105n^8-11060n^6+334054n^4-2973140n^2+4370361)/3111680$ |
| $T_{10}(\epsilon)=\epsilon^{10}+R_{8,10}\epsilon^8+R_{6,10}\epsilon^6+R_{4,10}\epsilon^4+R_{2,10}\epsilon^2+R_{0,10}$ | | |
| | where | $R_{8,10}=-15(3n^2-91)/76$ |
| | | $R_{6,10}=+21(15n^4-930n^2+10507)/2584$ |
| | | $R_{4,10}=-5(21n^6-1995n^4+47775n^2-245737)/10336$ |
| | | $R_{2,10}=+3(105n^8-13580n^6+514990n^4-6039260n^2+13782993)/1074944$ |
| | | $R_{0,10}=-63(n^2-1)(n^2-9)(n^2-25)(n^2-49)(n^2-81)/47297536$ |

* Note,—If the $R_{kt}$ coefficients, including $R_{tt}(=1)$, are replaced by the corresponding $S_{kt}$ (values in Table XII), then these same $f(\epsilon)$ give the respective values of $V_t(\epsilon)$. See Eqs. (21), (28) and (77).

[15] F. E. Allan, Proc. Roy. Soc. Edinburgh **50**, 310 (1929–30).

the coefficient of $\epsilon^k$ in the expression for $T_t$, as given in Table III. The new recursion relation is

$$R_{k+1,\,t+1}=R_{kt}-\frac{t^2(n^2-t^2)}{4(4t^2-1)}R_{k+1,\,t-1}. \tag{6}$$

By the use of this relation I have verified all of Allan's expressions. Since such expressions *must* be used in the numerical calculation of *extrapolated* values of $u_j(\epsilon)$, it seems necessary to present them in full. They apply *only* to the special abscissa scale that we have designated by $\epsilon$, where $\epsilon$ runs from $-q$ to $+q$, and $2q+1=n$ $=$number of observations. In fact, it is the symmetry about $\epsilon=0$ that causes the coefficients of alternate powers of $\epsilon$ to become zero, thus greatly simplifying the expressions for $T_t$. Stated more specifically, $R_{kt}$ (and $S_{kt}$) $=0$ if $t-k=$odd integer.

In Table III the coefficient of $\epsilon^t$ in $T_t(\epsilon)$ is really $R_{tt}$, but $T_t$ has been deliberately chosen so as to make $R_{tt}$ equal to unity. Under this condition the coefficient of $T_t$ in Eq. (5) is actually $a_t$, the coefficient of $\epsilon^t$ in the power-series form of polynomial, as already noted. We thus have a complete formulation of the $T_t$ portion of Eq. (5), and by the use of Eq. (6) any desired explicit expressions can be derived for values of $t$ higher than 10.

The remaining problem is the evaluation of $a_t$ in Eq. (5), the quantity which alone involves the actual observations $y$ of a given problem. Tchebycheff[9] derived for $a_t$ the following expression:

$$a_t=\sum_\epsilon y\cdot T_t/\sum_\epsilon T_t^2\equiv\sum_\epsilon y\cdot T_t/M_t. \tag{7}$$

Thus[16] $M_t$ is used to denote $\sum_\epsilon T_t^2$ (compare Eq. (4)).

Putting Eq. (7) into Eq. (5) we get

$$u_j(\epsilon)=\sum_{t=0}^{j}(\sum_\epsilon y\cdot T_t/\sum_\epsilon T_t^2)T_t. \tag{8}$$

Thus Eq. (8), with the explicit values of $T_t$ given in the foregoing table, furnishes the complete least-squares' solution of a polynomial of degree $j$, expressed as a sum of orthogonal polynomials, with the special $\epsilon$ scale of abscissa.

When expressed as a *power series* in $\epsilon$, the function may be written

$$u_j(\epsilon)=\sum_{k=0}^{j}a_{kj}\epsilon^k. \tag{9}$$

Hence $a_{kj}$ is the coefficient of $\epsilon^k$ in the polynomial of degree $j$. It is, in general, necessary to use such a double subscript, since the present method of deriving the least-squares' solution of a polynomial of degree $j$ involves all of the polynomials of lower degree. In this same symbolism, the *final* coefficient in Eq. (9) is $a_{jj}$. Also, in Eq. (5), $a_t$ is more correctly written as $a_{tt}$ since it is the coefficient of $\epsilon^t$ in the polynomial of degree $t$. However, because of the repeated occurrence of $a_{tt}$ (or $a_{jj}$) in the present treatment, it is convenient to simplify the symbol to $a_t$ (or $a_j$).

A comparison of Eqs. (9) and (5), in connection with the explicit expressions for $T_t$ already given, shows that the $a_{kj}$ in Eq. (9) are given by

$$a_{kj}=\sum_{t=k}^{j}R_{kt}a_t. \tag{10}$$

Thus Eqs. (9) and (10), with the explicit values of $R_{kt}$ listed in Table III, lead to the least-squares' solution of a polynomial of degree $j$, in the form of a *power series* in $\epsilon$. These equations are used in the form of solution given by Birge and Shea.[3] Dr. Shea and the writer obtained a

[16] The question of a convenient set of symbols is a very perplexing one to which Dr. Weinberg and I have given a great deal of thought. When it is possible to formulate a certain relation in a variety of ways, as we shall find is the case with Eq. (7), and when tables of numerical values are given in various places in the literature, it is highly desirable that a *different* symbol be used for each different form, in order that there may be no confusion over what form is being tabulated. Thus one might express the series of possible forms of $T_t$ by the symbols $T_t$, $T_t'$, $T_t''$, etc., or by $T_t$, $T_t^*$, $T_t^{**}$, etc. But if a table of one of these alternative forms is being given, as is the case in this paper, so that constant reference to such an alternative form is required, it is obviously desirable to use for it the simplest possible symbol. For that reason we are using *different letters* for the different forms. In general the alphabetic letter immediately or closely *beyond* the original letter is used for the alternative form. Thus $V_t$ represents a certain new form of $T_t$, $N_t$ a new form of $M_t$, $b_t$ of $a_t$, $H_{kt}$ of $G_{kt}$, and $S_{kt}$ of $R_{kt}$. (See Table V). The present adopted symbols frequently differ from those used earlier (footnotes 3 and 4).

Many other differing sets of symbols appear in the literature, but none of them has seemed appropriate to our present purpose. We have avoided the use of Greek letters, in general, merely because such letters do not appear on ordinary typewriters. For that reason we reject the symbols $\xi_t$ and $\xi_t'$ in constant use by Fisher and his co-workers. They correspond to our $T_t$ and $V_t$, respectively.

general formula for $a_t$, corresponding to Eqs. (14), (15), and (16) ahead, as well as explicit expressions for the necessary $R_{kt}$ up to $t=5$. We did not then notice the relations between $a_t$ and $R_{kt}$, as shown by Eq. (7) with the values of $T_t$ given in Table III.

### C3. Alternative Forms of Orthogonal Polynomials

Let us now return to Tchebycheff's fundamental Eq. (8). We note in this equation that $T_t$ occurs twice as a factor in *both* the numerator and the denominator. This simple fact furnishes the *possibility* of an endless variety of equally valid modifications of the equation. Several such modifications have already appeared in the literature, and in most cases the author has apparently been unaware of the previous closely related work of others. It therefore seems desirable to indicate the explicit relations between some of these modifications of Eq. (8).

Thus, because of the fact that $T_t$ appears to the same power in both numerator and denominator of Eq. (8), we can multiply $T_t$ by *any factor F, not a function of* $\epsilon$, without changing the values of $u_j(\epsilon)$. In order to obtain the Birge and Shea equation for $a_t$, we write

$$P_B = F_B \cdot T_t = 2t!T_t/(t!)^2. \qquad (11)$$

Putting Eq. (11) into Eq. (8) we get

$$u_j(\epsilon) = \sum_{t=0}^{j} (\sum_{\epsilon} y \cdot P_B / \sum_{\epsilon} P_B{}^2) P_B = \sum_{t=0}^{j} (\sum_{\epsilon} y \cdot P_B / K_t) T_t, \qquad (12)$$

where, with the incorporation also of $M_t$ from Eq. (7),

$$K_t = \sum_{\epsilon} P_B{}^2 / F_B = F_B \cdot M_t \qquad (13)$$

and, from Eqs. (12) and (5)

$$a_t = \sum_{\epsilon} y \cdot P_B / K_t. \qquad (14)$$

Equation (14) was used by Birge and Shea for calculating $a_t$, and $P_B$ was termed a "pair-factor." No special symbol was used for it. Expressions for $P_B$ and $K_t$ were derived in terms of factorials, but by changing the factorials into binomial coefficient form,[17] it is now found that a common integral factor ($t!$) exists. The new simplified expressions are

$$P_B'' = P_B/t! = 2t!T_t/(t!)^3 = \sum_{s=0}^{t} (-1)^s \binom{q-r}{s} \binom{2q-s}{2q-t} \binom{t+s}{s}, \qquad (15)$$

$$K_t'' = K_t/t! = 2t!M_t/(t!)^3 = t!\binom{2q+t+1}{2t+1}, \qquad (16)$$

in which[18] $r=|\epsilon|$, and as usual $2q+1=n=$ number of observations. Weinberg has derived Eqs. (15) and (16) in a far simpler way than that used earlier by Birge and Shea (see Section J). These equations contain, implicitly, the analytic expressions for $T_t$ and $M_t$ of Eq. (7).

A second modification of Eq. (8) is obtained by writing

$$P_M = F_M \cdot T_t = (-1)^t \frac{2t!(n-1-t)!T_t}{(t!)^2(n-1)!}. \qquad (17)$$

Then $P_M$ represents the *Legendre polynomial for discrete points*, and it has been tabulated and used by W. E. Milne[19] for the solution of the present problem.

---

[17] The binomial coefficient $\binom{m}{r}$ is defined as $m!/r!(m-r)!$. An alternative statement is $\binom{a}{b}=\binom{a}{c}$ $=a!/b!c!$, if $b+c=a$ and if all three letters represent positive integers. One advantage of using a binomial coefficient, wherever possible, is that it is necessarily an integer.

[18] The use of this new symbol $r$ is directly connected with the fact that in the original Birge and Shea method, as well as in its new modified form, two observations corresponding to equal positive and negative $\epsilon$ value form a "pair," which is to be multiplied by a common "pair-factor." Each such *pair* of observations is then designated by the value of $r$. We shall also use $r$ for probable error, but there can be no possible resulting confusion.

[19] Private communication from his student, George Pomeroy. Milne actually uses a form of $P_M$ applying to the range $x=0$ to $n-1$, whereas Eq. (17) applies to the range $\epsilon=-q$ to $+q$. Milne's own form is

$$P_M = \sum_{s=0}^{t} (-1)^s \binom{t}{s} \binom{t+s}{s} \frac{x^{(s)}}{(n-1)^{(s)}},$$

where $x^{(s)} \equiv x!/(x-s)!$ and $(n-1)^{(s)} \equiv (n-1)!/(n-s-1)!$ If $x^{(s)}$ is replaced by $x^s$ and $(n-1)^{(s)}$ is deleted, one gets an expression for one form of the Legendre polynomials, in which $x$ varies continuously.

The first practical use of Tchebycheff's orthogonal polynomials was made by Jordan[20] in a paper of fundamental importance. In this paper he carried out in detail the application of such polynomials to the case of uniformly spaced data. In connection with this work he calculated a table of values, not of $T_t$, as defined here, but of

$$P_J = F_J \cdot T_t = 2t!T_t/(t!)^2(2^t). \tag{18}$$

His table covers the range $t$ up to 5, and $n$ up to 20. It is undoubtedly the first table of such "pair-factors" to be published, but Dr. Shea and I in 1924 were unaware of its existence, and everyone else who has since published analogous tables, even up to the present time, seems to have been equally unaware of its existence.

Finally, Tchebycheff gave his orthogonal polynomials in the form[21]

$$P_T = F_T \cdot T_t = 2t!T_t/t!. \tag{19}$$

As already noted, the particular form $T_t$, as used by Sasuly[12] and others, has been so chosen that the coefficient of $\epsilon^t$ is unity (Table III).

### C4. The *a*-*T* and *b*-*V* Systems of Calculation

After this rather long digression on the various ways in which Eq. (8) can be expressed and has been expressed, we take up the question of the best actual use of this fundamental equation. It will be noted that Eq. (14), as used by Birge and Shea, gives the most direct calculation of $a_t$ for use in Eq. (5). But, as shown by Eqs. (12) and (13), $K_t$ is *not* the sum of the squares of the quantity $P_B$ that appears in the numerator of Eq. (12). On the contrary, it is that sum divided by $F_B$ (defined in Eq. (11)).

Now our recent study of probable errors has shown that it is most important to preserve the *form* of Eq. (8), regardless of the factor used. Thus let us write, as an entirely *general* expression,

$$u_j(\epsilon) = \sum_{t=0}^{j} \left(\sum_\epsilon y \cdot P / \sum_\epsilon P^2\right) P, \tag{20}$$

where $P = F \cdot T_t$ and $F$ is any factor not a function of $\epsilon$. Let the quantity inside the parentheses be denoted by $a_t'$. Then it is found that $\sum_\epsilon P^2$ is the *weight* of $a_t'$, to be used in getting the probable error of $a_t'$. If $F=1$, then $a_t' = a_t$ and $\sum_\epsilon P^2 = \sum_\epsilon T_t^2 = M_t =$ *weight* of $a_t$ on the basis of unit weight assigned to each observation $y$ (see Section J). Thus, if one desires to calculate probable errors, it is most convenient to use the *form* of Eq. (20). Furthermore, Eq. (20) is most convenient for calculating the value of $u_j(\epsilon)$ for any single value of $\epsilon$. In the Birge and Shea method Eq. (10) was used to get the coefficients $a_{kj}$ of the polynomial as a *power* series in $\epsilon$, and then any value of the polynomial was calculated by direct substitution in this power series. But Eq. (20) is more convenient for this purpose, after the necessary constants $a_t'$ have been evaluated.

The change from the use of Eqs. (14), (10), and (9) to an expression of the form of Eq. (20) constitutes the essential modification now presented of the original Birge and Shea method. The simplest way of getting from the values of $a_t'$ in Eq. (20) to the values of the coefficients of any desired power series (in either $\epsilon$ *or* in the original $x$) will be discussed presently.

The rapid and efficient use of Eq. (20) for the calculation of values of $u_j(\epsilon)$ requires tables of values of $P$ and of $\sum P^2$ for the desired values of $t$ and $n$. It is obvious that for such calculations integer values of $P$ (and hence also of $\sum P^2$) are highly desirable. Let us therefore examine, with this idea in mind, the various expressions already given for $P$.

In the first place, the fact that $P_B''$ of Eq. (15) consists of a product of binomial coefficients shows that it is always an integer. Then $P_B$ of Eq. (11) is also an integer, but larger by a factor $t!$ than $P_B''$. The $P_T$ of Eq. (19) is also an integer but larger than $P_B''$ by a factor $(t!)^2$.

On the other hand, most of the values of $T_t$ are non-integers, as can be seen easily by examining the explicit expressions listed in Table III. Nearly one-half of the values of $P_J$ of Eq. (18), as given in Jordan's table,[20] are also non-integers, and from Eq. (17) most of the values of $P_M$ must also be non-integers, especially in the case of larger values of $n$.

[20] Charles Jordan, Proc. London Math. Soc. (2) **20**, 297 (1921). This paper includes a brief description of Tchebycheff's original work. We are indebted to Dr. E. U. Condon for calling our attention to Jordan's contributions. (Note that Charles Jordan and Karl Jordan are the same person.)

[21] See the footnote p. 34 in Sasuly's book (footnote 12) or p. 299 of Jordan's paper (footnote 20).

As a matter of fact, $P_B''$ represents the *smallest* possible set of integers valid for a given value of $t$ and for all values of $n$. On the other hand, for a *particular* value of $n$ and of $t$ there is usually a particular factor[22] that can be removed from the entire set of values of $P_B''$. The values of the new integral set of "pair-factors" thus obtained are denoted by the symbol $V_t$. As will appear in a moment, it is necessary in calculating the coefficients $a_{kj}$ of Eq. (10) to know the relation of the $V_t$ values, not to $P_B''$ or $P_B$, but to the more fundamental $T_t$ of Eqs. (5) and (8). For this relation Fisher and his co-workers use the symbol $\lambda$, but the logical symbol for this paper, as shown in Eq. (28) ahead, is $S_{tt}$. Thus, by definition,

$$V_t \equiv \lambda_t T_t \equiv S_{tt} T_t. \tag{21}$$

We now replace $T_t$ by $V_t$ in the fundamental Eq. (8) and thus get, with the use of Eqs. (7) and (21),

$$u_j(\epsilon) = \sum_{t=0}^{j} (\sum_\epsilon y \cdot V_t / \sum_\epsilon V_t^2) V_t \equiv \sum_{t=0}^{j} b_t V_t. \tag{22}$$

Thus, by definition,

$$b_t = \sum_\epsilon y \cdot V_t / \sum_\epsilon V_t^2 \equiv \sum_\epsilon y \cdot V_t / N_t = a_t / S_{tt}. \tag{23}$$

In Eq. (23), $N_t$ is the "weight" of $b_t$ just as $M_t$ in Eqs. (7), (13), and (16) is the weight of $a_t$, as already noted.

Hence Eq. (5) may be replaced, for purposes of numerical calculation, by Eq. (22), as a result of the important relation, which follows directly from Eqs. (21) and (23),

$$a_t T_t = b_t V_t. \tag{24}$$

If we desire to know the calculated values of the polynomial at either the observed values of $\epsilon$ or at any extrapolated values, and also the probable errors of the function at such values of $\epsilon$—but *not* the coefficients of the polynomial in power-series form—then it is possible to carry through all the necessary calculations in terms of (1) values of $a_t$ and $T_t$ *or* (2) values of $b_t$ and $V_t$. *We shall denote the first as the a–T system of calculation and the second as the b–V system.*

For each symbol in the $a$–$T$ system there is a "corresponding" symbol in the $b$–$V$ system. Thus $N_t(=\sum_\epsilon V_t^2)$ corresponds to $M_t(=\sum_\epsilon T_t^2)$. Full details are shown in Table V ahead. The basis for the choice of the corresponding *pairs* of symbols is stated in footnote 16. The factor connecting each such pair of symbols is $S_{tt}$ or some power of it. As already noted, the values of $S_{tt}$ have been so chosen as to give integers, and the *smallest possible* integers, for the quantities of most common occurrence in the calculations of the $b$–$V$ system. That system, therefore, possesses a very great advantage over the $a$–$T$ system in which the major portion of the corresponding quantities are non-integers.

In the Birge and Shea paper the values of the so-called "pair-factors," listed in an extensive table, are just the values of $V_t$. But instead of calculating $b_t$ by means of Eq. (23), Birge and Shea calculated $a_t$ by means of Eq. (14), except that all common integral factors were first removed from numerator and denominator, giving

$$a_t = \sum_\epsilon y \cdot V_t / K_t^*, \tag{25}$$

where

$$K_t^* = \sum_\epsilon V_t^2 / S_{tt} = N_t / S_{tt}. \tag{26}$$

The values of $K_t^*$ were listed by Birge and Shea (and denoted $K$) in the same table as those of $V_t$. In the *modified* Birge and Shea method now being presented *we keep all calculations in the b–V system* in order to utilize the important advantage of that system which has just been mentioned. Hence we again list (in Table XIII) the values of $V_t$, but in place of $K_t^*$ we now list the values of $N_t$, which are related to those of $K_t^*$ by Eq. (26). For convenience, we have calculated the listed values of $N_t$ from Eq. (26), using the values of $S_{tt}$ listed in Table XII. But, in principle, the values of $N_t$ are to be obtained from the more basic $M_t$ as expressed by Eq. (16), and with the use of Eq. (31). One then employs the listed values of $V_t$ and $N_t$ (Table XIII) in Eq. (23) to obtain $b_t$.

If now one desires the values of $a_{kj}$, defined by Eq. (9), or the values of the corresponding coefficients of the power series in $x$, it becomes neces-

[22] Let us denote this factor by $f$ so that $P_B''/f \equiv V_t$ of Eq. (21). Then from Eqs. (15) and (21) we get $f \cdot S_{tt} = 2t!/(t!)^3$. As will appear later, this is a very valuable relation, connecting the $S_{tt}$ values listed in Table XII of Section I with the $f$ values listed in Table VI of Section D. Further details appear in Section D.

sary to formulate a new equation in the $b-V$ system, "corresponding" to Eq. (10) of the $a-T$ system, which was used by Birge and Shea. This new equation is

$$a_{kj}=\sum_{t=k}^{j} S_{kt}\cdot b_t, \tag{27}$$

where

$$S_{kt}=S_{tt}\cdot R_{kt}. \tag{28}$$

We now have the basis for the choice of symbol $S_{tt}$. As already stated, the various expressions for $T_t$ listed in Table III actually start with the term $R_{tt}\cdot\epsilon^t$, but $T_t$ is so defined that all $R_{tt}$ equal unity. This condition on $R_{tt}$ is satisfied by Eq. (28) by putting $k=t$. In the present paper we list values of $S_{kt}$ (Table XII) *in place of* the table of $R_{kt}$ values published by Birge and Shea. In the new table the values of $S_{tt}$ are necessarily included. They are the basic numbers connecting the $a-T$ and $b-V$ systems. Aside from these values of $S_{tt}$, the contents of Table XII have not before appeared in print.

Furthermore, since $S_{kt}$ "corresponds" to $R_{kt}$ and $V_t$ to $T_t$; if we substitute $S_{kt}$ for $R_{kt}$ in Table III, we will get explicit expressions for $V_t$ as a $f(\epsilon)$. Hence, $S_{kt}$ is best defined as the coefficient of $\epsilon^k$ in the expression for $V_t$ as $f(\epsilon, n)$.

In 1927, Birge and Shea published a table of values of $V_t$ for $t=1$ to 5, and for $n$ up to 30. So far as I knew until a short time ago, this was the first table of such values to be published. As already noted, the table by Jordan,[20] published in 1921, gives values of $P_J$, Eq. (18), most of which are not integers. But quite recently Dr. V. A. Nekrassoff has informed me that a table similar to that of Birge and Shea appears in a book by Khotimsky[23] published in 1925. I have not yet seen the book.

Then, in 1938, a new table of $V_t$ values was published by Fisher and Yates[24] covering the region $t=1$ to 5 and $n$ up to 52. The method presented by these authors is based upon the use of Eq. (22). Their work, in turn, seems to be based essentially on two earlier papers by Fisher.[25,26] In the earlier of these two papers Fisher gives the results of what appears to be an independent derivation of Tchebycheff's Eq. (5). In neither of these papers nor in Fisher's well-known book[13] is there any mention of Tchebycheff's earlier work on orthogonal polynomials. In fact, in his first paper Fisher names such expressions "uncorrelated polynomials" which, to the statistician, is a much more revealing designation.

Finally, in 1942, Anderson and Houseman[27] extended the Fisher and Yates' table to $n=104$. Their table is thus by far the most extensive now in print and their paper includes a detailed account of the use of Eq. (22) with an illustrative problem. But the much older table of Birge and Shea is not mentioned in any of these papers.

To summarize the discussion thus far: By means of the table of values of $V_t$ and

$$\sum_{\epsilon} V_t^2(=N_t),$$

(Table XIII), and Eqs. (22) and (23), one can obtain with amazing speed and accuracy the calculated values of the least-squares' polynomial $u_j(\epsilon)$ for each value of $\epsilon$, from $-q$ to $+q$, i.e., at abscissa points corresponding to the given data. We can also calculate any *extrapolated* value of $u_j(\epsilon)$ but here the process takes much longer. We still use Eqs. (22) and (23), but now we have available no numerical value of $V_t$ in Table XIII. Hence, we are forced to calculate $V_t$ from the analytic expressions of Table III, with the $R_{kt}$ replaced by $S_{kt}$ and with the numerical values of $S_{kt}$ from Table XII, for the appropriate value of $n$. In this connection one must remember to multiply each $\epsilon^t$ term in $T_t$ by $S_{tt}$, which thus replaces the implicit $R_{tt}(=1)$ of Table III.

Furthermore, we can calculate the values of $a_{kj}$, the coefficients of $u_j(\epsilon)$ expressed as a *power series* in $\epsilon$, by the use of Eq. (27). Then the most rapid way to obtain the coefficients of $u_j(e)$, where $e=\epsilon h$, is undoubtedly to divide each $a_{kj}$ coefficient by $h^k$ and replace $\epsilon$ by $e$, exactly as has been done in the problem of Section B. The

[23] V. Khotimsky, *Graduation of Statistical Series by Least Squares* (Moscow and Leningrad, 1925), in Russian. This publication is mentioned also by A. C. Aitken, Proc. Roy. Soc. Edinburgh **54**, 1 (1933).

[24] R. A. Fisher and F. Yates, *Statistical Tables for Biological, Agricultural and Medical Research* (Oliver and Boyd, London, 1938), see Table XXIII.

[25] R. A. Fisher, J. Agric. Sci. **11**, 107 (1921).

[26] R. A. Fisher, Phil. Trans. Roy. Soc. London **B213**, 89 (1924).

[27] R. L. Anderson and E. E. Houseman, Research Bulletin 297, Agric. Exp. Station, Iowa State College, 1942.

coefficients of $u_j(x)$ can then be obtained by a Horner shift of $-x_0$, where $e=x-x_0$, as also carried out in Section B. But this method, while the most rapid, does not furnish *explicitly* the quantities needed to get the *probable errors* of the new coefficients. Hence a method suitable for this latter purpose will be given in a moment.

With the values of $b_t$ obtained from Eq. (23) one can immediately calculate $\sum_\epsilon [v_j(\epsilon)]^2$, where $v_j(\epsilon)=y_\epsilon-u_j(\epsilon)=$ residual of any observation $y_\epsilon$. The necessary equation (derived in Section J) is

$$\sum_\epsilon [v_j(\epsilon)]^2=\sum_\epsilon y^2-\sum_{t=0}^{j} M_t a_t^2, \tag{29}$$

where $M_t=\sum_\epsilon T_t^2=$ statistical *weight* of $a_t$, in Eq. (7), as already noted. Eq. (29), which is a type fundamental in analysis of orthogonal functions, was first applied by Jordan[20] to the problem under discussion. It is given by Anderson and Houseman[27] without reference to source, and it may have been derived independently by R. A. Fisher. It is also derived in Section J.

The "corresponding" equation in the $b$–$V$ system is

$$\sum_\epsilon [v_j(\epsilon)]^2=\sum_\epsilon y^2-\sum_{t=0}^{j} N_t b_t^2, \tag{30}$$

where $N_t=\sum_\epsilon V_t^2=$ weight of $b_t$ in Eq. (23), as already noted. Values of $N_t$ are listed in Table XIII together with the $V_t$ values. A typical series of values of $\sum_\epsilon [v_j(\epsilon)]^2$ for $j=0$ to 5, as yielded by the difference of $\sum_\epsilon y^2$ and the successive sums $N_0b_0^2$, $N_0b_0^2+N_1b_1^2$, $N_0b_0^2+N_1b_1^2+N_2b_2^2$, etc., has been given in Section A. The great advantage of Eq. (30), resulting from its orthogonal form, is that in the process of calculating the desired $\sum_\epsilon [v_j(\epsilon)]^2$ for the $j$th degree polynomial, we incidentally discover just the effect on the value of $\sum v^2$ of *each* least-squares' polynomial of degree $t=0$ to $j$. This fact has been pointed out in Section A.

### C5. Coefficients, Weights, and Probable Errors

We now consider the important subject of the *probable error* of each of the quantities of interest to us. We note, in the first place, that $M_t(=\sum_\epsilon T_t^2)$ is the statistical weight of $a_t$. This fact is proved in Section J, but it has presumably been proved earlier by others and it is used by Fisher and Yates.[24] The "corresponding" weight of $b_t$ is then $N_t(=\sum_\epsilon V_t^2)$, and it follows from Eq. (21) that

$$N_t=S_{tt}^2\cdot M_t. \tag{31}$$

Now if $Z$ is a function of $z_1$, $z_2$, etc., where $z_1, z_2, \cdots$ are *independently observed* quantities, and if $p_1, p_2, \cdots$ are the assigned *weights* of $z_1, z_2, \cdots$, then the *resulting weight* $p_Z$ of $Z$, as given by the law of propagation of errors, is

$$1/p_Z=\sum_i (\partial Z/\partial z_i)^2/p_i. \tag{32}$$

In connection with the problem now being discussed, the orthogonal character of the solution is again of major importance since, as proved in Section J, the quantities $b_t$ *(or $a_t$) act like independently observed quantities.* It is believed that this proof is new. Thus in the case of *any* quantity that can be written as an explicit function of the $b_t$ or the $a_t$, we can immediately obtain an expression for its "weight" by applying propagation of errors.

Then, knowing the weight $p$, one gets the probable error

$$r_p=r'/\sqrt{p}, \tag{33}$$

where $r'=$ probable error of a hypothetical quantity of *unit* weight. It is calculated from

$$r'=0.6745[\sum v^2/(n-s)]^{\frac{1}{2}}. \tag{34}$$

This equation applies to the case of $n$ unweighted observations fitted to a function of $s$ undetermined constants. For a polynomial of degree $j$ we have $s=j+1$.

Thus, from Eqs. (5) and (32), the *weight* $p_j(\epsilon)$ of the function $u_j(\epsilon)$ at any specified value of $\epsilon$, is given by

$$1/p_j(\epsilon)=\sum_{t=0}^{j} (T_t^2/M_t), \tag{35}$$

since $M_t(=\sum_\epsilon T_t^2)$ is the weight of $a_t$ in Eq. (5). Similarly, in the $b$–$V$ system, from Eqs. (22) and (32)

$$1/p_j(\epsilon)=\sum_{t=0}^{j} (V_t^2/N_t), \tag{36}$$

since $N_t(=\sum_\epsilon V_t^2)$ is the weight of $b_t$ in Eq. (22). It is Eq. (36) that is to be used in actual calculations since values of both $V_t$ and $N_t$ are listed in Table XIII.

Next, the weights $p_{kj}$ of the coefficients $a_{kj}$ of Eq. (9) are given, from Eqs. (10) and (32), as

$$1/p_{kj}=\sum_{t=k}^{j} (R_{kt}^2/M_t), \quad (37)$$

or in the $b-V$ system, from Eqs. (27) and (32), as

$$1/p_{kj}=\sum_{t=k}^{j} (S_{kt}^2/N_t). \quad (38)$$

We shall now use $a_{kj}'$ for the coefficients of the power series in $e(=\epsilon h)$, and $a_{kj}''$ for those of the power series in $x(=\epsilon h+x_0)$. Thus

$$u_j(x)=\sum_{k=0}^{j} a_{kj}''x^k, \quad (39)$$

where $x$ gives the true abscissa value at which the ordinate $y$ is observed, and $u_j(x)$ is the corresponding calculated ordinate.

In order to get the weights of the coefficients $a_{kj}''$, we must first express such coefficients explicitly as functions of $a_t$ or of $b_t$ and then apply propagation of errors. It is for just this reason that the method used in getting $u_e$ and $u_x$ in Section B is not suitable if one wishes to calculate the probable errors of such coefficients. The conversion of a $f(\epsilon)$ to a $f(x)$, where $\epsilon=(x-x_0)/h$, keeping all relations in literal form, is quite simple and direct. The writing of the resulting equations is, however, simplified by replacing the oft-occurring $x_0$ by $m$. Thus $m(=x_0)$ is that value of $x$ for which $\epsilon=0$. In the Horner shift from $f(x-m)$ to $f(x)$, the amount of the shift is $-m$.

We now give the explicit expressions to be used in calculating the $a_{kj}''$ coefficients of the final $f(x)$ in power-series form and their weights $p_{kj}''$. As usual, such expressions may be formulated in either the $a-T$ system or the $b-V$ system.

In the $a-T$ system

$$a_{kj}''=\sum_{t=k}^{j} G_{kt}a_t, \quad (40)$$

and

$$1/p_{kj}''=\sum_{t=k}^{j} (G_{kt}^2/M_t). \quad (41)$$

In the $b-V$ system

$$a_{kj}''=\sum_{t=k}^{j} H_{kt}b_t, \quad (42)$$

and

$$1/p_{kj}''=\sum_{t=k}^{j} (H_{kt}^2/N_t). \quad (43)$$

Numerical values of $N_t(=\sum_\epsilon V_t^2)$ appear in Table XIII. It seems desirable to list explicit expressions for $H_{kt}(=S_{tt}\cdot G_{kt})$ in detail, since they are to be used in actual calculations, and the present method of obtaining $a_{kj}''$ with the aid of a table of $S_{kt}$ values has not been presented before. If in the expressions which follow (Table IV) one merely replaces $S_{kt}$ by $R_{kt}$, then $H_{kt}$ becomes $G_{kt}$ of Eq. (40) in the $a-T$ system. It is therefore unnecessary to list explicit expressions for $G_{kt}$.

Table IV contains all expressions needed in

TABLE IV. Expressions for $a_{kj}''$ and $H_{kt}$. ($a_{kj}''$=coefficient of $x^k$ in a power series of degree $j$, for values of $x$ spaced at intervals of $h$, where $m$ = "middle" value of $x$, i.e., where $\epsilon=0$. Numerical values of $S_{kt}$ listed in Table XII.)

| | | |
|---|---|---|
| | $a_{0j}''=H_{00}b_0+H_{01}b_1+\cdots H_{0j}b_j$ | (42:0) |
| where | $H_{00}=S_{00}=1$ | |
| | $H_{01}=-S_{11}(m/h)$ | |
| | $H_{02}=S_{02}+S_{22}(m/h)^2$ | |
| | $H_{03}=-[S_{13}(m/h)+S_{33}(m/h)^3]$ | |
| | $H_{04}=S_{04}+S_{24}(m/h)^2+S_{44}(m/h)^4$ | |
| | $H_{05}=-[S_{15}(m/h)+S_{35}(m/h)^3+S_{55}(m/h)^5]$ | |
| | $a_{1j}''=H_{11}b_1+H_{12}b_2+\cdots H_{1j}b_j$ | (42:1) |
| where | $H_{11}=S_{11}/h$ | |
| | $H_{12}=-S_{22}(2m/h^2)$ | |
| | $H_{13}=S_{13}/h+S_{33}(3m^2/h^3)$ | |
| | $H_{14}=-[S_{24}(2m/h^2)+S_{44}(4m^3/h^4)]$ | |
| | $H_{15}=S_{35}(3m^2/h^3)+S_{55}(5m^4/h^5)$ | |
| | $a_{2j}''=H_{22}b_2+H_{23}b_3+\cdots H_{2j}b_j$ | (42:2) |
| where | $H_{22}=S_{22}/h^2$ | |
| | $H_{23}=-S_{33}(3m/h^3)$ | |
| | $H_{24}=S_{24}/h^2+S_{44}(6m^2/h^4)$ | |
| | $H_{25}=-[S_{35}(3m/h^3)+S_{55}(10m^3/h^5)]$ | |
| | $a_{3j}''=H_{33}b_3+H_{34}b_4+\cdots H_{3j}b_j$ | (42:3) |
| where | $H_{33}=S_{33}/h^3$ | |
| | $H_{34}=-S_{44}(4m/h^4)$ | |
| | $H_{35}=S_{35}/h^3+S_{55}(10m^2/h^5)$ | |
| | $a_{4j}''=H_{44}b_4+H_{45}b_5+\cdots$ | (42:4) |
| where | $H_{44}=S_{44}/h^4$ | |
| | $H_{45}=-S_{55}(5m/h^5)$ | |
| | $a_{55}''=H_{55}b_{55}$ | (42:5) |
| where | $H_{55}=S_{55}/h^5$ | |

Eqs. (42) and (43) for polynomials up to $j=5$. In order to make Eq. (42) more explicit, its special forms are given and designated (42:0) to (42:5) for $k=0$ to 5, followed by the actual expressions for $H_{kt}$ needed in each such form.

If, in Table IV, $m=0$ and $h=1$, so that $f(x)$ becomes $f(\epsilon)$, then the various $H_{kt}$ become merely the $S_{kt}$ and Eq. (42) simplifies to Eq. (27). Furthermore, Eq. (43) for the weights then becomes Eq. (38).

### C6. Table of Important Equations

In concluding this section of the paper, it is desirable to summarize all important relations by means of a table in the form of the corresponding symbols of the $a-T$ and $b-V$ systems. The relation between each pair of symbols involves the factors $S_{tt}$(=Fisher's $\lambda_t$), the values of which, up to $n=30$, are given in our Table XII, up to $n=52$ by Fisher and Yates,[24] and up to $n=104$ by Anderson and Houseman.[27] The last column of Table V lists the equations used in calculating all quantities (functions, coefficients, and weights) thus far discussed.

## D. LEAST-SQUARES' POLYNOMIAL FITTING OF DATA IN TERMS OF FINITE DIFFERENCES

As already noted, the desired coefficients of the least-squares' polynomial may be expressed explicitly in terms of the finite differences of the observations, as well as in terms of the observations themselves. An example of the use of finite differences, applied to data from an Atwood machine, has been outlined in Section A. By this method it is the value of $a_t$ that is obtained most directly from the finite differences of order $t$, since $t!a_t$ is always a certain *weighted average* of such differences. The problem is then to determine the appropriate weights.

TABLE V. Relations of the $a-T$ (polynomials with unit leading coefficient) and $b-V$ (polynomials with least integer ordinates) systems. $\epsilon=(x-x_0)/h$, $x_0$ (or $m$) = middle value of $x$, and $h$ = constant interval of $x$. Numerical values of $V_t$ and $N_t$ are given in Table XIII, those of $S_{kt}$ are given in Table XII. Explicit expressions for $H_{kt}$ are found in Table IV.

| $a-T$ system | $b-V$ system | Relation | Relevant equations |
|---|---|---|---|
| $a_t$ | $b_t$ | $b_t=a_t/S_{tt}$ (23) | $a_t=\dfrac{\Sigma y\cdot T_t}{M_t},\quad b_t=\dfrac{\Sigma y\cdot V_t}{N_t}$ (7)(23) |
| $T_t$ | $V_t$ | $V_t=S_{tt}T_t$ (21) | $u_j(\epsilon)=\sum_{t=0}^{j} a_tT_t=\sum_{t=0}^{j} b_tV_t$ (5)(22) |
| $R_{kt}$ | $S_{kt}$ | $S_{kt}=S_{tt}R_{kt}$ (28) | $a_{kj}=\sum_{t=k}^{j} R_{kt}\cdot a_t=\sum_{t=k}^{j} S_{kt}\cdot b_t$ (10)(27) |
| $G_{kt}$ | $H_{kt}$ | $H_{kt}=S_{tt}G_{kt}$ | $a_{kj}''=\sum_{t=k}^{j} G_{kt}\cdot a_t=\sum_{t=k}^{j} H_{kt}\cdot b_t$ (40)(42) |
| $M_t(=\Sigma_\epsilon T_t^2)$ | $N_t(=\Sigma_\epsilon V_t^2)$ | $N_t=S_{tt}^2\cdot M_t$ (31) | $\dfrac{1}{p_j(\epsilon)}=\sum_{t=0}^{j}\dfrac{T_t^2}{M_t}=\sum_{t=0}^{j}\dfrac{V_t^2}{N_t}$ (35)(36)<br>$\dfrac{1}{p_{kj}}=\sum_{t=k}^{j}\dfrac{R_{kt}^2}{M_t}=\sum_{t=k}^{j}\dfrac{S_{kt}^2}{N_t}$ (37)(38)<br>$\dfrac{1}{p_{kj}''}=\sum_{t=k}^{j}\dfrac{G_{kt}^2}{M_t}=\sum_{t=k}^{j}\dfrac{H_{kt}^2}{N_t}$ (41)(43) |

$$u_j(\epsilon)=\sum_{k=0}^{j} a_{kj}\cdot\epsilon^k \quad (9) \qquad u_j(x)=\sum_{k=0}^{j} a_{kj}''x^k \quad (39)$$

$$\Sigma_\epsilon [v_j(\epsilon)]^2=\Sigma_\epsilon y^2-\sum_{t=0}^{j} M_ta_t^2=\Sigma_\epsilon y^2-\sum_{t=0}^{j} N_tb_t^2 \quad (29)(30)$$

The final result of the work on finite differences by Weinberg and the writer appears as Eqs. (54)–(56) ahead, with the values of the various symbols given by Eqs. (46), (51), and (53). The detailed use of the method constitutes Section F7 ahead. But in order to show the relation of our work to the earlier work of Sasuly, as well as to establish certain very important relations between the coefficients appearing in the formulas of the present section and those appearing in the equations of Section C, it is necessary to give the following rather detailed discussion.

Sasuly[12] considers the use of finite differences of the data for the present problem on pp. 47–58 of his book. He outlines there a method by which the weighted averages of the $\delta^t y$ values, for successive values of $t$, can be formulated, starting from the original observations. Special formulas are given for $t=1$, 2, 3, and 4, holding for the abscissa scale $x=1$ to $n$. Thus for $t=1$ and 2 he writes

$$1!a_1=\frac{6}{n(n^2-1)}\sum_{i=1}^{n-1} i(n-i)\Delta y_i. \tag{44}$$

$$2!a_2=\frac{30}{n(n^2-1)(n^2-4)} \times\sum_{i=1}^{n-2} i(i+1)(n-1)(n-i-1)\Delta^2 y_i. \tag{45}$$

His succeeding two formulas are increasingly complex and will not be quoted. Sasuly[12] also gives on p. 318 of his book an empirical generalization of these expressions that is equivalent to our Eq. (52).

The method used by the writer in obtaining the desired generalization is as follows. It was first noted that the factor preceding the summation in Eqs. (44) and (45) is just the $1/K_t$ of Eq. (14), the equation derived by Birge and Shea for the calculation of $a_t$ from the observations. The generalization of the factors following the summation sign is obvious by mere inspection. There are, however, several modifications that can well be made. In the first place, all of the weighting factors are symmetrical about the central finite difference. Hence these finite differences may be handled in pairs, the two differences comprising a pair having the same weight. For the purpose of designating a pair we again employ the symbol $r=|\epsilon|$, which is zero for the central difference if the number of differences is odd, and $\frac{1}{2}$ for the central *pair* of differences if the number is even.

In the second place, because of the arrangement in pairs it is convenient to use central differences together with our special abscissa $\epsilon$, which runs from $-q$ to $+q$ for the observations, but from $-\nu$ to $+\nu$ for a given set of finite differences of order $t$, where

$$2\nu=2q-t=n-t-1. \tag{46}$$

In this connection it should be noted that the observations themselves should now be considered as the finite differences of zero order, for which $t=0$ and $\nu=q$. Thus the equation for $t=0$, corresponding to Eq. (44) for $t=1$, is necessarily $a_0=\sum y_i/n$, which is just the familiar arithmetic average of the observations.

With the use, then, of central differences and the symbols $r$ and $\nu$ we get as the generalized formula for the coefficient $a_t$ of a polynomial of degree $t$, when the abscissae are spaced at *unit* interval,

$$t!a_t=\frac{1}{K_t}\sum_{r=0 \text{ or } \frac{1}{2}}^{\nu}(\delta^t y_r+\delta^t y_{-r})Q_t, \tag{47}$$

where

$$K_t=\frac{(2\nu+2t+1)!(t!)^2}{(2\nu)!(2t+1)!}=(t!)^2\binom{2\nu+2t+1}{2t+1} \tag{48}$$

and

$$Q_t=\frac{(\nu+t-r)!(\nu+t+r)!}{(\nu-r)!(\nu+r)!} =(t!)^2\binom{\nu+t-r}{t}\binom{\nu+t+r}{t}. \tag{49}$$

It should be noted that since $t!a_t$ is a weighted average of the finite differences $\delta^t y$, with weights $Q_t$, it necessarily follows that $K_t$ is merely the $\sum Q_t$ over *all* $\delta^t y$ values.

Replacing $2\nu$ in Eq. (48) by its equivalent $2q-t$, from Eq. (46), we see that the $K_t$ of Eq. (16) is just the $K_t$ of Eq. (48), as already stated. This identity is not entirely a matter of coincidence. In each case the $K_t$ is a factor that converts non-integer values into integers although, as it will appear, not the smallest possible factor. The identity does, however, lead to valuable numerical relations.

Thus let us rewrite Eq. (14) in the form

$$a_t=\sum_\epsilon y\cdot P_B/K_t\equiv\sum_\epsilon y\cdot P_B^*, \qquad (14')$$

and Eq. (47) in the form

$$a_t=\sum_\epsilon \delta^t y\cdot Q_t/t!K_t\equiv\sum_\epsilon \delta^t y\cdot Q_t^*. \qquad (47')$$

By comparing Eqs. (14′) and (47′) we see that, according to the theorem of summation by parts (see Section J), the complete coefficients $P_B^*$ of the observations $y$ in Eq. (14′) must be just the finite differences of order $t$ of the complete coefficients $Q_t^*$ of the $\delta^t y$ in Eq. (47′). It will be understood in future references to the quantities $Q_t$ that $t$ successive null values are adjoined beyond the range of $\epsilon$ in Eq. (47′), i.e., for $\epsilon=-\nu-1\ldots-\nu-t$ and $\epsilon=\nu+1\ldots\nu+t$. These null values are consistent with Eq. (49) for $Q_t$ as a function of $r=|\epsilon|$. Hence, if one had derived Eqs. (48) and (49), from which the values of $Q_t^*(=Q_t/t!K_t)$ immediately follow, but had not yet derived any formulas for the "pair-factors" of Eqs. (14), (15), etc., then one *could* have obtained all numerical values of $P_B^*$ by merely constructing a table of differences of the $Q_t^*$ values.[28]

Inspection of the $P_B^*$ and $Q_t^*$ values will show that in general they are not integers. Then $K_t$, as defined by Eq. (48) or by Eq. (16), may be considered as merely one possible integer factor that converts a given set of $Q_t^*$ values into the values $Q_t/t!$ which, by Eq. (49), are necessarily integers. If now the corresponding set of $P_B^*$ values in Eq. (14′) is multiplied by the *same* factor $K_t$, one must obtain *integer* values of $P_B$ (as defined by Eq. (15)), since these new $P_B$ values must be, in turn, merely the finite differences of the new integral $Q_t/t!$ values. Of course, Eq. (15) indicates the integral nature of the $P_B$ but, as noted, we are now assuming that this equation has not yet been derived.

Furthermore, a comparison of Eqs. (14′) and (47′) shows that

$$\sum_\epsilon y\cdot P_B=\sum_\epsilon \delta^t y\cdot Q_t/t!. \qquad (50)$$

[28] For $t$ an even integer, the resulting finite differences give just the values of $P_B^*$, but for $t$ odd, the results give the *negative* of $P_B^*$.

The sets of quantities $P_B$ and $Q_t/t!$ are, however, *not* the *smallest* sets of integers for a given value of $t$. As shown explicitly by Eqs. (15) and (49), the actual smallest sets of integers, independent of $n$, are obtained by removing a common integer factor ($t!$) from $P_B$ to give the $P_B''$ values of Eq. (15), and from $Q_t/t!$ to give, from Eq. (49),

$$Q_t'=Q_t/(t!)^2=\binom{\nu+t-r}{t}\binom{\nu+t+r}{t}. \qquad (51)$$

Then Eq. (50) takes on the new form

$$\sum_\epsilon y\cdot P_B''=\sum_\epsilon \delta^t y\cdot Q_t', \qquad (50')$$

where again the $P_B''$ values for a given $t$ are just the finite differences of order $t$ of the $Q_t'$ values.

We can now use the *integer* values $Q_t'$ and still preserve the form of Eq. (47) by writing

$$t!a_t=(1/K_t')\sum_{r=0\text{ or }\frac{1}{2}}^{\nu}(\delta^t y_r+\delta^t y_{-r})Q_t', \qquad (52)$$

where

$$K_t'=K_t/(t!)^2=\binom{2\nu+2t+1}{2t+1}. \qquad (53)$$

It is actually Eqs. (51)–(53) that have been derived by Weinberg with the use of direct algebraic methods (see Section J).

Finally, if there are *special* integer factors $f$, for special values of $n$ that may be removed from the respective sets of $Q_t'$ values to give what we shall term $W_t$ values then, by Eq. (50′), these *same* factors $f$ must be removable from the $P_B''$ values to give the $V_t$ values, as mentioned in footnote 22. For, again, the new $V_t$ values must be merely the finite differences of order $t$ of the new $W_t$ values. We thus get as a *second* new form of Eq. (50)

$$\sum_\epsilon y\cdot V_t=\sum_\epsilon \delta^t y\cdot W_t. \qquad (50'')$$

To preserve the form of Eq. (52) or of Eq. (47) we write

$$t!a_t=(1/L_t)\sum_{r=0\text{ or }\frac{1}{2}}^{\nu}(\delta^t y_r+\delta^t y_{-r})W_t, \qquad (54)$$

where

$$W_t=Q_t'/f=Q_t/(t!)^2f \qquad (55)$$

and, necessarily, from Eq. (52),

$$L_t=K_t'/f=K_t/(t!)^2f. \qquad (56)$$

Weinberg has now been able to devise a general method for determining the $f$ values of Eqs. (55) and (56). Such information greatly simplifies the numerical calculation of the $W_t$ and $L_t$ values starting with the derived Eqs. (51) and (53). The values of $f$ thus found are listed in Table VI, for the range $t=1$ to 5, and $n$ to 30. The resulting values of $W_t$ and $L_t$, over the same range, are listed in Table XIV, which is entirely new. As just noted, these $W_t$ and $L_t$ represent the smallest possible sets of integers for the evaluation of $a_t$ by Eq. (54).[29]

Since $t!a_t$ of Eq. (54) is a properly weighted average value of $\delta^t y$, it follows that $L_t$ is necessarily the sum of all the weighting factors $W_t$ of a given set, just as $K_t$ of Eq. (47) is the sum of the weighting factors $Q_t$, as already noted. But since the weights are always distributed symmetrically about $\epsilon=0$, only the common weight $W_t$ of a given value of $r=|\epsilon|$ is recorded in Table XIV. Hence the values of $L_t$, in terms of the *recorded* sets of $W_t$, are given by

$$L_t=2\sum_{r=\frac{1}{2}}^{\nu} W_t$$

for an *even* number of $\delta^t y$ values (half-integer $r$ values), and

$$L_t=W_t(\text{for } r=0)+2\sum_{r=1}^{\nu} W_t$$

for an odd number of $\delta^t y$ values (integer $r$ values).

The calculation of numerical values of $Q_t'$ and $K_t'$, from Eqs. (48), (49), (51), (53) is really extremely simple, because of certain recursion relations.[30] The necessary relation between $L_t$ and the sum of the $W_t$ values furnishes a complete numerical check on the listed values.

It may be noted that the analytic expression for $t!a_t$ assumes an equally simple form, if one uses the abscissae $x=0$ to $2\nu$ in place of $\epsilon=-\nu$ to $+\nu$. Thus we obtain, in place of Eqs. (52) and (51),

$$t!a_t=(1/K_t')\sum_{x=0}^{2\nu} Q_t''\Delta^t y_x, \tag{57}$$

where

$$Q_t''=\binom{t+x}{t}\binom{2\nu+t-x}{t} =\frac{(x+t)!(2\nu+t-x)!}{(2\nu-x)!x!(t!)^2}. \tag{58}$$

It should be emphasized that the $x$ values, like those of $\epsilon$, are assumed to be spaced at *unit* interval. The $K_t'$ values are given by Eq. (53) as before.

Weinberg's derivation actually leads to Eqs. (57), (58), and I then transformed his results to Eqs. (52), (51) in order to check my own Eqs. (47), (49), which in turn represent a generalization and transformation of Sasuly's specific expressions for $t=1$ to 4. It should further be noted that Jordan[20] developed a formula for the orthogonal polynomial $T_j$ as the $j$th difference of a polynomial of degree $2j$, which can readily be transformed into Eq. (57) by repeated summation by parts (see Section J).

Although it is a very simple matter to calculate any desired $a_t$ value by the use of Eq. (54) and Table XIV, it is a far more laborious process to obtain the probable error $r_{tt}$ in $a_t(=a_{tt})$. But the same situation exists when a given value of $a_t$ is calculated directly in terms of the observations, as shown by the equations and discussion of Sections C5 and C6. A numerical calculation of all probable errors of interest is made in Sections

---

[29] It was only after these $f$ values had been determined that we noted the necessary relation between them and the $S_{tt}$ values of Table XII, cited in footnote 22 ($f\cdot S_{tt}=2t!/(t!)^3$), which follows from the necessary relation of the $V_t$ and $W_t$ values just discussed. We do not know how the $S_{tt}$ values (to $n=104$) published by Anderson and Houseman[27] were actually obtained. Those used by Birge and Shea[3] (to $n=30$) were found purely empirically by studying each set of $P_B$ values of Eq. (15) for a possible common integer factor. We have now used our $f$ values (extended to $n=104$) to verify *all* of the $S_{tt}$ values published by Anderson and Houseman[27] (their $\lambda_t$).

[30] As an example of the simplicity of the numerical work involved, consider the set of values for $t=4$ and $n=24$. The successive values of $Q_t'$, Eq. (51), are obtained from the respective values for $n=23$ by multiplying by the successive factors 14/10, 15/11, 16/12, $\cdots$23/19. Then for $n=25$, where we have an additional fourth-difference coefficient, the initial $Q_t'(r=\frac{1}{2})$ for $n=24$ is multiplied first by 14/10 and then by 15/11 to give the initial two values of $Q_t'$ for $n=25$. The succeeding values are gotten from the remainder of those for $n=24$ by the use of the factors 16/12, 17/13, $\cdots$24/20. The quoted figures are merely to illustrate two quite obvious general rules, which it does not seem necessary to quote, one applying to an even number of difference coefficients, and the other to an odd number. By their use all values of both $Q_t'$ and $K_t'$ are easily derived from preceding values. Then the $f$ factor listed in Table VI is removed to give the final $W_t$ and $L_t$ values of Table XIV.

F4 and F6, and in Section F7 an explicit equation and a calculation are given for the probable error of a certain $a_t$ coefficient, whose value has been obtained from finite differences.

The present paper is devoted primarily to the discussion of a complete least-squares' solution, including probable errors, explicitly in terms of the observations. A corresponding complete solution could certainly be worked out in terms of finite differences. We are, however, confining the present discussion of finite differences to the calculation only of the $a_t$ values. For just that reason the formulas quoted, such as Eq. (54), are designed to lead directly to the value of $a_t$ rather than of $b_t$ which, for purely numerical reasons, is found to be especially convenient when one desires calculated values of the function at various points and the corresponding probable errors of the function.

TABLE VI. Values of $f$ in Eqs. (55) and (56).

| $n$ | $t=1$ | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 2 | 1 | | | | |
| 3 | 2 | 1 | | | |
| 4 | 1 | 3 | 1 | | |
| 5 | 2 | 3 | 4 | 1 | |
| 6 | 1 | 2 | 2 | 5 | 1 |
| 7 | 2 | 3 | 20 | 5 | 6 |
| 8 | 1 | 3 | 5 | 5 | 3 |
| 9 | 2 | 1 | 4 | 5 | 14 |
| 10 | 1 | 6 | 2 | 7 | 21 |
| 11 | 2 | 3 | 4 | 35 | 84 |
| 12 | 1 | 1 | 5 | 10 | 14 |
| 13 | 2 | 3 | 20 | 5 | 36 |
| 14 | 1 | 6 | 2 | 5 | 9 |
| 15 | 2 | 1 | 4 | 1 | 2 |
| 16 | 1 | 3 | 1 | 5 | 21 |
| 17 | 2 | 3 | 20 | 35 | 42 |
| 18 | 1 | 2 | 10 | 35 | 7 |
| 19 | 2 | 3 | 4 | 5 | 84 |
| 20 | 1 | 3 | 1 | 2 | 6 |
| 21 | 2 | 1 | 4 | 5 | 4 |
| 22 | 1 | 6 | 10 | 5 | 9 |
| 23 | 2 | 3 | 20 | 5 | 126 |
| 24 | 1 | 1 | 1 | 35 | 7 |
| 25 | 2 | 3 | 4 | 7 | 42 |
| 26 | 1 | 6 | 2 | 5 | 21 |
| 27 | 2 | 1 | 20 | 5 | 4 |
| 28 | 1 | 3 | 5 | 10 | 6 |
| 29 | 2 | 3 | 4 | 5 | 12 |
| 30 | 1 | 2 | 2 | 1 | 7 |

In spite of the limited extent of the present presentation in terms of finite differences, it is believed that Eq. (54), in connection with our new Table XIV, will prove of great value in many experimental situations such as in the calculation of acceleration by means of an Atwood's machine, considered briefly in Section A.

## E. SPECIAL TREATMENT OF FIRST DEGREE POLYNOMIAL

The first-degree polynomial represents merely one special case of the general methods discussed in Sections C and D. But for this commonly used function, the equations needed for the least-squares' solution are so simple that it is convenient to write them in explicit form. Incidentally, a direct derivation of these equations shows, in principle, one method of deriving the more general equations already listed. In fact, the following method is actually the one used by Birge and Shea[3] in obtaining the needed results for the first-, second-, and third-degree polynomials. The method, however, becomes intolerably complex for polynomials of higher degree.

Let us start with the standard equations for the least-squares' solution of $n$ unweighted observations (distributed in any way on the $x$ axis) in the form of a first-degree polynomial. For this case

$$u_1(x)=a_{01}+a_{11}x, \tag{59}$$

where

$$a_{01}=\left(\sum y \sum x^2-\sum xy \sum x\right)/D, \tag{60}$$

$$a_{11}=\left(n\cdot\sum xy-\sum x \sum y\right)/D, \tag{61}$$

and

$$D=n\cdot\sum x^2-\left(\sum x\right)^2. \tag{62}$$

If now we restrict ourselves to the *special* case of $n$ observations $y_1, y_2, y_3, \cdots, y_n$, corresponding to $x=1, 2, 3, \cdots, n$, we may employ the well-known expressions

$$\sum_1^n x=n(n+1)/2, \tag{63}$$

$$\sum_1^n x^2=n(n+1)(2n+1)/6. \tag{64}$$

Substituting Eqs. (63), (64), into (61), we obtain after a little algebraic reduction,

$$a_{11}=\frac{6}{n(n^2-1)}\{(y_n-y_1)(n-1)$$
$$+(y_{n-1}-y_2)(n-3)+\cdots\}. \quad (65)$$

This equation for $a_{11}$ may be found in various texts on physical measurements. On examining it, we note that each *pair* of observations is multiplied by a factor that equals the interval $\Delta x$ between the two observations comprising the pair; in other words, a certain interval $\Delta y$ is multiplied by its corresponding interval $\Delta x$.

Furthermore, the reciprocal of the factor outside the summation is just the sum of alternate squares from $(n-1)$ to 1 (or 0). Thus

$$n(n^2-1)/6=(n-1)^2+(n-3)^2+\cdots 1^2 \text{ or } 0^2. \quad (66)$$

To be specific, Eq. (65) for an *even* number of observations, such as $n=6$, has the form

$$a_{11}=\frac{5(y_6-y_1)+3(y_5-y_2)+(y_4-y_3)}{5^2\quad+\quad 3^2\quad+\quad 1^2}, \quad (67)$$

but for an *odd* number, such as $n=5$,

$$a_{11}=\frac{4(y_5-y_1)+2(y_4-y_2)+0(y_3-y_3)}{4^2\quad+\quad 2^2\quad+\quad 0^2}. \quad (68)$$

It should be noted, from Eq. (68), that in the case of an odd number of observations the *middle* observation (here $y_3$) does not appear and, hence, has no effect on the value of the slope.

If now we understand that observations are to be combined in pairs in the *particular* way shown by Eq. (67) or Eq. (68), i.e., last and first, next to last and second, etc., then Eq. (65) might be written

$$a_{11}=\frac{\sum_i \Delta y_i\Delta x_i}{\sum_i (\Delta x_i)^2}=\frac{\sum_i (\Delta y_i/\Delta x_i)(\Delta x_i)^2}{\sum_i (\Delta x_i)^2}. \quad (69)$$

The second form of Eq. (69) shows that $a_{11}$ is the *weighted average* of a series of slopes $\Delta y_i/\Delta x_i$, each of which is weighted proportional to the *square* of the interval, $\Delta x$, covered by the slope. Thus two points twice as far apart as another set of two points contribute four times as much weight in determining the slope $a_{11}$ of the least-squares' straight line through the data.

In Section D we have given the weights to be assigned to the slopes (first differences) furnished by *successive* points. But Eq. (69) applies to the quite different pairing of points shown by Eq. (65).

We now continue the least-squares' solution by substituting Eqs. (63) and (64) into (60). The result after some algebraic reduction is

$$a_{01}=a_{00}-\frac{(n+1)}{2}a_{11}, \quad (70)$$

where

$$a_{00}=\sum y/n, \quad (71)$$

and where $a_{11}$ is given by Eq. (65). Thus in the case of the first-degree equation we already see emerging the pattern of the least-squares' solution for the case of unweighted, equally spaced data. Any coefficient (here $a_{01}$) can be expressed as the sum of terms containing the final coefficient of the polynomial in question and the final coefficients of the polynomials of all lower degrees. These coefficients are here merely $a_{11}$ and $a_{00}$. In general they are $a_{tt}$, a symbol that has been simplified to $a_t$ in Sections C and D. The *general* formula for any coefficient $a_{kj}$ is given by Eq. (10) of Section C.

The proper expressions for $R_{kt}$ in Eq. (10) are different from those now being derived, since we are here using the scale $x=1$ to $n$ in place of $\epsilon=-q$ to $+q$. Thus from Eqs. (70), (71), and (59) we have

$$u_1(x)=\left\{a_{00}-\frac{(n+1)}{2}a_{11}\right\}+a_{11}x. \quad (72)$$

But with $x=\epsilon+x_0=\epsilon+(n+1)/2$ we get

$$u_1(\epsilon)=a_{00}+a_{11}\epsilon. \quad (73)$$

We have now expressed $u_1(\epsilon)$ in the form of Eq. (5), i.e., as the sum of orthogonal polynomials,

$$u_1(\epsilon)=a_0T_0+a_1T_1, \quad (74)$$

where $T_0=1$ and $T_1=\epsilon$, in agreement with Table III.

Another interesting, but lengthy process for calculating the least-squares' slope of a straight line is by means of the equation

$$a_{11}=\sum \Delta y/\sum \Delta x, \quad (75)$$

where each $\Delta x$ corresponds to a $\Delta y$, and the $\Delta y$'s are obtained by combining the observations in pairs in *every possible way*. Thus for 5 observations $y_1$ to $y_5$ we obtain the list

| $\Delta y$ | $\Delta x$ |
|---|---|
| $y_2-y_1$ | 1 |
| $y_3-y_1$ | 2 |
| $y_4-y_1$ | 3 |
| $y_5-y_1$ | 4 |
| $y_3-y_2$ | 1 |
| $y_4-y_2$ | 2 |
| $y_5-y_2$ | 3 |
| $y_4-y_3$ | 1 |
| $y_5-y_3$ | 2 |
| $y_5-y_4$ | 1 |

Then $\sum \Delta y=4(y_5-y_1)+2(y_4-y_2)$, and $\sum \Delta x=20=4^2+2^2$, in agreement with Eq. (68). This last method has been used in elementary laboratories, but I do not recommend it. I see no reason, however, why Eq. (65) or the even more explicit forms, such as Eqs. (67), (68), should not be used even in an elementary physics course. In fact, the essential object of this section is to call attention to the simplicity of equations such as (67) and (68), which obviously can be written down from memory without reference to any more general formula.

## F. EXPLICIT DIRECTIONS AND ILLUSTRATIVE PROBLEM

### F1. Values of $b_t$ and $\sum v^2$ (Model Form 1)

The object of this section is to present detailed directions for the use of the results contained in Sections C and D. The clearest method of presenting such material, it has always seemed to the writer, is by means of an illustrative problem. Although it is obviously necessary to refer repeatedly to the equations of Sections C and D, the present section is, so far as possible, independent of the preceding sections. It is hoped, therefore, that potential users of the method will be able from this section alone (when used in connection with the various tables of numerical values) to follow the necessary steps.

Our object is to fit data, equally spaced along the $x$ axis, and unweighted, to a polynomial of any desired degree by the *method of least squares*, and to obtain the probable errors of various calculated quantities. Although the general formulas given in Table III of Section C2 extend up to the *tenth*-degree polynomial, the various tables of numerical values presented here extend only to the fifth degree (and to 30 observations).

As a sample problem we choose seven observations that are closely satisfied by a fourth-degree polynomial. A sixth-degree function will exactly satisfy all seven points and, due to the deliberately chosen symmetry of the fourth differences, the sum of the squares of the residuals is no smaller for the fifth-degree solution than for that of the fourth degree. The assumed observations have also been so chosen that all essential calculated quantities appear as simple terminating decimals and hence the *exact* value of each can be and is obtained. All computations have been made on a 10-key calculating machine. The methods advocated in this paper require, for rapid work, such a machine rather than an adding-tabulating machine. This point is discussed in Section G.

The first five observations of the seven now presented have been fitted *exactly* to a fourth-degree polynomial, in Section B, and directions for such work are given there. The process given in Section C and now illustrated (Table VII) is, however, equally valid in such a case.

For the method discussed in Section C it is not necessary to calculate the differences since all results are in terms of the observations. The $\epsilon$ scale of abscissas is the standard scale in terms of which the results are first obtained. It is defined by $\epsilon=(x-m)/h=e/h$, where $h=5$, the constant interval in terms of $x$, and $m$ (also denoted $x_0$) is the *middle* value of $x(=15)$. For an

TABLE VII. Standard illustrative problem. Data and differences.

| $x$ | $\epsilon$ | $y$ | $\delta y$ | $\delta^2 y$ | $\delta^3 y$ | $\delta^4 y$ |
|---|---|---|---|---|---|---|
| 0 | $-3$ | 0 | | | | |
| | | | 2.10 | | | |
| 5 | $-2$ | 2.10 | | 4.41 | | |
| | | | 6.51 | | 0.42 | |
| 10 | $-1$ | 8.61 | | 4.83 | | 49.35 |
| | | | 11.34 | | 49.77 | |
| 15 | 0 | 19.95 | | 54.60 | | 51.66 |
| | | | 65.94 | | 101.43 | |
| 20 | $+1$ | 85.89 | | 156.03 | | 49.35 |
| | | | 221.97 | | 150.78 | |
| 25 | $+2$ | 307.86 | | 306.81 | | |
| | | | 528.78 | | | |
| 30 | $+3$ | 836.64 | | | | |

even number of observations $m$ is half-way between the center *pair* of values of $x$. The values of $\epsilon$ always run from $-q$ to $+q$ where $2q+1=n$ = number of observations.

The observations are now to be combined in pairs, the last and first, next to last and second, etc. Hence they are listed in columns 5 and 7 of model form 1, with the final $y$ (largest value of $x$) at the head of column 7 and the initial $y$ at the head of column 5. The observations comprising each pair then lie on the same horizontal line. The pair is designated by the value of $r=|\epsilon|$ in column 6. For an odd number of observations, as here, there is only one observation corresponding to $r=0$, and the last space in column 5 is left vacant. For an even number of observations the last pair has $r=\frac{1}{2}$ (or $\epsilon=\pm\frac{1}{2}$).

Column 8 is the sum of columns 7 and 5, and column 4 is the difference. The numbers in columns 2, 3, 9, and 10 are the values of $V_t$, as copied from Table XIII, for $n=7$. The missing column 1, to the left of column 2, would be headed $V_5$ and would be used in the case of a fifth-degree solution. These values of $V_t$ are called "pair-factors" by Birge and Shea.[3] The values of $N_t$, occupying the first horizontal line below the data (row (1)), are likewise taken from Table XIII for $n=7$. The value of $N_0$ is merely $n$, and is placed in column 8. These values of $N_t$ differ numerically from the values designated as $K$ by Birge and Shea[3] and by $K_t^*$ here, since $K_t^*=N_t/S_{tt}$ by Eq. (26). The $K$ values of Birge and Shea lead directly to the values of $a_t$ whereas the $N_t$ values lead directly to $b_t$, as shown in Section C, Eqs. (25) and (23).

Row (2) gives the algebraic sum of the products $\sum_\epsilon y\cdot V_t$ for each value of $t$ from 0 to $j$. Each of these sums can be obtained in one continuous process on any calculating machine. Actually the values of $V_t$ for negative $\epsilon$ are the negative of those for positive $\epsilon$, if $t$ is an odd integer, but there is no such change of sign for $t$ an even integer. These relations are evident from Table III and they are automatically satisfied by multiplying the successive *differences* in column 4 first by $V_1$ and then by $V_3$, and the successive *sums* in column 8, first by $V_2$ and then by $V_4$. Since the value of $V_0$ is always unity, the $\sum_\epsilon y\cdot V_0$ is obtained by merely adding column 8, and the sum is recorded in row (2), column 8.

Thus in row (2),

$$3198.72=3\times 836.64+2\times 305.76+77.28,$$

and

$$554.40=3\times 836.64-7\times 309.96+94.50+6\times 19.95.$$

Then row (3) ($=b_t$) is obtained by dividing each row (2) result by the corresponding value of $N_t$ in row (1). Rows (4) and (5) ($b_t^2$ and $N_t\cdot b_t^2$) are then to be calculated. The results

TABLE VIII. Values of $\Sigma_\epsilon\,[v_j(\epsilon)]^2$ from Eq. (30).

| | | |
|---|---|---|
| $\Sigma_\epsilon\, y^2=$ | 802,597.9059 | (exact) |
| $-N_0b_0^2=$ | $-$227,178.1575 | |
| $\Sigma\, v_0^2=$ | 575,419.7484 | |
| $-N_1b_1^2=$ | $-$365,421.7728 | |
| $\Sigma\, v_1^2=$ | 209,997.9756 | |
| $-N_2b_2^2=$ | $-$173,709.9525 | |
| $\Sigma\, v_2^2=$ | 36,288.0231 | |
| $-N_3b_3^2=$ | $-$34,292.1600 | |
| $\Sigma\, v_3^2=$ | 1,995.8631 | |
| $-N_4b_4^2=$ | $-$1,995.8400 | |
| $\Sigma\, v_4^2=$ | 0.0231 | (exact) |

MODEL FORM 1. Values of $b_t$ from the observations.

| | (2) $V_3$ | (3) $V_1$ | (4) diff. | (5) $y_{-q}\cdots$ | (6) $r$ | (7) $y_q\cdots$ | (8) sum | (9) $V_2$ | (10) $V_4$ |
|---|---|---|---|---|---|---|---|---|---|
| | +1 | +3 | 836.64 | 0 | 3 | 836.64 | 836.64 | +5 | +3 |
| | −1 | +2 | 305.76 | 2.10 | 2 | 307.86 | 309.96 | 0 | −7 |
| | −1 | +1 | 77.28 | 8.61 | 1 | 85.89 | 94.50 | −3 | +1 |
| | 0 | 0 | — | — | 0 | 19.95 | 19.95 | −4 | +6 |
| (1) | 6 | 28 | | | $N_t$ | | 7 | 84 | 154 |
| (2) | 453.6 | 3198.72 | | | $\Sigma_\epsilon\, y\cdot V_t$ | | 1261.05 | 3819.90 | 554.40 |
| (3) | 75.6 | 114.24 | | | $b_t$ | | 180.150 | 45.475 | 3.60 |
| (4) | 5715.36 | 13050.7776 | | | $b_t^2$ | | 32454.0225 | 2067.975625 | 12.96 |
| (5) | 34292.16 | 365421.7728 | | | $N_t\cdot b_t^2$ | | 227178.1575 | 173709.9525 | 1995.84 |

given in these last two rows are needed only for calculating the sums of squares of residuals (and hence, finally, any desired probable errors). These last sums are given by Eq. (30) which is now repeated for convenience.

$$\sum_{\epsilon} [v_j(\epsilon)]^2 = \sum_{\epsilon} y^2 - \sum_{t=0}^{j} N_t b_t^2. \quad (30)$$

Hence we get, by repeated subtraction, the value of $\sum_\epsilon v^2$ for polynomials of successive degree $j=0, 1, 2$, etc. The magnitude of $\sum v^2$ is probably the best criterion of the fit of the data to the polynomial used. The figures for our sample problem are given in Table VIII.

Ordinarily the calculations need not be carried to so many digits, but if one wishes an *accurate* value of the residuals for a curve that fits the data very closely, it is necessary to retain all digits shown. Here all results are *exact*, both in model form 1 and in the $\sum_\epsilon [v_t(\epsilon)]^2$ calculations.

As already noted, the fourth differences are symmetric about $\epsilon=0$. This fact leads to $b_5=0$, and hence $\sum v_5^2$ also equals 0.0231. The $\sum v_6^2$ is necessarily zero, since a sixth-degree polynomial exactly fits any given seven points. It is thus evident that a fourth-degree polynomial gives a satisfactory fit, and a far better fit than does one of any lower degree. The same information is shown by the constancy, or lack of constancy, of the various columns of finite differences, as given in Table VII. But the successive values of $\sum v^2$ put this information in quantitative form.

One therefore has the following rule for choosing the degree of the polynomial to be employed. Increase the degree until the value of $\sum v^2$ has dropped to a roughly constant value. The size of this final value is a measure of the goodness of fit. It should, of course, be remembered that the higher the degree of the polynomial and thus the greater the number of undetermined coefficients, the better the fit will necessarily be. But each added degree should, in general, produce a very *large proportional* decrease in the value of $\sum v^2$. If it does not do so, there is little justification for the added degree. Thus, in this case, the first-degree polynomial is little better than that of zero degree, the second degree makes a considerable improvement, the third degree still greater (proportional) improvement, and the fourth degree an enormously greater improvement. But the fifth degree produces no further improvement at all. Hence the fourth-degree polynomial, but one of no lower degree, is to be considered a satisfactory representation of the data.

## F2. Values of $u_j(\varepsilon)$ and $v_j(\varepsilon)$ for Observed Points (Model Form 2)

The next information that we obtain easily and rapidly consists of the calculated values and the probable errors of each observation. With somewhat more labor we can determine the calculated value and probable error of any *extrapolated* point. It is very common in physical science to obtain an analytic representation of a set of data primarily in order to calculate a certain extrapolated value (possibly the ordinate at $x=0$ for experimental data extending over some positive range of values $x_1$ to $x_2$). Presumably because of the fact that the standard method for the least-squares' solution of a polynomial and for the calculation of its probable error is a laborious process, even for one of the second degree, the calculation of the *probable error* of an extrapolated value is practically never carried out by physical scientists.

But the representation of a polynomial as a *sum of orthogonal polynomials*, as is done in Section C, makes such a calculation comparatively simple, as will be illustrated immediately. There are numerous cases in the literature where an extrapolated value has been calculated and later used as if it were both reliable and important. But if the probable error of such an extrapolated point had been calculated, it is probably no exaggeration to say that in a substantial fraction of the cases the error thus obtained would have been found to be so large as to nullify completely any significance that might be attached to the result. Especially in the case of polynomials of high degree, the probable errors increase with extreme rapidity as we move beyond the limits of the experimental data. This fact has been emphasized and illustrated in an important paper by Schultz.[31]

[31] H. Schultz, "The standard error of a forecast from a curve," J. Am. Stat. Assoc. 25, 139 (1930), p. 159.

MODEL FORM 2. Calculation of $u_j(\epsilon)$ and $v_j(\epsilon)$ values. $u_j(\epsilon)=b_0V_0+b_1V_1+\cdots b_jV_j$, $u_0(\epsilon)=b_0V_0=b_0=180.15$, $u_4(\epsilon)=180.15+114.24\,V_1+45.475\,V_2+75.6\,V_3+3.6\,V_4$. (Values of $V_t$ and $b_t$ from model form 1.)

| $\epsilon$ | $y$ | $y-u_0(\epsilon)$ $=v_0(\epsilon)$ | $b_1V_1$ $=114.24\,V_1$ | $u_0(\epsilon)+b_1V_1$ $=u_1(\epsilon)$ | $y-u_1(\epsilon)$ $=v_0(\epsilon)-b_1V_1$ $=v_1(\epsilon)$ | $b_2V_2$ $=45.475\,V_2$ | $u_1(\epsilon)$ $+b_2V_2$ $=u_2(\epsilon)$ |
|---|---|---|---|---|---|---|---|
| −3 | 0 | −180.15 | −342.72 | −162.57 | +162.57 | +227.375 | +64.805 |
| −2 | 2.10 | −178.05 | −228.48 | −48.33 | +50.43 | 0 | −48.330 |
| −1 | 8.61 | −171.54 | −114.24 | +65.91 | −57.30 | −136.425 | −70.515 |
| 0 | 19.95 | −160.20 | 0 | +180.15 | −160.20 | −181.900 | −1.750 |
| +1 | 85.89 | −94.26 | +114.24 | +294.39 | −208.50 | −136.425 | +157.965 |
| +2 | 307.86 | +127.71 | +228.48 | +408.63 | −100.77 | 0 | +408.630 |
| +3 | 836.64 | +656.48 | +342.72 | +522.87 | +313.77 | +227.375 | +750.245 |

| $\epsilon$ | $y-u_2(\epsilon)$ $=v_1(\epsilon)-b_2V_2$ $=v_2(\epsilon)$ | $b_3V_3$ $=75.6\,V_3$ | $u_2(\epsilon)+b_3V_3$ $=u_3(\epsilon)$ | $y-u_3(\epsilon)$ $=v_2(\epsilon)-b_3V_3$ $=v_3(\epsilon)$ | $b_4V_4$ $=3.6\,V_4$ | $u_3(\epsilon)+b_4V_4$ $=u_4(\epsilon)$ | $y-u_4(\epsilon)$ $=v_3(\epsilon)-b_4V_4$ $=v_4(\epsilon)$ | $x$ |
|---|---|---|---|---|---|---|---|---|
| −3 | −64.805 | −75.6 | −10.795 | +10.795 | +10.8 | +0.005 | −0.005 | 0 |
| −2 | +50.430 | +75.6 | +27.270 | −25.170 | −25.2 | +2.070 | +0.030 | 5 |
| −1 | +79.125 | +75.6 | +5.085 | +3.525 | +3.6 | +8.685 | −0.075 | 10 |
| 0 | +21.700 | 0 | −1.750 | +21.700 | +21.6 | +19.850 | +0.100 | 15 |
| +1 | −72.075 | −75.6 | +82.365 | +3.525 | +3.6 | +85.965 | −0.075 | 20 |
| +2 | −100.770 | −75.6 | +333.030 | −25.170 | −25.2 | +307.830 | +0.030 | 25 |
| +3 | +86.395 | +75.6 | +825.845 | +10.795 | +10.8 | +836.645 | −0.005 | 30 |

In contrast to the laborious method of determinants used by Schultz, we now have the following rapid method for calculating ordinates and probable errors. From Section C we rewrite

$$u_j(\epsilon)=\sum_{t=0}^{j} b_tV_t, \tag{22}$$

in which both the $b_t$ and the $V_t$ values[32] required for each value of $\epsilon$, over the range of the data, are given in model form 1.

Just as in the case of $\sum v^2$ in Eq. (30), the form of Eq. (22) shows that in obtaining each calculated value $u_j(\epsilon)$ for the polynomial of degree $j$, we simultaneously obtain the calculated value for each polynomial of lower degree. In fact, as shown by the various equations of Section C, *every* desired result, when calculated by an orthogonal polynomial method, is given by a *sum* of terms. If one desires merely the result corresponding to a polynomial of stated degree, all that need be recorded is the actual sum. But by recording each successive *partial* sum one obtains simultaneously the solutions for *all* polynomials of lower degree. In the case of the specific problem treated in the present section we shall in every case record and use such partial sums, partly to show the possibility of such additional results and partly to illustrate in detail the improvement in the fitting of a given set of data to a polynomial as the degree of the polynomial is increased (from 0 to 4).

There is a very significant interpretation of the *change* in the $u_j(\epsilon)$ values as $j$ increases. Let us start with the zero-degree polynomial. The solution is merely $u_0(\epsilon)=b_0V_0=b_0=\sum y/n$, the arithmetic average of the observations ($=180.15$ in model form 1). Suppose that we now calculate the *residuals* $v_0(\epsilon)[=y-u_0(\epsilon)]$ for such a solution. Obviously they are large, but let us attempt to fit these residuals $v_0(\epsilon)$ to a *first*-degree polynomial. In other words, we now consider the $v_0(\epsilon)$ as new *observed* values. We then find for the *calculated* values of $v_0$ just the *second* term of Eq. (22), i.e., $b_1V_1$. Then the difference of the

[32] As just mentioned, each value of $V_t$ for a *negative* value of $\epsilon$ when $t$ is an *odd* integer, is the *negative* of that given in model form 1 (as taken from Table XIII). The magnitude of $\epsilon$ is given by $r$.

observed and calculated values $[v_0(\epsilon)-b_1V_1]$ gives the residuals $v_1$ of the first-degree least-squares' polynomial $u_1(\epsilon)$. Next we consider the $v_1(\epsilon)$ as new observed values and attempt to fit them to a *second*-degree polynomial. The result is then just the *third* term of Eq. (22), i.e., $b_2V_2$. The difference $[v_1(\epsilon)-b_2V_2]$ in turn gives the residuals $v_2(\epsilon)$ of the second-degree polynomial $u_2(\epsilon)$.

We thus find that each *orthogonal polynomial* $b_tV_t$ of Eq. (22) actually represents the least-squares' polynomial of degree $t$ for the *residuals* (*considered as observations*) of the least-squares' solution of *one lower* degree. As a series of algebraic steps the foregoing process reads

$$\begin{aligned}
u_0(\epsilon)&=b_0V_0,\\
v_0(\epsilon)&=y-u_0(\epsilon),\\
u_1(\epsilon)&=u_0(\epsilon)+b_1V_1,\\
v_1(\epsilon)&=y-u_1(\epsilon)=y-(u_0(\epsilon)+b_1V_1)\\
&=v_0(\epsilon)-b_1V_1,\\
u_2(\epsilon)&=u_1(\epsilon)+b_2V_2,\\
v_2(\epsilon)&=y-u_2(\epsilon)=y-(u_1(\epsilon)+b_2V_2)\\
&=v_1(\epsilon)-b_2V_2, \quad (76)
\end{aligned}$$

etc.

This interpretation of the orthogonal polynomial $b_tV_t$, of degree $t$, as the analytic function best fitting a series of residuals $v_{t-1}(\epsilon)$, which result from the use of a function of degree $t-1$, appears to be highly significant.

Having found that a fourth-degree polynomial represents a satisfactory fit of the data of model form 1, one would ordinarily merely calculate and add the five terms of Eq. (22) that are necessary to obtain each value of $u_4(\epsilon)$. But in order to show the foregoing relations between successive orthogonal polynomials and successive residuals we include these residuals in model form 2, which covers the range of the data $\epsilon=-3$ to $+3$.

Then, in model form 3, we will calculate the extrapolated values out to $\epsilon=-6$ and $+6$. Next we will obtain, in model form 4, the *weights* of all points over the range $\epsilon=-6$ to $+6$, for the polynomials of all degrees from zero to four and finally in model form 5 the corresponding *probable errors* in all these cases. Thus we obtain a complete picture of the fitting of our seven assumed observations, not only to a polynomial of the fourth degree, but also to polynomials of all lower degrees.

In connection with model form 2, we again call attention to the relations mentioned in footnote 32.

### F3. Values of $u_j$ for Extrapolated Points. (Model Form 3)

We will next calculate *extrapolated* values of $u_j$ out to $\epsilon=-6$ and $+6$ by the use of the same Eq. (22) that we have just used for the range $\epsilon=-3$ to $+3$. But now we must first *calculate* values of $V_t$ for the desired values of $\epsilon$, since the values of $V_t$ in Table XIII cover only the range of the observations. The equations for $V_t$ are given in Table III, provided we substitute $S_{kt}$ for $R_{kt}$. The values of $S_{kt}$ are given in Table XII. Here we wish the values for $n=7$. For convenience, the needed expressions from Table III in terms of $S_{kt}$ are now listed. These expressions can obviously be used for *any* value of $\epsilon$, either interpolated or extrapolated.

$$\begin{aligned}
V_0&=S_{00}=1,\\
V_1&=S_{11}\epsilon,\\
V_2&=S_{22}\epsilon^2+S_{02},\\
V_3&=S_{33}\epsilon^3+S_{13}\epsilon,\\
V_4&=S_{44}\epsilon^4+S_{24}\epsilon^2+S_{04}. \quad (77)
\end{aligned}$$

In model form 3 we list the values of $V_t$ as calculated by the foregoing expressions and also, as in model form 2, the values of the successive orthogonal polynomials $b_tV_t$ and the successive least-squares' polynomials $u_j(\epsilon)$. But since there are no residuals in the extrapolated region, the five columns of $v_j(\epsilon)$ values of model form 2 no longer appear.

The method of calculating $u_j(\epsilon)$ values given by Eq. (22) and used in model form 2, is certainly the most rapid one for values of $\epsilon$ corresponding to the observations. But in the case of other values of $\epsilon$, whether interpolated or extrapolated, we may, in place of model form 3, calculate the coefficients $a_{kj}$ of the power series in $\epsilon$ by means of Eq. (27), as will be done presently (Section F5) and then evaluate the function, Eq. (9), for the desired value of $\epsilon$. Actually these two processes give the values of $u_j(\epsilon)$ in terms of the $S_{kt}$ of Table XII and differ only in the order in which the various algebraic processes are performed.

### F4. Values of Weights and Probable Errors of Calculated Points, for Both the Observations and the Region of Extrapolation. (Model Forms 4 and 5)

We now consider the important question of the weights and probable errors of the $u_j(\epsilon)$ values already calculated. Later we will calculate the coefficients and their probable errors, of the various polynomials in power series form. Most of these equations for calculating weights and probable errors are believed to be new. They are based on the important fact, proved in the Appendix (Section J), that because of the orthogonal character of the solution, the quantities $b_t$ act like *independently observed* quantities to which we can apply the law of propagation of errors.

The relations needed for calculating the weights and probable errors of $u_j(\epsilon)$ are covered in Eqs. (31), (32), (33), (34), and (36) of Section C. The weight of $b_t$ is $N_t(=\sum_\epsilon V_t^2)$ and the values of $N_t$ for $n=7$, have already been used in model form 1. Then, applying propagation of errors—Eq. (32)—to Eq. (22), we get for the weight $p_j(\epsilon)$ of $u_j(\epsilon)$ the result already given in Eq. (36), namely

$$1/p_j(\epsilon)=\sum_{t=0}^{j}(V_t^2/N_t). \tag{36}$$

Using the values of $V_t$ already listed in model forms 1 and 3, for the several desired values of $\epsilon$ and the values of $N_t$ from model form 1, we obtain the results given in model form 4. As usual, in the process of getting values of $p_4(\epsilon)$ we also get weights for all polynomials of lower degree. The weight is obviously the same for equal plus and minus values of $\epsilon$ since only the square of $V_t$ is involved and, hence, we need to list only $|\epsilon|(=r)$, just as is done in model form 1.

Equation (36) shows that we are concerned with a sum of reciprocals of weights and, furthermore, it is convenient to calculate only the reciprocal $1/p_j(\epsilon)$ since from Eq. (33) multiplication of $[1/p_j(\epsilon)]^{\frac{1}{2}}$ by $r'$, the probable error for unit weight, gives us the final desired probable error. In the case of reciprocals of weights we encounter for the first time more complex repeating decimals. This result is inevitable for the case of 7 observations since $1/p_0(\epsilon)=1/n=1/7$. The results given in model form 4 are carried to 5 or 6 decimals in order to show the precise relations, but in all practical work involving probable errors, slide-rule accuracy is of course fully sufficient.

To get the probable errors $r_j(\epsilon)$ for the several $u_j(\epsilon)$ values, whose weights $p_j(\epsilon)$ are listed in model form 4, we need first the probable error $r_j'$ of a hypothetical quantity of *unit* weight, for a

Model Form 3. Extrapolated values of $u_j(\epsilon)$. (Values of $V_t$ calculated by Eq. (77), with values of $S_{kt}$ for $n=7$ from Table XII.) $u_0(\epsilon)=b_0V_0=b_0=180.15$.

| $\epsilon$ | $V_1$ | $b_1V_1 = 114.24V_1$ | $u_0(\epsilon)+b_1V_1 = u_1(\epsilon)$ | $V_2$ | $b_2V_2 = 45.475V_2$ | $u_1(\epsilon)+b_2V_2 = u_2(\epsilon)$ |
|---|---|---|---|---|---|---|
| −6 | −6 | −685.44 | −505.29 | 32 | +1455.200 | 949.91 |
| −5 | −5 | −571.20 | −391.05 | 21 | +954.975 | 563.92 |
| −4 | −4 | −456.96 | −276.81 | 12 | +545.700 | 268.89 |
| +4 | +4 | +456.96 | +637.11 | 12 | +545.700 | 1182.81 |
| +5 | +5 | +571.20 | +751.35 | 21 | +954.975 | 1706.32 |
| +6 | +6 | +685.44 | +865.59 | 32 | +1455.200 | 2320.79 |

| $\epsilon$ | $V_3$ | $b_3V_3 = 75.6V_3$ | $u_2(\epsilon)+b_3V_3 = u_3(\epsilon)$ | $V_4$ | $b_4V_4 = 3.6V_4$ | $u_3(\epsilon)+b_4V_4 = u_4(\epsilon)$ | $x$ |
|---|---|---|---|---|---|---|---|
| −6 | −29 | −2192.4 | −1242.49 | 561 | 2019.6 | +777.11 | −15 |
| −5 | −15 | −1134.0 | −570.08 | 231 | 831.6 | +261.52 | −10 |
| −4 | −6 | −453.6 | −184.71 | 66 | 237.6 | +52.89 | −5 |
| +4 | +6 | +453.6 | +1636.41 | 66 | 237.6 | +1874.01 | +35 |
| +5 | +15 | +1134.0 | +2840.32 | 231 | 831.6 | +3671.92 | +40 |
| +6 | +29 | +2192.4 | +4513.19 | 561 | 2019.6 | +6532.79 | +45 |

MODEL FORM 4. Weights $p_j(\epsilon)$ of calculated values $u_j(\epsilon)$. $1/p_0(\epsilon)=1/N_0=1/n=1/7$ here.

| $\|\epsilon\|$ | $V_1^2/N_1$ | $1/N_0 + V_1^2/N_1 = 1/p_1(\epsilon)$ | $V_2^2/N_2$ | $1/p_1(\epsilon) + V_2^2/N_2 = 1/p_2(\epsilon)$ |
|---|---|---|---|---|
| 6 | 1.285714 | 1.428571 | 12.19047 | 13.61904 |
| 5 | 0.892857 | 1.035714 | 5.25 | 6.28571 |
| 4 | 0.571428 | 0.714285 | 1.714285 | 2.42857 |
| 3 | 0.321428 | 0.464285 | 0.297619 | 0.76190 |
| 2 | 0.142857 | 0.285714 | 0 | 0.28571 |
| 1 | 0.035714 | 0.178571 | 0.10714 | 0.28571 |
| 0 | 0 | 0.142857 | 0.19047 | 0.33* |

| $\|\epsilon\|$ | $V_3^2/N_3$ | $1/p_2(\epsilon) + V_3^2/N_3 = 1/p_3(\epsilon)$ | $V_4^2/N_4$ | $1/p_3(\epsilon) + V_4^2/N_4 = 1/p_4(\epsilon)$ |
|---|---|---|---|---|
| 6 | 140.16* | 153.7856* | 2043.6428 | 2197.428 |
| 5 | 37.50 | 43.7857 | 346.5 | 390.286 |
| 4 | 6 | 8.42857 | 28.2857 | 36.714 |
| 3 | 0.16* | 0.92857 | 0.05844 | 0.987 |
| 2 | 0.16* | 0.45238 | 0.31818 | 0.7706 |
| 1 | 0.16* | 0.45238 | 0.00649 | 0.4589 |
| 0 | 0 | 0.33* | 0.23376 | 0.5671 |

* Singly repeating decimals are indicated by a star. Thus 140.16* = 140.16666 ⋯.

polynomial of degree $j$, where $j=0$ to 4. The needed value of $r_j'$ is given by Eq. (34) of Section C, in which the number of undetermined coefficients $s$ is $j+1$. Thus

$$r_j'=0.6745\left[\frac{\sum [v_j(\epsilon)]^2}{n-(j+1)}\right]^{\frac{1}{2}}. \qquad (78)$$

The values of $\sum_\epsilon [v_j(\epsilon)]^2$ have been calculated from Eq. (30), immediately following model form 1. Using these values in Eq. (78) we obtain, for $n=7$,

$$r_0'=208.89, \quad r_1'=138.23, \quad r_2'=64.25,$$
$$r_3'=17.397, \quad r_4'=0.07249.$$

Then, from Eq. (33),

$$r_j(\epsilon)=r_j'[1/p_j(\epsilon)]^{\frac{1}{2}}. \qquad (79)$$

Model form 5 gives the resulting values of $r_j(\epsilon)$ for the calculated values $u_j(\epsilon)$ of each polynomial, as listed in model forms 2 and 3, with the use of the values of $p_j(\epsilon)$ listed in model form 4. Again it is to be recalled that the $u_j(\epsilon)$ for the same positive and negative value of $\epsilon$ in model forms 2 and 3, have the same weight and probable error, and hence in both model forms 4 and 5 we need list only the values of $|\epsilon|$. In these two cases we use $|\epsilon|$ rather than the equivalent symbol $r$ in order to avoid any possible confusion with the symbol for probable error.

The values of $r_j(\epsilon)$, the probable error of the function, in model form 5 are worthy of careful study. As one proceeds into the region of extrapolation the predominating term in $1/p_j(\epsilon)$ becomes $V_j^2/N_j$ and the predominating term in $V_j$, for sufficiently large $\epsilon$, is $\epsilon^j$. Hence the resulting value of $r_j(\epsilon)$, which depends on $[1/p_j(\epsilon)]^{\frac{1}{2}}$, varies as $\epsilon^j$. This rule is very approximate, as shown by an examination of the actual expressions for $V_t$ in Eq. (77) in connection with the values of $S_{kt}$ in Table XII. But it is important to notice that in a rough sort of way the probable errors of a first-degree polynomial, as one proceeds into the region of extrapolation, vary linearly with $\epsilon$, those for a second-degree polynomial vary as the square of $\epsilon$, etc. It is just this fact that causes the great uncertainty of extrapolated values when calculated by means of a polynomial of high degree.

On the other hand, within the region of the data (here $\epsilon=-3$ to $+3$) the probable error remains more or less constant with *small* maxima

MODEL FORM 5. Probable errors $r_j(\epsilon)$ for $u_j(\epsilon)$ values of model forms 2 and 3. Weights $p_j(\epsilon)$ in model form 4. Values of $r_j'$ following Eq. (78). $r_0(\epsilon)=208.89(1/7)^{\frac{1}{2}}=78.95$ = constant.

| $\|\epsilon\|$ | $[1/p_1(\epsilon)]^{\frac{1}{2}}$ | $r_1'[1/p_1(\epsilon)]^{\frac{1}{2}} = r_1(\epsilon)$ | $[1/p_2(\epsilon)]^{\frac{1}{2}}$ | $r_2'[1/p_2(\epsilon)]^{\frac{1}{2}} = r_2(\epsilon)$ |
|---|---|---|---|---|
| 6 | 1.1952 | 165.21 | 3.690 | 237.08 |
| 5 | 1.0177 | 140.67 | 2.507 | 161.07 |
| 4 | 0.8451 | 116.82 | 1.5584 | 100.13 |
| 3 | 0.6814 | 94.19 | 0.8729 | 56.084 |
| 2 | 0.5345 | 73.89 | 0.5345 | 34.341 |
| 1 | 0.4226 | 58.42 | 0.5345 | 34.341 |
| 0 | 0.3780 | 52.25 | 0.5774 | 37.095 |

| $\|\epsilon\|$ | $[1/p_3(\epsilon)]^{\frac{1}{2}}$ | $r_3'[1/p_3(\epsilon)]^{\frac{1}{2}} = r_3(\epsilon)$ | $[1/p_4(\epsilon)]^{\frac{1}{2}}$ | $r_4'[1/p_4(\epsilon)]^{\frac{1}{2}} = r_4(\epsilon)$ |
|---|---|---|---|---|
| 6 | 12.401 | 215.740 | 46.88 | 3.398 |
| 5 | 6.617 | 115.116 | 19.75 | 1.432 |
| 4 | 2.903 | 50.503 | 6.06 | 0.439 |
| 3 | 0.9636 | 16.764 | 0.9935 | 0.0720 |
| 2 | 0.6726 | 11.701 | 0.8778 | 0.0636 |
| 1 | 0.6726 | 11.701 | 0.6774 | 0.0491 |
| 0 | 0.5774 | 10.045 | 0.7531 | 0.0546 |

and minima values. In fact, for a polynomial of degree $j$ there are, within the range of observations, $j$ minima and $j-1$ maxima values of the probable error of the function, all symmetrically located about $\epsilon=0$. The center point, $\epsilon=0$, corresponds to a small maximum of probable error for $j$ = even integer and to a small minimum for $j$ = odd integer.[33] One notes in this connection that the "center of mass" of the observed points is located at $\epsilon=0$, $y=b_0=\sum y/n$.

As a final remark on the probable errors in the region of extrapolation, I note that *in addition* to the fact that they are large, they have no real significance *unless* there is reason to believe that a polynomial of the $j$th degree is a valid function in the region considered. Thus if one, from theoretical considerations, believes that a linear relation actually exists even out to a certain extrapolated point, then one can use the value of the point thus calculated and can take its calculated probable error as a trustworthy measure of its uncertainty. But if the true but unknown function deviates from linearity in the extrapolated region, obviously the entire process breaks down. For just this reason, when polynomials are used as purely *empirical* functions for the *smoothing* of data, extrapolated values, and their probable errors become completely meaningless even though calculated correctly. The importance of this point cannot be overemphasized.

### F5. Coefficients of Power Series

We have now completed the discussion of the results that can be obtained very quickly from the data given in model form 1. These results have included the calculated values and probable errors of the *function* expressed as a polynomial of any degree up to the fourth (in the actual problem treated), at points corresponding to the observations as well as at certain extrapolated points.

The remaining desired information consists of (1) the *coefficients* $a_{kj}$ of the function expressed as a *power series* in $\epsilon$, and the coefficients $a_{kj}''$ of the corresponding power series in $x$, and (2) the probable errors of all such coefficients. In the usual treatment of orthogonal polynomials in the literature no rapid method is given for obtaining the results of (1) and, as already stated, it is believed that the method here presented of getting the probable errors of the coefficients is entirely new. In the present section we treat the coefficients. In Section F6 we shall consider the probable errors.

The simplest coefficients to obtain are those for the power series in $\epsilon$. These are given by Eq. (27), i.e.,

$$a_{kj}=\sum_{t=k}^{j} S_{kt}b_t. \tag{27}$$

As usual, due to the orthogonal character of the solution, it is possible to obtain the coefficients of all lower degree polynomials at the same time that one obtains those for the desired polynomial. For convenience, the specific expressions for all $a_{kj}$ coefficients up to $j=4$ are now given together with the numerical results for our sample problem. The needed values of $b_t$ from model form 1, and those of $S_{kt}$ (for $n=7$) from Table XII, are[34]

$$b_0=180.15,\quad b_1=114.24,\quad b_2=45.475,\quad b_3=75.6,\quad b_4=3.6,$$

$$S_{00}=1 \text{ (always)},\quad S_{11}=1,\quad S_{02}=-4,\quad S_{22}=1,\quad S_{33}=0.16^*,$$

$$S_{13}=-1.16^*,\quad S_{44}=0.583^*,\quad S_{24}=-5.583^*,\quad S_{04}=6.$$

The various $a_{kj}$ values are then, by Eq. (27),

$$\begin{aligned}
&a_{00}=a_{01}=S_{00}b_0=180.15,\\
&a_{02}=a_{03}=S_{00}b_0+S_{02}b_2=-1.75,\\
&a_{04}(=a_{05})=S_{00}b_0+S_{02}b_2+S_{04}b_4=19.85,\\
&a_{11}=a_{12}=S_{11}b_1=114.24,\\
&a_{13}=a_{14}=S_{11}b_1+S_{13}b_3=26.04,\\
&a_{22}=a_{23}=S_{22}b_2=45.475,\\
&a_{24}(=a_{25})=S_{22}b_2+S_{24}b_4=25.375,\\
&a_{33}=a_{34}=S_{33}b_3=12.6,\\
&a_{44}(=a_{45})=S_{44}b_4=2.1.
\end{aligned} \tag{80}$$

All of these results are exact. In fact, as stated earlier, the experimental data have been de-

[33] Using the least-squares' solution of polynomials by means of determinants, Schultz, footnote 31, pp. 155–160, discusses these facts about the probable error of a polynomial in the interpolated and extrapolated regions. He gives the explicit mathematical expressions for $j=1$, 2 and 3, needed to locate on the $\epsilon$ axis the various maxima and minima of probable error.

[34] As noted in model form 4, a star denotes a repeating decimal digit. Thus $0.16^*=0.16666\cdots$.

liberately chosen to make such simple exact results possible. The successive least-squares' solutions, corresponding to Eq. (9), are then

$$\begin{aligned} u_0(\epsilon) &= +180.15, \\ u_1(\epsilon) &= +180.15+114.24\epsilon, \\ u_2(\epsilon) &= -1.75+114.24\epsilon+45.475\epsilon^2, \\ u_3(\epsilon) &= -1.75+26.04\epsilon+45.475\epsilon^2+12.6\epsilon^3, \\ u_4(\epsilon) &= +19.85+26.04\epsilon+25.375\epsilon^2 \\ &\qquad +12.6\epsilon^3+2.1\epsilon^4. \end{aligned} \tag{81}$$

One does not ordinarily have any real need for an analytic expression in terms of $\epsilon$ in power-series form since its calculated values can be more rapidly obtained by the methods already presented (model forms 2 and 3). But the coefficients of the corresponding $f(x)$ may have real theoretical interest in physical science, and we now proceed to calculate the coefficients $a_{kj}''$ of this new function, where $\epsilon=(x-m)/h$, as used throughout the paper. The necessary equation, as given in Section C5, is

$$a_{kj}''=\sum_{t=k}^{j} H_{kt}\cdot b_t, \tag{42}$$

and explicit expressions for the $H_{kt}$ are given in connection with Eqs. (42:0) to (42:5) in Section C5. We shall not repeat these expressions, but shall merely give the resulting values of $H_{kt}$ for the present problem.

From Table VII, in Section F1, we find

$$\begin{aligned} h &= 5 = \text{interval between values of } x, \\ m &= x_0 = +15 \\ &= \text{“middle” value of } x \text{ (for which } \epsilon=0). \end{aligned}$$

Then from the expressions for $H_{kt}$ just mentioned, and with the values of $S_{kt}$ used in Eq. (80), we get

$$\begin{aligned} &H_{00}=1, \quad H_{01}=-3, \quad H_{02}=+5, \\ &\qquad\qquad H_{03}=-1, \quad H_{04}=+3, \\ &H_{11}=0.2, \quad H_{12}=-1.2, \\ &\qquad\qquad H_{13}=+2/3, \quad H_{14}=-5.9, \\ &H_{22}=0.04, \quad H_{23}=-0.06, \quad H_{24}=+311/300, \\ &H_{33}=1/750, \quad H_{34}=-7/125, \quad H_{44}=7/7500. \end{aligned}$$

All of these values of $H_{kt}$ are exact.

The values of $a_{kt}''$ are given by the successive partial sums of Eqs. (42:0) to (42:5). These results are now listed in Table IX. The first row contains the designation of the successive terms of these equations, and the second row contains the *value* of the successive *partial sums* of the terms, yielding the value of the coefficient stated in the third row under each such partial sum. Hence, from Eq. (39), the successive least-squares' solutions expressed as power series in $x$ are

$$\begin{aligned} u_0(x) &= +180.15, \\ u_1(x) &= -162.57+22.848x, \\ u_2(x) &= +64.805-31.722x+1.819x^2, \\ u_3(x) &= -10.795+18.678x \\ &\qquad -2.717x^2+0.1008x^3, \\ u_4(x) &= +0.005-2.562x+1.015x^2 \\ &\qquad -0.1008x^3+0.00336x^4. \end{aligned} \tag{82}$$

As in the case of the corresponding Eq. (81), all equations in (82) are exact.

## F6. Probable Errors of Coefficients of Power Series

Our final problem is to calculate the weights and probable errors of the coefficients $a_{kj}$ of Eq.

TABLE IX. Values of $a_{kj}''$.

| (1) Terms | $H_{00}b_0$ | $H_{01}b_1$ | $H_{02}b_2$ | $H_{03}b_3$ | $H_{04}b_4$ |
|---|---|---|---|---|---|
| (2) Sums | +180.15 | −162.57 | +64.805 | −10.795 | +0.005 |
| (3) Coefficients | $a_{00}''$ | $a_{01}''$ | $a_{02}''$ | $a_{03}''$ | $a_{04}''$ |

| (1) Terms | $H_{11}b_1$ | $H_{12}b_2$ | $H_{13}b_3$ | $H_{14}b_4$ |
|---|---|---|---|---|
| (2) Sums | +22.848 | −31.722 | +18.678 | −2.562 |
| (3) Coefficients | $a_{11}''$ | $a_{12}''$ | $a_{13}''$ | $a_{14}''$ |

| (1) $H_{22}b_2$ | $H_{23}b_3$ | $H_{24}b_4$ | (1) $H_{33}b_3$ | $H_{34}b_4$ | (1) $H_{44}b_4$ |
|---|---|---|---|---|---|
| (2) +1.819 | −2.717 | +1.015 | (2) +0.1008 | −0.1008 | (2) +0.00336 |
| (3) $a_{22}''$ | $a_{23}''$ | $a_{24}''$ | (3) $a_{33}''$ | $a_{34}''$ | (3) $a_{44}''$ |

(81) and of the coefficients $a_{kj}''$ of Eq. (82). The weights $p_{kj}$ of the coefficients $a_{kj}$ are given by Eq. (38), namely,

$$1/p_{kj}=\sum_{t=k}^{j}(S_{kt}^2/N_t). \qquad (38)$$

The required values of $S_{kt}$ (from Table XII) precede Eq. (80), and those of $N_t$ (from Table XIII) appear in model form 1. With these values we obtain, from Eq. (38), the results listed in Table X.

In ordinary work all of these results would naturally be expressed in decimals, with only slide-rule accuracy necessary. But in the foregoing expressions the exact results are recorded.

With the foregoing values of $p_{kj}$, the probable errors $r_{kj}$ of the coefficients $a_{kj}$ are to be obtained by the use of Eq. (33), which now takes the specific form

$$r_{kj}=r_j'(1/p_{kj})^{\frac{1}{2}} \qquad (83)$$

in place of the special form Eq. (79) used in connection with model form 5. The numerical values of $r_j'$, the probable error for unit weight, have already been calculated from Eq. (78) and listed following that equation for use in model form 5. With these values one obtains, for example,

$$r_{13}=r_3'(1/p_{13})^{\frac{1}{2}}=17.397(397/1512)^{\frac{1}{2}}=8.9144,$$
$$r_{24}=r_4'(1/p_{24})^{\frac{1}{2}}=0.07249(679/3168)^{\frac{1}{2}}=0.03356.$$

Instead of listing separately the remaining values of $r_{kj}$, we insert them directly into Eq. (81) and thus get as our final power-series polynomials, expressed as a function of $\epsilon$,

$$\begin{aligned}
u_0(\epsilon)&=180.15\pm78.95,\\
u_1(\epsilon)&=(180.15\pm52.25)+(114.24\pm26.12)\epsilon,\\
u_2(\epsilon)&=-(1.75\pm37.09)+(114.24\pm12.14)\epsilon\\
&\quad+(45.475\pm7.010)\epsilon^2,\\
u_3(\epsilon)&=-(1.75\pm10.04)+(26.04\pm8.914)\epsilon\\
&\quad+(45.475\pm1.898)\epsilon^2\\
&\quad+(12.6\pm1.184)\epsilon^3,\\
u_4(\epsilon)&=(19.85\pm0.05459)\\
&\quad+(26.04\pm0.03714)\epsilon\\
&\quad+(25.375\pm0.03356)\epsilon^2\\
&\quad+(12.6\pm0.004932)\epsilon^3\\
&\quad+(2.1\pm0.003407)\epsilon^4.
\end{aligned} \qquad (84)$$

As in the case of Eq. (81), the coefficients $a_{kj}$ of Eq. (84) are exact. All probable errors are here recorded to four significant figures although in practical work two significant figures in the probable errors, with the corresponding two doubtful figures in the quantities themselves, are always quite sufficient. Since the value of $u_j(\epsilon)$ at $\epsilon=0$ is merely $a_{0j}$, the respective probable errors $r_{0j}$ in Eq. (84) are also the probable errors of the function at $\epsilon=0$, which have already been given in model form 5 in the row $\epsilon=0$ and in the respective $r_j(\epsilon)$ columns.

TABLE X. Values of $p_{kj}$.

$$\begin{aligned}
&1/p_{00}=1/p_{01}=S_{00}/N_0=1/7\\
&1/p_{02}=1/p_{03}=1/p_{00}+S_{02}^2/N_2=1/3\\
&1/p_{04}(=1/p_{05})=1/p_{02}+S_{04}^2/N_4=131/231\\
&1/p_{11}=1/p_{12}=S_{11}/N_1=1/28\\
&1/p_{13}=1/p_{14}=1/p_{11}+S_{13}^2/N_3=397/1512\\
&1/p_{22}=1/p_{23}=S_{22}/N_2=1/84\\
&1/p_{24}(=1/p_{25})=1/p_{22}+S_{24}^2/N_4=679/3168\\
&1/p_{33}=1/p_{34}=S_{33}^2/N_3=1/216\\
&1/p_{44}(=1/p_{45})=S_{44}^2/N_4=7/3168
\end{aligned}$$

We now proceed to calculate the weights and probable errors of the coefficients $a_{kj}''$ of the various $u_j(x)$ listed in Eq. (82). The weights are given by Eq. (43), namely,

$$1/p_{kj}''=\sum_{t=k}^{j}(H_{kt}^2/N_t). \qquad (43)$$

The required values of $H_{kt}$ have already been calculated, from Eqs. (42:0) to (42:5), in connection with the evaluation of the coefficients $a_{kj}''$ of Eq. (82). The values of $N_t$ have already been used in obtaining the weights $p_{kj}$ of Table X. They appear in model form 1 and are taken from Table XIII for $n=7$.

As in the case of the $a_{kj}''$ coefficients, a tabular arrangement (Table XI) is used for the successive partial sums given by Eq. (43). The third row gives the designation of the numerical values in the second row.

Just as in the case of the values of $p_{kj}$, it is far simpler and quicker to express all of the foregoing weights as decimals, with only slide-rule accuracy necessary. The exact values are here recorded just to show their simplicity or complexity, as the case may be.

The final calculation is that of the probable errors $r_{kj}''$ of the $a_{kj}''$ coefficients. Here we again use Eq. (83) with $p_{kj}$ replaced by $p_{kj}''$. We also, as in the case of $r_{kj}$, do not list separately the resulting $r_{kj}''$ values but merely insert them in the $u_j(x)$ of Eq. (82). We thus obtain

$$\begin{aligned}
u_0(x) &= 180.15 \pm 78.95,\\
u_1(x) &= -(162.57 \pm 94.19)\\
&\quad + (22.848 \pm 5.225)x,\\
u_2(x) &= (64.805 \pm 56.08)\\
&\quad - (31.722 \pm 8.756)x\\
&\quad + (1.819 \pm 0.2804)x^2,\\
u_3(x) &= -(10.795 \pm 16.76)\\
&\quad + (18.678 \pm 5.295)x\\
&\quad - (2.717 \pm 0.4328)x^2\\
&\quad + (10.08 \pm 0.9470)\cdot 10^{-2}x^3,\\
u_4(x) &= (5.0 \pm 72.02)\cdot 10^{-3}\\
&\quad - (2.562 \pm 0.04092)x\\
&\quad + (1.015 \pm 0.006318)x^2\\
&\quad - (10.08 \pm 0.03295)\cdot 10^{-2}x^3\\
&\quad + (33.60 \pm 0.05452)\cdot 10^{-4}x^4.
\end{aligned} \tag{85}$$

In Eqs. (85), just as in Eqs. (84), all coefficients are exact, and the probable errors are recorded to four significant figures, although no more than two figures are of any significance in experimental work.

A portion of the probable errors in Eqs. (85) has already appeared in model form 5, just as in the case of Eqs. (84). In fact, the probable errors $r_{0j}''$ of the absolute term in Eqs. (85) are just the probable errors of the function at $x=0$ or $\epsilon=-3$, and are given in the row $|\epsilon|=3$ of model form 5.

### F7. Calculation of $a_t$ Values from Finite Differences. (Model Form 6)

In model form 1 we have calculated values of $b_t$ directly from the observations. From Eq. (27) we can then get $a_{tt}$ (normally denoted $a_t$) $=S_{tt}b_t$, where $a_t$ is the coefficient of $\epsilon^t$ in a polynomial of degree $t$. It is also the coefficient for any other abscissa scale for which the observations are spaced at unit interval, since a shift in the absolute value of the abscissa does not affect the coefficient of the highest degree term in a power series.

In Section D we have shown that $t!a_t$ can be considered as the weighted average of the finite differences $\delta^t y$, and Eq. (54) expresses such a weighted average. The necessary weights $W_t$ and sums of weights $L_t$ are listed in Table XIV. We shall now apply Eq. (54) to our sample problem.

TABLE XI. Values of $1/p_{kj}''$.

| (1) Terms | $H_{00}^2/N_0$ | $H_{01}^2/N_1$ | $H_{02}^2/N_2$ | $H_{03}^2/N_3$ | $H_{04}^2/N_4$ |
|---|---|---|---|---|---|
| (2) Sums | 1/7 | 13/28 | 16/21 | 13/14 | 76/77 |
| (3) $1/p_{kj}''$ | $1/p_{00}''$ | $1/p_{01}''$ | $1/p_{02}''$ | $1/p_{03}''$ | $1/p_{04}''$ |

| (1) $H_{11}^2/N_1$ | $H_{12}^2/N_2$ | $H_{13}^2/N_3$ | $H_{14}^2/N_4$ |
|---|---|---|---|
| (2) 1/700 | 13/700 | 1751/18900 | 132509/415800 |
| (3) $1/p_{11}''$ | $1/p_{12}''$ | $1/p_{13}''$ | $1/p_{14}''$ |

| (1) $H_{22}^2/N_2$ | $H_{23}^2/N_3$ | $H_{24}^2/N_4$ |
|---|---|---|
| (2) 1/52500 | 13/21000 | 15043/1,980,000 |
| (3) $1/p_{22}''$ | $1/p_{23}''$ | $1/p_{24}''$ |

| (1) $H_{33}^2/N_3$ | $H_{34}^2/N_4$ | (1) $H_{44}^2/N_4$ |
|---|---|---|
| (2) 1/3,375,000 | 767/37,125,000 | (2) 7/1,237,500,000 |
| (3) $1/p_{33}''$ | $1/p_{34}''$ | (3) $1/p_{44}''$ |

The necessary finite differences of the observations are given in Table VII of Section F1.

Since the weights are always symmetrical about the center point, the finite differences can be combined in pairs, as shown in Eq. (54), i.e.,

$$t!a_t = \frac{1}{L_t} \sum_{r=0 \text{ or } \frac{1}{2}}^{\nu} (\delta^t y_r + \delta^t y_{-r}) W_t, \tag{54}$$

but with a small number of $\delta^t y$ values it is simpler to list the entire set of values.[35] With the values

[35] As stated in Section D, the values of the "pair-factors" $V_t$, listed in Table XIII and used in model form 1, may be proved (see Section J) to be merely the finite differences of order $t$ of the "weights" $W_t$, listed in Table XIV and now used in model form 6. Historically, the values of $V_t$ were first calculated by Birge and Shea (footnote 3) from the equations derived by them, and then, many years later, the values of $W_t$ were calculated from the equations of Section D. Actually, however, the simplest way to obtain the $V_t$ values is by merely differencing the $W_t$ values, since the equations for $W_t$ are simpler to handle than those for $V_t$.

In order to obtain the necessary $n$ values of $V_t$ for a set of $n$ observations from the available $n-t$ values of $W_t$, we attach $t$ zeros at either end of the set of $W_t$ values. The justification for this procedure lies in the fact that $W_t(\epsilon)$ is a polynomial of degree $2t$, whose $2t$ roots occur at unit intervals on either side of the region of observation, immediately adjacent to it. There are thus *no* roots within the region of observation (i.e., all values of $W_t$ listed in Table XIV are positive). Furthermore, since the values of $V_t$ are the finite differences of order $t$ of the $W_t$ values, it follows that $V_t(\epsilon)$ is a polynomial of degree $t$, all of whose roots occur *within* the range of the observations, so that each complete set of $n$ values of $V_t$ (*not* the "half-set" listed in Table XIII, for use in connection with the observations arranged in pairs) changes sign $t$ times.

As an illustration of the above facts, let us obtain the values of $V_2$ for $n=7$, listed in model form 1, from the values of $W_2$ for $n=7$, listed in model form 6, and similarly the values of $V_1$ from those of $W_1$. In all cases the coefficients are listed in *descending* order of $\epsilon$, and as stated in footnote

MODEL FORM 6. Values of $a_t$ from finite differences. $a_0 = b_0 = \Sigma\, y/n = 180.15$ as in Eq. (80).

| $\epsilon$ | | $\delta y_\epsilon$ | $W_1$ | $\delta^2 y_\epsilon$ | $W_2$ | $\delta^3 y_\epsilon$ | $W_3$ | $\delta^4 y_\epsilon$ | $W_4$ |
|---|---|---|---|---|---|---|---|---|---|
| 5/2 | | 528.78 | 3 | | | | | | |
| | 2 | | | 306.81 | 5 | | | | |
| 3/2 | | 221.97 | 5 | | | 150.78 | 1 | | |
| | 1 | | | 156.03 | 10 | | | 49.35 | 3 |
| 1/2 | | 65.94 | 6 | | | 101.43 | 2 | | |
| | 0 | | | 54.60 | 12 | | | 51.66 | 5 |
| −1/2 | | 11.34 | 6 | | | 49.77 | 2 | | |
| | −1 | | | 4.83 | 10 | | | 49.35 | 3 |
| −3/2 | | 6.51 | 5 | | | 0.42 | 1 | | |
| | −2 | | | 4.41 | 5 | | | | |
| −5/2 | | 2.10 | 3 | | | | | | |
| | | $L_t=$ 28 | | 42 | | 6 | | 11 | |
| | | $\Sigma\, \delta^t y \cdot W_t =$ 3198.72 | | 3819.90 | | 453.60 | | 554.40 | |
| | | $t!a_t=$ 114.24 | | 90.95 | | 75.60 | | 50.40 | |
| | | $a_t=$ 114.24 | | 45.475 | | 12.60 | | 2.10 | |

of $W_t$ and $L_t$ for $n=7$, from Table XII, we get the values of $a_t$ as shown in model form 6. These values of $a_t$ check[36] with the final coefficients of the various $u_j(\epsilon)$ functions of Eq. (84).

As stated in Section D, one may also calculate the probable error $r_{jj}$ of $a_j(=a_{jj})$, but the process is far longer than that involved in evaluating $a_j$ itself. Suppose, for instance, that we wish to calculate $a_4$ and its probable error $r_{44}$. To get $a_4$ we use only the $\delta^4 y_\epsilon$ values and their weights $W_4$ in model form 6. One can also get $a_4$ with equal ease from model form 1 by the use of column 10, headed $V_4$, in connection with the data listed in columns 5 and 7. The quantity evaluated there is $b_4$ but $a_4$ follows immediately from Eq. (23) $(a_j = S_{jj} b_j)$.

All calculations of probable error already made in this paper have been in terms of the $b$–$V$ system, and Tables XII, XIII, XIV are designed specifically for use in that system. Hence if we have evaluated $a_4$ from finite differences and wish to get its probable error, we must convert all values, where necessary, to the $b$–$V$ system.

28, one thus gets the values of $+V_2$ (for which $t$ is even), but of $-V_1$ (for which $t$ is odd).

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $W_2$ | 0 | | 0 | | 5 | | 10 | | 12 | | 10 | | 5 | | 0 | | 0 |
| First difference | | 0 | | −5 | | −5 | | −2 | | +2 | | +5 | | +5 | | 0 | |
| $V_2$ | | | +5 | | 0 | | −3 | | −4 | | −3 | | 0 | | +5 | | |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $W_1$ | 0 | | 3 | | 5 | | 6 | | 6 | | 5 | | 3 | | 0 |
| $-V_1$ | | −3 | | −2 | | −1 | | 0 | | +1 | | +2 | | +3 | |

[36] It should be noted also that the respective values of $\Sigma\, \delta^t y \cdot W_t$ in model form 6 are identical with the corresponding values of $\Sigma\, y \cdot V_t$ in model form 1, in agreement with Eq. (50″) of Section D.

The first quantity needed for the evaluation of $r_{44}$ is $\sum_\epsilon [v_4(\epsilon)]^2$, and as Table VIII shows in detail, we need for this purpose not only the value of $b_4$ (as gotten from $a_4$) but also all other $b_t$ values for $t$ less than $j(=4)$. Hence, the complete model form 6 must be calculated up to $t=4$. The $N_t$ values of Table VIII, or Eq. (30), are taken as usual from Table XIII. This calculation of $\sum v_j^2$, as Table VIII shows, must in general be carried out with great accuracy if the final result is to be at all reliable.

One next follows the process outlined in Section F6. The weight $p_{jj}$ of the final coefficient $a_{jj}$ of a polynomial of degree $j$ has the specially simple form $N_j/S_{jj}^2$, as shown by Eq. (38) and Table X. Finally, the probable error $r_{jj}$ is given by Eq. (33) of Section C5, which now takes the special form

$$r_{jj} = r_j'(1/p_{jj})^{\frac{1}{2}}, \tag{86}$$

with $r_j'$ given by Eq. (78) of Section F4. For convenience all of the above steps, which involve Eqs. (23), (30), (38), and (86), may be combined in the following expression, which is explicit in the $a_t$ values that have presumably been obtained by model form 6.

$$r_{jj} = 0.6745\, S_{jj} \left\{ \frac{\sum_\epsilon [v_j(\epsilon)]^2}{N_j(n-j-1)} \right\}^{\frac{1}{2}}, \tag{87}$$

where, from Eqs. (30) and (23),

$$\sum_\epsilon [v_j(\epsilon)]^2 = \sum_\epsilon y^2 - \sum_{t=0}^{j} N_t (a_t/S_{tt})^2. \tag{88}$$

With the various values of $a_t$ calculated in model form 6, and with the values of $S_{tt}$ from Table XII and of $N_t$ from Table XIII (all for $n=7$), we get, from Eq. (88) with $j=4$, $\sum v_4^2 = +0.0231$, just as in Table VIII. Then from Eq. (87), $r_{44}=0.3407\times10^{-2}$. Hence, $a_{44}=2.10 \pm 0.0034_{07}$, checking the result already shown in the $u_4(\epsilon)$ function of Eq. (84), Section F6.

## F8. Summary

The explicit directions that have been given in Sections F1 to F7 may be summarized as follows.

(F1). A given set of equally spaced, unweighted data, such as that given in Table VII, is to be fitted to a polynomial of degree $j$, by the method of least squares. The original abscissa scale $x$ is replaced by a new scale $\epsilon$, in which $\epsilon$ proceeds by unit intervals from $-q$ to $+q$, where $n=2q+1$ = number of observations. Model form 1, in which $r=|\epsilon|$, then leads to the calculated values of $b_t$, where $t$ varies from 0 to $j$. The values of $V_t$ and $N_t$ appearing in model form 1 are taken directly from Table XIII for the proper value of $n$.

Knowing $b_t$ and $N_t$ one easily obtains the sum of the squares of the residuals (Eq. (30) and Table VIII) for each degree of polynomial $t$ from 0 to $j$, and one is thus able to decide on the proper degree to use (the fourth degree, for the data of Table VII).

(F2). The calculated value $u_j(\epsilon)$ of a least-squares' polynomial of degree $j$ corresponding to each and every observed $y(\epsilon)$ follows immediately, by Eq. (22), from the values of $b_t$ and $V_t$. Full details for the illustrative data of Table VII are given in model form 2, which contains also the value $v_j(\epsilon)$ of each corresponding residual, for polynomials of degree $j=0$ to 4.

A rearrangement of the terms of Eq. (22) leads to a very significant interpretation of the orthogonal polynomials $b_tV_t$ for successive values of $t$,—namely, each such orthogonal polynomial represents the least-squares' polynomial of degree $t$, fitted to the residuals $v_{t-1}(\epsilon)$ of the polynomial of one lower degree, considered as a set of observations.

(F3). For any value of $\epsilon$ not corresponding to an observation, whether extrapolated or interpolated, one cannot get numerical values of $V_t$ from Table VIII but must calculate them by Eq. (77) with the use of values of $S_{kt}$ listed in Table XII. Then Eq. (22) may be used, as before, to get each desired calculated value $u_j(\epsilon)$, as shown in detail in model form 3 for certain extrapolated points.

(F4). To determine the probable error of the *function* $u_j(\epsilon)$ at any value of $\epsilon$ we must first get the weight $p_j(\epsilon)$ of the function, by Eq. (36). The values of $V_t$ and $N_t$ for the observed values of $\epsilon$ already appear in model form 1, and the values of $V_t$ for extrapolated values of $\epsilon$ (with the same values of $N_t$ as before) in model form 3. Full details of the calculation of $p_j(\epsilon)$ appear in model form 4.

To get the probable error $r_j(\epsilon)$ corresponding to the weight $p_j(\epsilon)$, we need in addition to $p_j(\epsilon)$ only the probable error $r_j'$ of a hypothetical point of unit weight for the $j$th degree polynomial, as given by Eq. (78). Numerical values of $r_j'$ for $j=0$ to 4 follow that equation, and the resulting values of $r_j(\epsilon)$ are given by Eq. (79), as shown in detail in model form 5. There follows a brief description of the variation with $\epsilon$ of such probable errors, with emphasis on the rapid increase of the errors in the region of extrapolation.

(F5). In order to obtain the coefficients $a_{kj}$ of the *power series* $u_j(\epsilon)$, we use Eq. (27), as shown in detail in Eq. (80) with the numerical results given by Eq. (81). The necessary values of $S_{kt}$ in Eq. (80) are taken directly from Table XII for $n=7$.

To obtain the coefficients $a_{kj}''$ of the power series $u_j$ as a function of the original $x$ scale of abscissas (see Table VII of Section F1) we use Eq. (42) with the detailed expressions for $H_{kt}$ given in connection with Eqs. (42:0) to (42:5). The resulting numerical values of $H_{kt}$, for the standard illustrative problem, are given in Section F5 and are used in Table IX to calculate the $a_{kj}''$ coefficients. The final $u_j(x)$ are listed in Eq. (82).

(F6). The weights $p_{kj}$ of the coefficients $a_{kj}$ of the power series in $\epsilon$ are given by Eq. (38), for which detailed expressions are also listed with the resulting numerical values in Table X. The needed values of $S_{kt}$ are taken from Table XII, just as in Section F5. The probable errors $r_{kj}$ corresponding to the weights $p_{kj}$ are given by

Eq. (83). The numerical results $a_{kj} \pm r_{kj}$ are included in Eq. (84), which thus represents the final set of least-squares' solutions as $f(\epsilon)$ for $j=0$ to 4 of the standard illustrative problem.

The weights $p_{kj}''$ of the coefficients $a_{kj}''$ of the corresponding power series in $x$ are given by Eq. (43), with the numerical values listed in Table XI. Then Eq. (83) is again used, with $p_{kj}$ replaced by $p_{kj}''$, to obtain the probable errors $r_{kj}''$ corresponding to the weights $p_{kj}''$. The final power-series polynomials $u_j$ in terms of the original abscissa scale $x$ for $j=0$ to 4 with coefficients $a_{kj}'' \pm r_{kj}''$ are listed in Eq. (85). Exact values are recorded for all coefficients, and all probable errors and weights are here given to several more digits than have any significance in practical work.

(F7). In place of a solution explicitly in terms of the observations, as carried out in Sections F1 to F6, we may obtain the value of any single $a_t$, the coefficient of $\epsilon^t$ (*or* of $x^t$ if $x$ varies by *unit* intervals), of a polynomial of degree $t$ in terms of the finite differences $\delta^t y_\epsilon$. The necessary formula appears as Eq. (54), with numerical values of $W_t$ and $L_t$ listed in Table XIV. The actual calculation of the $a_t$ values, with $t=0$ to 4, for the standard data of Table VII, is given in model form 6.

To obtain the probable error $r_{jj}$ of $a_j(=a_{jj})$ in a polynomial of degree $j$, we must first calculate *each* $a_t$ for $t=0$ to $j$, as is done in model form 6, for $j=4$. Then Eqs. (87) and (88) give the explicit process for the calculation of $r_{jj}$. The necessary values of $S_{tt}$ appear in Table XII, and those of $N_t$ in Table XIII, for the specified number of observations.

Tables XII, XIII, and XIV, at the end of the paper, cover polynomials up to the fifth degree for any number of observations up to 30. The material of Table XIII only, up to $n=52$, appears in the reference of footnote 24, and up to $n=104$ in the reference of footnote 27.

## G. ALTERNATIVE PROCESSES

In Section A it is noted that much work has been done on the least-squares' fitting of polynomials to equally spaced data. A partial list of references on the subject is given in Section H. Any adequate account of these alternative processes would require in itself a paper far longer than the present one. We discuss here in detail merely two such alternative processes that lead to a solution in power-series form and hence are closely related to the methods just discussed. A few very general remarks are added on the factorial form of solution.

Near the beginning of Section C there has been presented a scheme for labeling the various proposed methods. In terms of that scheme the two alternative methods now to be discussed are (1) $A1\beta$, H. T. Davis,[11] and (2) $A3\beta$, Kerawala.[10]

### G1. H. T. Davis Method, $A1\beta$. (Model Form 7)

This method, as its designation indicates, leads to a result in power-series form making explicit use of power moments of the observations, but *not* making use of orthogonal polynomials. As will be demonstrated, the Davis method not only loses the vital advantages of the orthogonal solution, but is also a longer and less accurate method than that advocated in this paper and also than that proposed by Kerawala.[10] The original references for the Davis method are Davis and Latshaw[37] and Davis,[38] but full details,

[37] H. T. Davis and V. V. Latshaw, Annals Math. (2) **31**, 52 (1930).

[38] H. T. Davis, Annals Math. Stat. **4**, 154 (1933). In this article Davis tabulates the numerical factors needed for the $-q$ to $+q$ range of abscissa. He also discusses the so-called Gram polynomials, which are orthogonal polynomials in a special form first studied by J. P. Gram (J. f. Math. **94**, 41 (1883)). On page 158 of his paper Davis writes, in regard to Gram polynomials: "This method has since been more fully investigated by Edward Condon (Univ. of Calif. Pub. in Math. **2**, 55–66 (1927)) and his work was made the basis of a method for obtaining least squares polynomials by R. T. Birge and J. D. Shea. The work of the latter, however, while effecting a simplification, does not reduce the problem to its simplest form."

Both the facts and the implications of these remarks are so incorrect that the true circumstances should be noted. In the first place Condon states in his paper: "This investigation grew out of a desire to provide a more general basis for the work of Birge and Shea (Phys. Rev. **24**, 206(1924))." In other words, Condon's investigation was not carried out until some time after Birge and Shea had reported their work to the American Physical Society. Condon did revive and extend the work of Gram and he suggested the name "Gramian polynomials." But the chief purpose of his paper was to get a direct derivation of the formula for $a_t$, which had been obtained in an inelegant and laborious way by Birge and Shea (footnote 3). Condon, however, failed in this attempt, and the first direct derivation was obtained in 1942 by Weinberg (footnote 4).

In the second place, the Birge and Shea method, although possibly not in the simplest form from the standpoint of mathematical elegance, is certainly far simpler and far more accurate numerically than the method presented ten

MODEL FORM 7. Davis method.

| $r^3$ | $r$ | diff. | $y_{-q}\cdots$ | $r$ | $y_q\cdots$ | sum | $r^2$ | $r^4$ |
|---|---|---|---|---|---|---|---|---|
| 27 | 3 | 836.64 | 0 | 3 | 836.64 | 836.64 | 9 | 81 |
| 8 | 2 | 305.76 | 2.10 | 2 | 307.86 | 309.96 | 4 | 16 |
| 1 | 1 | 77.28 | 8.61 | 1 | 85.89 | 94.50 | 1 | 1 |
| 0 | 0 | — | | 0 | 19.95 | 19.95 | 0 | 0 |
| 25112.64 | 3198.72 | | | $m_t$ | $(=\Sigma_\epsilon\, \epsilon^t y)$ | 1261.05 | 8864.10 | 72821.70 |

with the complete numerical tables of factors, are given also in Davis' book.[11]

As is well known, the normal equations for the least-squares' solution of a polynomial of degree $j$ are formed by equating the calculated power moments, from $t=0$ to $j$, to the corresponding power moments of the observations, $\sum x^t y$. The calculated moments involve various $\sum x^i$, where $i$ runs from 0 to $2j$, multiplied by the coefficients $a_{kj}$ of the power series whose values are to be obtained by least squares. If the values of $x$ are equally spaced and at unit interval we thus need to know for the calculated moments only the sums of powers of successive integers. Davis[11] considers the solution for $x=-p$ to $+p$, where his $p$ is just our $q$ and will be so designated hereafter. He gives extensive tables of numerical factors (pp. 326–359 of his book). He *also* gives less complete tables of factors (pp. 370–385) for the case $x=1$ to $n$ (where $n=2q+1$). The latter set of tables is, of course, completely superfluous, since when one has a solution in the form $f(e)=f(x-x_0)$, the transformation to $f(x)$ can be made very quickly by a simple Horner shift (synthetic division), as illustrated in Section B.

For the case $x=-q$ to $+q$, i.e., $n=2q+1$ observations in all, the Davis tables run to $q=150$ for the first-degree polynomial ($j=1$), to $q=100$ for $j=2$, to $q=50$ for $j=3$, and to $q=25$ for $j=4$, 5, 6, and 7. In order to use our standard illustrative problem we will consider his fourth-degree solution which takes the form[39]

$$\begin{aligned} a_{04}&=Am_0+Bm_2+Cm_4,\\ a_{24}&=Bm_0+Dm_2+Em_4,\\ a_{44}&=Cm_0+Em_2+Fm_4,\\ a_{14}&=A'm_1+B'm_3,\\ a_{34}&=B'm_1+C'm_3. \end{aligned} \tag{89}$$

These $a_{kj}$ coefficients apply to Davis' $x=-q$ to $+q$. Hence his $x$ is identical with our $\epsilon$ and the coefficients of Eq. (89) are identical with the $a_{kj}$ coefficients of Eq. (9). The needed factors $A$, $B$, etc., whose numerical values are listed in the tables, are more or less complex functions of $n$, the number of observations.[40]

Davis gives the numerical values of the factors $A$, $B$, etc. as decimals to ten significant figures, and the accuracy of the solution is necessarily limited by this fact. In the method advocated in the present paper, all of the corresponding factors (the $S_{kt}$ of Table XII) appear as simple *terminating* decimals (or at the worst as thirds) and hence there is no limit on the resulting accuracy arising from this cause. This question of the accuracy of the solution as limited by the calculations themselves is very important, as will appear later.

The calculation of the power moments $m_t$ involves multiplication of the observations $y$ by the various values of $x^t$, and for $x(=\epsilon)=-q$ to $+q$ it is evident that we can shorten the process by combining the observations in sums and differences of pairs, just as has been done in model form 1. In fact, model form 1 can easily be adapted to the Davis method, although Davis himself suggests no such special form. Hence we

---

years later by Davis, as is shown in detail in the present section. This distinction between mathematical elegance and simplicity (including *accuracy*) of numerical calculation is emphasized more than once in the present paper just because it is so often misunderstood and hence ignored.

[39] Davis uses $M_t$ for the power moment. But since we have already employed this symbol in a different sense, in Eq. (7), we adopt $m_t$ for $\Sigma\, x^t y$.

[40] Davis gives these functions in terms of $q$ (=his $p$) and also uses many special symbols, so that it is difficult to visualize by inspection the exact dependence on the number of observations $n$. But Kerawala (footnote 10) does derive and list explicit $f(n)$, for $j=1$ to 5, as discussed in Section G2.

now proceed, by the use of our model form 7 and Eqs. (89), to obtain the $a_{kj}$ values for our standard illustrative set of data given in Table VII of Section F1. The corresponding solution, by our own method, consists in the calculation of the $b_t$ values in model form 1, and then the calculation of the $a_{kj}$ values by Eqs. (80) with the results given in Eqs. (81). But here, for brevity, we carry out only the fourth-degree solution. As usual, $r=|\epsilon|=|x|$ of Davis.

The value of $m_0$ is that under the "sum" column, and the other $m_t$ values lie in the respective $r^t$ columns. It should be noted first that the $m_t$ values are, in general, considerably larger than the $\sum y\cdot V_t$ values of model form 1, which are obtained by an exactly similar process, i.e., as an algebraic sum of products. The reason for these larger values of $m_t$ lies first in the larger average value of $r^t$, as compared with $V_t$, and second in the fact that for a given sum a portion of the $V_t y$ products are negative and a portion positive, whereas all $r^t y$ products in a given sum are positive.

We now use Eqs. (89) to evaluate $a_{kj}$ with the values of $A$, $B$, etc. from Davis' tables,[11] for $n=7$. It is just this process that causes the main limitation in the accuracy of the Davis method and hence it is here given in complete detail.

For $n=7$

$$A=0.5670995671,\quad B=-0.265151515\cdots,$$
$$C=0.02272727\cdots,\quad D=0.214330808\cdots,$$
$$E=-0.021148989\cdots,\quad A'=0.2625661376,$$
$$B'=-0.032407407\cdots,\quad C'=0.004629629629\cdots.$$

Hence, by Eq. (89),

$$a_{04}=+715.1409091-2350.329546+1655.038636=19.849999,$$
$$a_{24}=-334.3693182+1899.849715-1540.105398=25.374999,$$
$$a_{44}=+28.66022728-187.4667614+160.9065341=2.1000000,$$
$$a_{14}=+839.87555\cdots-813.83555\cdots=+26.0400\cdots,$$
$$a_{34}=-103.66222\cdots+116.26222\cdots=+12.6000\cdots.$$

All of the values of $a_{kt}$ thus calculated agree with the exact values of Eq. (81) to at least eight significant figures. But the number of multiplications and additions is now much greater, and the numerical size and complexity of the factors are very much larger. The results just presented have been obtained on a ten-key calculating machine and hence all products are given to ten digits. In any case, they cannot in general be trusted to more than ten digits because the factors $A$, $B$, etc. are given by Davis only to that number of digits. (But many of these factors, as shown in the figures quoted, involve simple repeating decimals and hence can be extended indefinitely.)

As Eq. (89) shows, the set of nine terms for $k$ even, or the set of four for $k$ odd, may be said to occupy a complete matrix, whereas in Eq. (80) we have only the main diagonal and one side of such a matrix. Thus the determination of $a_{04}$, $a_{24}$, and $a_{44}$ by Eq. (89) requires a total of nine multiplications and six additions, whereas the determination of these same three coefficients by Eq. (80) requires only six multiplications and three additions.[41]

The loss of accuracy in the Davis method results just from the fact that to obtain $a_{44}$, for instance, we add algebraically three terms that *almost cancel.* To be specific, the sum of the two positive terms in our illustrative problem differs from the negative term by only 1.1 percent of either. Thus if the individual terms are good to ten digits, the resulting value of $a_{44}$ is good only to eight digits. On the other hand, in the method illustrated in Section F the tabulated factors are exact, and if a ten-key machine is used the resulting values of $a_{kj}$ and other calculated quantities are in general good to ten digits.

Actually all methods for obtaining a least-squares' solution are in principle the same in that they all lead to the same numerical result for a given set of data. It is, in fact, merely the *order* in

[41] In the original Birge and Shea method,[3] where the $a_{kj}$ coefficients are calculated from Eq. (10) in which all $R_{tt}=1$, in place of Eq. (27) or Eq. (80), there are only *three* required multiplications and three additions. As has been shown in detail in Sections C and F, the adoption of the complete $b-V$ system of calculation in place of the Birge and Shea method does make the calculation of both the $a_{kj}$ and the $a_{kj}''$ coefficients slightly longer. On the other hand, it greatly shortens the evaluation of the calculated values of the function. The main reason, however, for the use of the $b-V$ system with its accompanying necessary replacement of the $R_{kt}$ table of Birge and Shea by the $S_{kt}$ Table XII of the present paper (see Table V), is the simplification of the numerical factors appearing in the various necessary calculations.

which the various necessary algebraic processes are performed that differs from one method to another. This point appears to me of sufficient importance to be considered with some care. Thus, just how does it happen that the Section F method gives the correct result to more places than does the Davis method, with the same calculating machine?

The answer is as follows. Let us consider the calculation of $b_4$, from which $a_{44}$ immediately follows by Eq. (23). We evaluate $\sum y \cdot V_t$, which involves both positive and negative terms. For $t=4$ we have 2724.12 for the sum of the positive terms and 2169.72 for the sum of the negative terms, with the difference $+554.40$, as the recorded value of $\sum y \cdot V_4$. This difference is 20 percent of the positive sum (compared to only one percent in the Davis method). But suppose that this process is carried out on a six-key machine. The values of $y$ are given only to five digits and hence all of the operations in model form 1 could be carried out on such a machine with the same complete accuracy in the determination of $b_t$ as that indicated in the model form.

The point is simply that on a six-key machine the lower dial (in which each *product* appears) has room for *twelve* digits, and all the *cancelling* of terms is done *on this lower dial*, in the continuous process of getting $\sum y \cdot V_4$. In other words, a *six*-key machine has here effectively the accuracy of a *twelve*-key machine in the crucial cancellations that limit the final accuracy of the result! But in the Davis method the cancellations occur with product terms of the form $a_i b_i$ in which the accuracy of both $a_i$ and $b_i$, and hence of their product, is limited just to the number of keys on the machine (unless we go through a laborious process of multiplication by parts with all auxiliary additions and subtractions done on paper). Thus for the Davis method a six-key machine has only six-key accuracy in handling the various cancelling terms whose algebraic sum may be good to a far *smaller* number of digits.

A more typical illustration of the difference in accuracy of the two methods is furnished by the sample problem used by Birge and Shea.[3] In that problem 25 observations (actually 25 observed spectral lines of a band series) are fitted to a fourth-degree polynomial. In that paper a six-key calculating machine was used and the final calculated values are good to six digits so far as the accuracy of the calculation itself is concerned. But the results for the same problem, when carried out by the Davis method with a *nine*-key machine, are definitely *less* accurate than those obtained by the Birge and Shea method with a six-key machine.

Thus the Davis method, in addition to being definitely longer and involving the handling of much larger factors, also requires that all intermediate results be obtained correctly to more digits for a given final accuracy, than does the method advocated in this paper. The Davis method suffers, in fact, from just the well-known defect of the standard solution of simultaneous linear equations by determinants. Thus, due to the almost complete cancellation of terms in the determination of the unknown first evaluated (which sets the accuracy of all the others), it is often necessary, in the case of four such simultaneous equations, to obtain correctly all products to ten digits in order to get final results correct to possibly five digits.

In conclusion it should again be noted that the Davis method gives the result in non-orthogonal form and thus lacks the numerous important advantages of the orthogonal solution advocated here. Furthermore, there has been published, so far as I am aware, no method for getting the probable errors of any of the results when derived by the Davis method. Because of the non-orthogonal character of the solution, the necessary method will almost certainly be found to be very laborious.

### G2. The Kerawala Method, A3β. (Model Form 8)

As indicated by its designation, the Kerawala method[10] yields a least-squares' solution in terms of a power series and explicitly in terms of the observations but in non-orthogonal form. It is, in fact, just the obvious simplification of the original Birge and Shea method that is possible if one relinquishes the advantages of the orthogonal solution.

In the Birge and Shea method the values of $a_t(=a_{tt})$ are first calculated from the observations in a model form analogous to the present model form 1, and then the $a_{kj}$ coefficients of $f(\epsilon)$ are calculated by Eq. (10). In the modified process presented in Sections C and F the $b_t$ values are calculated from the observations in model form 1 and the $a_{kj}$ coefficients are then calculated by Eq. (27). It is fairly obvious that one can combine the two steps of either method so that each coefficient $a_{kj}$ is expressed explicitly as a function of the observations. This is just what Kerawala has done.

In order to derive the various needed factors he begins with the expressions for the $a_{kj}$ coefficients in terms of power moments $m_t$, exactly as Davis[11] has done. Kerawala independently derives expressions for the factors $A$, $B$, etc., and as noted in footnote 40 his results involve explicitly the number of observations $n$, whereas this dependence is rather hidden by special symbols in Davis' expressions.

As a very simple example the first coefficient of the second-degree solution (compare Eq. (83)) is given by

$$a_{02}=Am_0+Bm_2, \quad (90)$$

where,[42] according to Kerawala,

$$A=\frac{3(3n^2-7)}{4n(n^2-4)}, \quad B=-\frac{15}{n(n^2-4)}. \quad (91)$$

Kerawala now "dissects" $m_0$ and $m_2$ into the terms whose sum these power moments represent in order to get finally an explicit dependence of $a_{02}$ on the observations $y_i$. Thus for $n=7$, $A=1/3$ and $B=-1/21$. But the moments, for $n=7$, may be written as

$$m_0=(y_3+y_{-3})+(y_2+y_{-2})+(y_1+y_{-1})+y_0,$$
$$m_2=3^2(y_3+y_{-3})+2^2(y_2+y_{-2})+1^2(y_1+y_{-1}). \quad (92)$$

Hence the actual factor multiplying $(y_3+y_{-3})$ is given by the combination of Eqs. (90), (91), and (92) as $1(1/3)-9(1/21)=-2/21$. Similarly for $(y_2+y_{-2})$ one finds $1(1/3)-4(1/21)=+3/21$. The remaining two factors, for $(y_1+y_{-1})$ and $y_0$, are $+6/21$ and $+7/21$. Since these four factors have a common denominator $K(=21)$, we can multiply by the respective numerators, add the products to get the sum $(\sum)$, and divide by $K$ to get $a_{02}$ in complete analogy to model form 1. In other words, *all* factors used in the Kerawala method are expressed as integers, just as they are in the determination of $b_t$ in model form 1, or of $a_t$ in the original Birge and Shea method.[3]

For our standard illustrative problem (Table VII) the calculation may then be put in the following model form 8.[43] Hence $u_4(\epsilon)=19.85+26.04\epsilon+25.375\epsilon^2+12.60\epsilon^3+2.1\epsilon^4$, in agreement with Eq. (81).

The Kerawala tables cover the same range as those of the Birge and Shea paper[3] and of the present paper, namely, to $n=30$ for $j=1$ to 5. His tables for calculating $a_{jj}$ are identical with those of Birge and Shea, since in both cases this final coefficient is given by Eq. (25). Kerawala states that all the numerical factors (values of $V_t$ and $K_t^*$) published by Birge and Shea were thus checked and found to be correct. But his own paper is not free from typographical errors. Thus the factor $+3$ at the head of the last column of model form 8 is printed by Kerawala as $-3$.

It is obvious that the Kerawala method is far superior in every respect to that of Davis. The Kerawala method is, in fact, the most rapid and accurate one for evaluating the coefficients $a_{kj}$ of the power series in $\epsilon$ if one wishes to forego the advantages of the orthogonal form of solution.

### G3. Factorial Forms of Solution

It has been noted in Section C that the least-squares' solution of a polynomial may be carried out in terms of factorial moments, and may be expressed in factorial form. Such a solution, which is strongly favored by Sasuly,[12] is denoted B2 in Section C. I have myself made no extensive investigation of the relative merits of the power series form of solution and the factorial form. As already stated, physical scientists usually prefer the power series form since it is often directly related to theory, whereas statisticians are usually interested primarily in merely smoothing the data. In the latter case the explicit analytic form of the solution is immaterial.

[42] Kerawala uses no explicit symbols for the coefficients of $m_t$. Those given here are due to Davis,[11] who writes the $A$, $B$, etc. without subscripts although their values depend obviously on the degree of the polynomial, as well as on the number of observations. Thus in Eq. (89) one writes, more precisely, $a_{04}=A_4m_0+B_4m_2+C_4m_4$ and in Eq. (90), $a_{02}=A_2m_0+B_2m_2$.

[43] Kerawala publishes all the model forms given by Birge and Shea (footnote 3). However he refers only to the 1924 abstract and not the complete 1927 paper, which contains the model forms. No reference is made by Kerawala to the work of Davis.

MODEL FORM 8. Kerawala method.

| $k=3$ | $k=1$ | diff. | $y_{-q}\cdots$ | $r$ | $y_q\cdots$ | sum | $k=0$ | $k=2$ | $k=4$ |
|---|---|---|---|---|---|---|---|---|---|
| +1 | −22 | 836.64 | 0 | 3 | 836.64 | 836.64 | +5 | −13 | +3 |
| −1 | +67 | 305.76 | 2.10 | 2 | 307.86 | 309.96 | −30 | +67 | −7 |
| −1 | +58 | 77.28 | 8.61 | 1 | 85.89 | 94.50 | +75 | −19 | +1 |
| 0 | 0 | — | — | 0 | 19.95 | 19.95 | +131 | −70 | +6 |
| 36 | 252 | | | $K$ | | | 231 | 264 | 264 |
| 453.60 | 6562.08 | | | $\Sigma$ | | | 4585.35 | 6699.0 | 554.4 |
| +12.60 | +26.04 | | | $a_{kj}$ | $(=\Sigma/K)$ | | +19.850 | +25.375 | +2.10 |

In this connection the remarks of two statisticians, Anderson and Houseman,[27] are of interest. As stated in Section C, where their paper is listed, they merely extend to $n=104$ the tables of $V_t$ values already given to $n=52$ by Fisher and Yates[24] (and much earlier to $n=30$ by Birge and Shea[3]). Before commencing this rather laborious task of computation, Anderson and Houseman made a careful study of the time required with factorial moments as contrasted with that required by their method, which is identical with our model form 1. It may be noted that the evaluation of the factorial moments themselves requires only successive additions,[44] which are most conveniently made on a printing-adding machine, whereas the method advocated here (and by Anderson and Houseman) involves the successive multiplications and additions of Eq. (23), which are best performed on an electric calculating machine.

Anderson and Houseman state that if both printing-adding and calculating machines are available the summation method requires almost 30 percent more time on the average. This time is almost doubled when no adding machine is available. They also state that the relative efficiency of the product method is greater for high degree polynomials, but that if no comparison is desired of the fit of polynomials of various degree, as shown by the relative magnitudes of $\sum v_j^2$, Eq. (30), there may be little difference in the computing time of the two methods. Finally they state that there are fewer formulas involved in the product method, and hence it is learned more quickly by the computer.

Although Sasuly, in private correspondence, has indicated a very strong preference for the summation method, he has apparently never made any such actual comparison of the two methods, and the experience of Anderson and Houseman seems to demonstrate conclusively the advantage of the method advocated in this paper over a method involving factorial moments.

Anderson and Houseman also refer to the beautiful work of Aitken.[45] They state that the method is too involved for ordinary computing work, but that if a computer is to handle polynomials exclusively the Aitken method is worthy of consideration. It may be remarked, in closing this section, that the chief purpose of the extensive investigations of both Aitken and Jordan (see Section H for more complete references) is to eliminate so far as possible the need for extensive tables of numerical factors. Thus in the latest paper by Jordan,[46] tables that occupy only 23 rather small pages give all of the necessary factors for the polynomial solution up to the seventh degree and up to 100 observations. In contrast, the Anderson and Houseman paper[27] includes 62 pages of tables with far more figures to the page, and carries the solution, as already noted, to 104 observations, but only to $j=5$. The

[44] The factorial moment corresponding to the power moment $\sum_{i=1}^{n} x^t y_i$ is given by

$$S_n^{(t)}=\sum_{i=1}^{n}\binom{n+t-i}{t}y_i.$$

Since all binomial coefficients can be derived by successive summation, the same process obviously yields factorial moments.

[45] A. C. Aitken, Proc. Roy. Soc. Edinburgh **53**, 54 (1932).
[46] Charles Jordan, Annals Math. Stat. **3**, 257 (1932).

brevity thus attained by Jordan is, however, annulled in part by the increased complexity of the method. In fact Jordan's article like that of Aitken is evidently addressed to trained mathematicians, and is not likely to be read profitably by any one without such training.

### H. SELECTED BIBLIOGRAPHY

(Reference in this paper to items in the Bibliography is indicated by the number of the footnote in brackets [ ].)


A. C. Aitken, Proc. Roy. Soc. Edinburgh **53**, 54 (1932–33). [45].

A. C. Aitken, Proc. Roy. Soc. Edinburgh **54**, 1 (1933–34).

A. C. Aitken, Proc. Roy. Soc. Edinburgh **55**, 42 (1935).

A. C. Aitken, *Statistical Mathematics* (Oliver and Boyd, Edinburgh, Scotland, 1939).

F. E. Allan, Proc. Roy. Soc. Edinburgh **50**, 310 (1929–30). [15].

R. L. Anderson and E. E. Houseman, Research Bulletin 297, Agric. Exp. Station, Iowa State College (1942). [27].

R. T. Birge, Phys. Rev. **13**, 360 (1919). [2].

R. T. Birge and J. D. Shea, Univ. of Calif. Pub. in Math. **2**, 67 (1927). [3].

R. T. Birge, Phys. Rev. **40**, 207 (1932).

R. T. Birge and J. W. Weinberg, Phys. Rev. **68**, 106A (1945). [4].

E. U. Condon, Univ. of Calif. Pub. in Math. **2**, 55 (1927). [38].

H. T. Davis, *Tables of the Higher Mathematical Functions* (Principia Press, Bloomington, Indiana, 1935), II, 307–385. [11].

H. T. Davis, Ann. Math. Stat. **4**, 154 (1933). [38].

H. T. Davis and V. V. Latshaw. Ann. Math. **31**, 52 (1930). [37].

F. Esscher, Försäkringsaktiebolaget Skandia, Stockholm, 107 (1930).

F. Esscher, Medelanden fran Lunds Astron. Obs. (1920).

R. A. Fisher, J. Agric. Sci. **11**, 107 (1921). [25].

R. A. Fisher, Trans. Roy. Soc. London **B213**, 89 (1924). [26].

R. A. Fisher and F. Yates, *Statistical Tables for Biological, Agricultural and Medical Research* (Oliver and Boyd, London, England, 1938), Table XXIII. [24].

R. A. Fisher, *Statistical Methods for Research Workers* (Oliver and Boyd, Edinburgh, Scotland, 1946), 10th edition, pp. 147–156. [13].

J. P. Gram, J. f. Math. **94**, 41 (1883). [38].

Charles Jordan (or Karl Jordan), Proc. London Math. Soc. **20**, 297 (1921). [20].

Charles Jordan, *Statistique Mathématique* (Gauthier-Villars, Paris, France, 1927), pp. 25–32, 270–275, 290–292, 321–334.

Charles Jordan, J. Soc. Hong. Stat., 35 (1929).

Charles Jordan, Hungarian Com. on Ec. Stat. (1930).

Charles Jordan, Ann. Math. Stat. **3**, 257 (1932). [46].

S. M. Kerawala, Indian J. Phys. **15**, 241 (1941). [10].

V. Khotimsky (also written Chotimsky), *Graduation of Statistical Series by Least Squares* (Moscow and Leningrad, 1925), in Russian. [23].

G. J. Lidstone, J. Inst. Actuaries (Great Britain) **64**, 128 (1933).

Max Sasuly, *Trend Analysis in Statistics* (the Brookings Institute, Washington, D. C., 1934). [12].

Henry Schultz, J. Am. Stat. Assoc. **25**, 139 (1930). [31].

H. Schwerdt, Phys. Zeit. **22**, 312 (1921).

J. A. Shohat, *A Bibliography of Orthogonal Polynomials* (Nat. Research Council, Bulletin 103, 1940).

P. L. Tchebycheff, *Oeuvres*, Vol. I, pp. 203–230, 381–384, 473–498, 541–560, 701–702; Vol. II, pp. 219–242. (1854 to 1875.) [9].

J. W. Weinberg and R. T. Birge, Phys. Rev. **68**, 106A (1945). [4].

J. Weinberg, Phys. Rev. **62**, 304A (1942). [4].

## I. TABLES XII, XIII, XIV

TABLE XII. Values of $S_{kt}$ for Eq. (27). ($S_{00}=1.$) $S_{kt}$ is the coefficient of $\epsilon^k$ in $V_t(\epsilon)$, the orthogonal polynomial with least integer ordinates—*see* Table III and Eq. (77).

| $n$ | $\lambda_1 = S_{11}$ | $S_{22}\cdot R_{02} = S_{02}$ | $\lambda_2 = S_{22}$ | $S_{33}\cdot R_{13} = S_{13}$ | $\lambda_3 = S_{33}$ | $S_{44}\cdot R_{04} = S_{04}$ | $S_{44}\cdot R_{24} = S_{24}$ | $\lambda_4 = S_{44}$ | $S_{55}\cdot R_{15} = S_{15}$ | $S_{55}\cdot R_{35} = S_{35}$ | $\lambda_5 = S_{55}$ | $n$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 1 | −2 | 3 | —— | —— | —— | —— | —— | —— | —— | —— | 3 |
| 4 | 2 | −1.25 | 1 | −6.83* | 3.3* | —— | —— | —— | —— | —— | —— | 4 |
| 5 | 1 | −2 | 1 | −2.83* | 0.83* | 6 | −12.916* | 2.916* | —— | —— | —— | 5 |
| 6 | 2 | −4.375 | 1.5 | −8.416* | 1.6* | 2.953125 | −3.9583* | 0.583* | 24.097916* | −16.916* | 2.1 | 6 |
| 7 | 1 | −4 | 1 | −1.16* | 0.16* | 6 | −5.583* | 0.583* | 8.73* | −4.083* | 0.35 | 7 |
| 8 | 2 | −5.25 | 1 | −6.16* | 0.6* | 10.828125 | −7.4583* | 0.583* | 16.3635416* | −11.083* | 0.7 | 8 |
| 9 | 1 | −20 | 3 | −9.83* | 0.83* | 18 | −9.583* | 0.583* | 11.93* | −3.083* | 0.15 | 9 |
| 10 | 2 | −4.125 | 0.5 | −24.416* | 1.6* | 20.109375 | −8.5416* | 0.416* | 12.639583* | −2.583* | 0.1 | 10 |
| 11 | 1 | −10 | 1 | −14.83* | 0.83* | 6 | −2.083* | 0.083* | 4.76* | −0.7916* | 0.025 | 11 |
| 12 | 2 | −35.75 | 3 | −14.16* | 0.6* | 30.1640625 | −8.72916* | 0.2916* | 41.4177083* | −5.7083* | 0.15 | 12 |
| 13 | 1 | −14 | 1 | −4.16* | 0.16* | 84 | −20.583* | 0.583* | 22.56* | −2.625 | 0.0583* | 13 |
| 14 | 2 | −8.125 | 0.5 | −48.416* | 1.6* | 113.953125 | −23.9583* | 0.583* | 123.047916* | −12.25 | 0.23* | 14 |
| 15 | 1 | −56 | 3 | −27.83* | 0.83* | 756 | −137.916* | 2.916* | 737.53* | −63.583* | 1.05 | 15 |
| 16 | 2 | −21.25 | 1 | −126.83* | 3.3* | 196.828125 | −31.4583* | 0.583* | 91.722916* | −6.916* | 0.1 | 16 |
| 17 | 1 | −24 | 1 | −7.16* | 0.16* | 36 | −5.083* | 0.083* | 58.86* | −3.916* | 0.05 | 17 |
| 18 | 2 | −40.375 | 1.5 | −16.083* | 0.3* | 45.421875 | −5.7083* | 0.083* | 446.585416* | −26.416* | 0.3 | 18 |
| 19 | 1 | −30 | 1 | −44.83* | 0.83* | 396 | −44.583* | 0.583* | 46.43* | −2.4583* | 0.025 | 19 |
| 20 | 2 | −33.25 | 1 | −198.83* | 3.3* | 1218.8203125 | −123.64583* | 1.4583* | 801.5302083* | −38.2083* | 0.35 | 20 |
| 21 | 1 | −110 | 3 | −54.83* | 0.83* | 594 | −54.583* | 0.583* | 1466.76* | −63.2916* | 0.525 | 21 |
| 22 | 2 | −20.125 | 0.5 | −24.083* | 0.3* | 716.953125 | −59.9583* | 0.583* | 787.714583* | −30.916* | 0.23* | 22 |
| 23 | 1 | −44 | 1 | −13.16* | 0.16* | 858 | −65.583* | 0.583* | 67.4 | −2.416* | 0.016* | 23 |
| 24 | 2 | −143.75 | 3 | −286.83* | 3.3* | 145.546875 | −10.2083* | 0.083* | 1441.835416* | −47.416* | 0.3 | 24 |
| 25 | 1 | −52 | 1 | −77.83* | 0.83* | 858 | −55.416* | 0.416* | 283.53* | −8.583* | 0.05 | 25 |
| 26 | 2 | −28.125 | 0.5 | −168.416* | 1.6* | 1406.953125 | −83.9583* | 0.583* | 664.639583* | −18.583* | 0.1 | 26 |
| 27 | 1 | −182 | 3 | −18.16* | 0.16* | 1638 | −90.583* | 0.583* | 4064.76* | −105.2916* | 0.525 | 27 |
| 28 | 2 | −65.25 | 1 | −78.16* | 0.6* | 948.1640625 | −48.72916* | 0.2916* | 3138.8635416* | −75.5416* | 0.35 | 28 |
| 29 | 1 | −70 | 1 | −104.83* | 0.83* | 2184 | −104.583* | 0.583* | 1808.36* | −40.5416* | 0.175 | 29 |
| 30 | 2 | −112.375 | 1.5 | −224.416* | 1.6* | 12515.765625 | −559.7916* | 2.916* | 3554.585416* | −74.416* | 0.3 | 30 |

0.83* indicates 0.83333 · · ·, etc.

TABLE XIII. Values of $V_t$ and $N_t$ for Eq. (23). (See model form 1.)

| $n$ | $r$ | $V_1$ | $V_2$ | $V_3$ | $V_4$ | $V_5$ |
|---|---|---|---|---|---|---|
| 3 | 1 | +1 | +1 | | | |
| | 0 | 0 | −2 | | | |
| | $N_t=$ | 2 | 6 | | | |
| 4 | 3/2 | +3 | +1 | +1 | | |
| | 1/2 | +1 | −1 | −3 | | |
| | $N_t=$ | 20 | 4 | 20 | | |
| 5 | 2 | +2 | +2 | +1 | +1 | |
| | 1 | +1 | −1 | −2 | −4 | |
| | 0 | 0 | −2 | 0 | +6 | |
| | $N_t=$ | 10 | 14 | 10 | 70 | |
| 6 | 5/2 | +5 | +5 | +5 | +1 | +1 |
| | 3/2 | +3 | −1 | −7 | −3 | −5 |
| | 1/2 | +1 | −4 | −4 | +2 | +10 |
| | $N_t=$ | 70 | 84 | 180 | 28 | 252 |
| 7 | 3 | +3 | +5 | +1 | +3 | +1 |
| | 2 | +2 | 0 | −1 | −7 | −4 |
| | 1 | +1 | −3 | −1 | +1 | +5 |
| | 0 | 0 | −4 | 0 | +6 | 0 |
| | $N_t=$ | 28 | 84 | 6 | 154 | 84 |

TABLE XIII.—*Continued.*

| $n$ | $r$ | $V_1$ | $V_2$ | $V_3$ | $V_4$ | $V_5$ |
|---|---|---|---|---|---|---|
| 8 | 7/2 | +7 | +7 | +7 | +7 | +7 |
| | 5/2 | +5 | +1 | −5 | −13 | −23 |
| | 3/2 | +3 | −3 | −7 | −3 | +17 |
| | 1/2 | +1 | −5 | −3 | +9 | +15 |
| | $N_t=$ | 168 | 168 | 264 | 616 | 2184 |
| 9 | 4 | +4 | +28 | +14 | +14 | +4 |
| | 3 | +3 | +7 | −7 | −21 | −11 |
| | 2 | +2 | −8 | −13 | −11 | +4 |
| | 1 | +1 | −17 | −9 | +9 | +9 |
| | 0 | 0 | −20 | 0 | +18 | 0 |
| | $N_t=$ | 60 | 2772 | 990 | 2002 | 468 |
| 10 | 9/2 | +9 | +6 | +42 | +18 | +6 |
| | 7/2 | +7 | +2 | −14 | −22 | −14 |
| | 5/2 | +5 | −1 | −35 | −17 | +1 |
| | 3/2 | +3 | −3 | −31 | +3 | +11 |
| | 1/2 | +1 | −4 | −12 | +18 | +6 |
| | $N_t=$ | 330 | 132 | 8580 | 2860 | 780 |
| 11 | 5 | +5 | +15 | +30 | +6 | +3 |
| | 4 | +4 | +6 | −6 | −6 | −6 |
| | 3 | +3 | −1 | −22 | −6 | −1 |

TABLE XIII.—*Continued.*

| $n$ | $r$ | $V_1$ | $V_2$ | $V_3$ | $V_4$ | $V_5$ |
|---|---|---|---|---|---|---|
| | 2 | +2 | −6 | −23 | −1 | +4 |
| | 1 | +1 | −9 | −14 | +4 | +4 |
| | 0 | 0 | −10 | 0 | +6 | 0 |
| | $N_t=$ | 110 | 858 | 4290 | 286 | 156 |
| 12 | 11/2 | +11 | +55 | +33 | +33 | +33 |
| | 9/2 | +9 | +25 | −3 | −27 | −57 |
| | 7/2 | +7 | +1 | −21 | −33 | −21 |
| | 5/2 | +5 | −17 | −25 | −13 | +29 |
| | 3/2 | +3 | −29 | −19 | +12 | +44 |
| | 1/2 | +1 | −35 | −7 | +28 | +20 |
| | $N_t=$ | 572 | 12,012 | 5148 | 8008 | 15,912 |
| 13 | 6 | +6 | +22 | +11 | +99 | +22 |
| | 5 | +5 | +11 | 0 | −66 | −33 |
| | 4 | +4 | +2 | −6 | −96 | −18 |
| | 3 | +3 | −5 | −8 | −54 | +11 |
| | 2 | +2 | −10 | −7 | +11 | +26 |
| | 1 | +1 | −13 | −4 | +64 | +20 |
| | 0 | 0 | −14 | 0 | +84 | 0 |
| | $N_t=$ | 182 | 2002 | 572 | 68,068 | 6188 |
| 14 | 13/2 | +13 | +13 | +143 | +143 | +143 |
| | 11/2 | +11 | +7 | +11 | −77 | −187 |
| | 9/2 | +9 | +2 | −66 | −132 | −132 |
| | 7/2 | +7 | −2 | −98 | −92 | +28 |
| | 5/2 | +5 | −5 | −95 | −13 | +139 |
| | 3/2 | +3 | −7 | −67 | +63 | +145 |
| | 1/2 | +1 | −8 | −24 | +108 | +60 |
| | $N_t=$ | 910 | 728 | 97,240 | 136,136 | 235,144 |
| 15 | 7 | +7 | +91 | +91 | +1001 | +1001 |
| | 6 | +6 | +52 | +13 | −429 | −1144 |
| | 5 | +5 | +19 | −35 | −869 | −979 |
| | 4 | +4 | −8 | −58 | −704 | −44 |
| | 3 | +3 | −29 | −61 | −249 | +751 |
| | 2 | +2 | −44 | −49 | +251 | +1001 |
| | 1 | +1 | −53 | −27 | +621 | +675 |
| | 0 | 0 | −56 | 0 | +756 | 0 |
| | $N_t=$ | 280 | 37,128 | 39,780 | 6,466,460 | 10,581,480 |
| 16 | 15/2 | +15 | +35 | +455 | +273 | +143 |
| | 13/2 | +13 | +21 | +91 | −91 | −143 |
| | 11/2 | +11 | +9 | −143 | −221 | −143 |
| | 9/2 | +9 | −1 | −267 | −201 | −33 |
| | 7/2 | +7 | −9 | −301 | −101 | +77 |
| | 5/2 | +5 | −15 | −265 | +23 | +131 |
| | 3/2 | +3 | −19 | −179 | +129 | +115 |
| | 1/2 | +1 | −21 | −63 | +189 | +45 |
| | $N_t=$ | 1360 | 5712 | 1,007,760 | 470,288 | 201,552 |
| 17 | 8 | +8 | +40 | +28 | +52 | +104 |
| | 7 | +7 | +25 | +7 | −13 | −91 |
| | 6 | +6 | +12 | −7 | −39 | −104 |
| | 5 | +5 | +1 | −15 | −39 | −39 |
| | 4 | +4 | −8 | −18 | −24 | +36 |
| | 3 | +3 | −15 | −17 | −3 | +83 |

TABLE XIII.—*Continued.*

| $n$ | $r$ | $V_1$ | $V_2$ | $V_3$ | $V_4$ | $V_5$ |
|---|---|---|---|---|---|---|
| | 2 | +2 | −20 | −13 | +17 | +88 |
| | 1 | +1 | −23 | −7 | +31 | +55 |
| | 0 | 0 | −24 | 0 | +36 | 0 |
| | $N_t=$ | 408 | 7752 | 3876 | 16,796 | 100,776 |
| 18 | 17/2 | +17 | +68 | +68 | +68 | +884 |
| | 15/2 | +15 | +44 | +20 | −12 | −676 |
| | 13/2 | +13 | +23 | −13 | −47 | −871 |
| | 11/2 | +11 | +5 | −33 | −51 | −429 |
| | 9/2 | +9 | −10 | −42 | −36 | +156 |
| | 7/2 | +7 | −22 | −42 | −12 | +588 |
| | 5/2 | +5 | −31 | −35 | +13 | +733 |
| | 3/2 | +3 | −37 | −23 | +33 | +583 |
| | 1/2 | +1 | −40 | −8 | +44 | +220 |
| | $N_t=$ | 1938 | 23,256 | 23,256 | 28,424 | 6,953,544 |
| 19 | 9 | +9 | +51 | +204 | +612 | +102 |
| | 8 | +8 | +34 | +68 | −68 | −68 |
| | 7 | +7 | +19 | −28 | −388 | −98 |
| | 6 | +6 | +6 | −89 | −453 | −58 |
| | 5 | +5 | −5 | −120 | −354 | +3 |
| | 4 | +4 | −14 | −126 | −168 | +54 |
| | 3 | +3 | −21 | −112 | +42 | +79 |
| | 2 | +2 | −26 | −83 | +227 | +74 |
| | 1 | +1 | −29 | −44 | +352 | +44 |
| | 0 | 0 | −30 | 0 | +396 | 0 |
| | $N_t=$ | 570 | 13,566 | 213,180 | 2,288,132 | 89,148 |
| 20 | 19/2 | +19 | +57 | +969 | +1938 | +1938 |
| | 17/2 | +17 | +39 | +357 | −102 | −1122 |
| | 15/2 | +15 | +23 | −85 | −1122 | −1802 |
| | 13/2 | +13 | +9 | −377 | −1402 | −1222 |
| | 11/2 | +11 | −3 | −539 | −1187 | −187 |
| | 9/2 | +9 | −13 | −591 | −687 | +771 |
| | 7/2 | +7 | −21 | −553 | −77 | +1351 |
| | 5/2 | +5 | −27 | −445 | +503 | +1441 |
| | 3/2 | +3 | −31 | −287 | +948 | +1076 |
| | 1/2 | +1 | −33 | −99 | +1188 | +396 |
| | $N_t=$ | 2660 | 17,556 | 4,903,140 | 22,881,320 | 31,201,800 |
| 21 | 10 | +10 | +190 | +285 | +969 | +3876 |
| | 9 | +9 | +133 | +114 | 0 | −1938 |
| | 8 | +8 | +82 | −12 | −510 | −3468 |
| | 7 | +7 | +37 | −98 | −680 | −2618 |
| | 6 | +6 | −2 | −149 | −615 | −788 |
| | 5 | +5 | −35 | −170 | −406 | +1063 |
| | 4 | +4 | −62 | −166 | −130 | +2354 |
| | 3 | +3 | −83 | −142 | +150 | +2819 |
| | 2 | +2 | −98 | −103 | +385 | +2444 |
| | 1 | +1 | −107 | −54 | +540 | +1404 |
| | 0 | 0 | −110 | 0 | +594 | 0 |
| | $N_t=$ | 770 | 201,894 | 432,630 | 5,720,330 | 121,687,020 |
| 22 | 21/2 | +21 | +35 | +133 | +1197 | +2261 |
| | 19/2 | +19 | +25 | +57 | +57 | −969 |
| | 17/2 | +17 | +16 | 0 | −570 | −1938 |

TABLE XIII.—*Continued.*

| $n$ | $r$ | $V_1$ | $V_2$ | $V_3$ | $V_4$ | $V_5$ |
|---|---|---|---|---|---|---|
| | 15/2 | +15 | +8 | −40 | −810 | −1598 |
| | 13/2 | +13 | +1 | −65 | −775 | −663 |
| | 11/2 | +11 | −5 | −77 | −563 | +363 |
| | 9/2 | +9 | −10 | −78 | −258 | +1158 |
| | 7/2 | +7 | −14 | −70 | +70 | +1554 |
| | 5/2 | +5 | −17 | −55 | +365 | +1509 |
| | 3/1 | +3 | −19 | −35 | +585 | +1079 |
| | 1/2 | +1 | −20 | −12 | +702 | +390 |
| | $N_t=$ | 3542 | 7084 | 96,140 | 8,748,740 | 40,562,340 |
| 23 | 11 | +11 | +77 | +77 | +1463 | +209 |
| | 10 | +10 | +56 | +35 | +133 | −76 |
| | 9 | +9 | +37 | +3 | −627 | −171 |
| | 8 | +8 | +20 | −20 | −950 | −152 |
| | 7 | +7 | +5 | −35 | −955 | −77 |
| | 6 | +6 | −8 | −43 | −747 | +12 |
| | 5 | +5 | −19 | −45 | −417 | +87 |
| | 4 | +4 | −28 | −42 | −42 | +132 |
| | 3 | +3 | −35 | −35 | +315 | +141 |
| | 2 | +2 | −40 | −25 | +605 | +116 |
| | 1 | +1 | −43 | −13 | +793 | +65 |
| | 0 | 0 | −44 | 0 | +858 | 0 |
| | $N_t=$ | 1012 | 35,420 | 32,890 | 13,123,110 | 340,860 |
| 24 | 23/2 | +23 | +253 | +1771 | +253 | +4807 |
| | 21/2 | +21 | +187 | +847 | +33 | −1463 |
| | 19/2 | +19 | +127 | +133 | −97 | −3743 |
| | 17/2 | +17 | +73 | −391 | −157 | −3553 |
| | 15/2 | +15 | +25 | −745 | −165 | −2071 |
| | 13/2 | +13 | −17 | −949 | −137 | −169 |
| | 11/2 | +11 | −53 | −1023 | −87 | +1551 |
| | 9/2 | +9 | −83 | −987 | −27 | +2721 |
| | 7/2 | +7 | −107 | −861 | +33 | +3171 |
| | 5/2 | +5 | −125 | −665 | +85 | +2893 |
| | 3/2 | +3 | −137 | −419 | +123 | +2005 |
| | 1/2 | +1 | −143 | −143 | +143 | +715 |
| | $N_t=$ | 4600 | 394,680 | 17,760,600 | 394,680 | 177,928,920 |
| 25 | 12 | +12 | +92 | +506 | +1518 | +1012 |
| | 11 | +11 | +69 | +253 | +253 | −253 |
| | 10 | +10 | +48 | +55 | −517 | −748 |
| | 9 | +9 | +29 | −93 | −897 | −753 |
| | 8 | +8 | +12 | −196 | −982 | −488 |
| | 7 | +7 | −3 | −259 | −857 | −119 |
| | 6 | +6 | −16 | −287 | −597 | +236 |
| | 5 | +5 | −27 | −285 | −267 | +501 |
| | 4 | +4 | −36 | −258 | +78 | +636 |
| | 3 | +3 | −43 | −211 | +393 | +631 |
| | 2 | +2 | −48 | −149 | +643 | +500 |
| | 1 | +1 | −51 | −77 | +803 | +275 |
| | 0 | 0 | −52 | 0 | +858 | 0 |
| | $N_t=$ | 1300 | 53,820 | 1,480,050 | 14,307,150 | 7,803,900 |
| 26 | 25/2 | +25 | +50 | +1150 | +2530 | +2530 |
| | 23/2 | +23 | +38 | +598 | +506 | −506 |
| | 21/2 | +21 | +27 | +161 | −759 | −1771 |

TABLE XIII.—*Continued.*

| $n$ | $r$ | $V_1$ | $V_2$ | $V_3$ | $V_4$ | $V_5$ |
|---|---|---|---|---|---|---|
| | 19/2 | +19 | +17 | −171 | −1419 | −1881 |
| | 17/2 | +17 | +8 | −408 | −1614 | −1326 |
| | 15/2 | +15 | 0 | −560 | −1470 | −482 |
| | 13/2 | +13 | −7 | −637 | −1099 | +377 |
| | 11/2 | +11 | −13 | −649 | −599 | +1067 |
| | 9/2 | +9 | −18 | −606 | −54 | +1482 |
| | 7/2 | +7 | −22 | −518 | +466 | +1582 |
| | 5/2 | +5 | −25 | −395 | +905 | +1381 |
| | 3/2 | +3 | −27 | −247 | +1221 | +935 |
| | 1/2 | +1 | −28 | −84 | +1386 | +330 |
| | $N_t=$ | 5850 | 16,380 | 7,803,900 | 40,060,020 | 48,384,180 |
| 27 | 13 | +13 | +325 | +130 | +2990 | +16445 |
| | 12 | +12 | +250 | +70 | +690 | −2530 |
| | 11 | +11 | +181 | +22 | −782 | −10879 |
| | 10 | +10 | +118 | −15 | −1587 | −12144 |
| | 9 | +9 | +61 | −42 | −1872 | −9174 |
| | 8 | +8 | +10 | −60 | −1770 | −4188 |
| | 7 | +7 | −35 | −70 | −1400 | +1162 |
| | 6 | +6 | −74 | −73 | −867 | +5728 |
| | 5 | +5 | −107 | −70 | −262 | +8803 |
| | 4 | +4 | −134 | −62 | +338 | +10058 |
| | 3 | +3 | −155 | −50 | +870 | +9479 |
| | 2 | +2 | −170 | −35 | +1285 | +7304 |
| | 1 | +1 | −179 | −18 | +1548 | +3960 |
| | 0 | 0 | −182 | 0 | +1638 | 0 |
| | $N_t=$ | 1638 | 712,530 | 101,790 | 56,448,210 | 2,032,135,560 |
| 28 | 27/2 | +27 | +117 | +585 | +1755 | +13455 |
| | 25/2 | +25 | +91 | +325 | +455 | −1495 |
| | 23/2 | +23 | +67 | +115 | −395 | −8395 |
| | 21/2 | +21 | +45 | −49 | −879 | −9821 |
| | 19/2 | +19 | +25 | −171 | −1074 | −7866 |
| | 17/2 | +17 | +7 | −255 | −1050 | −4182 |
| | 15/2 | +15 | −9 | −305 | −870 | −22 |
| | 13/2 | +13 | −23 | −325 | −590 | +3718 |
| | 11/2 | +11 | −35 | −319 | −259 | +6457 |
| | 9/2 | +9 | −45 | −291 | +81 | +7887 |
| | 7/2 | +7 | −53 | −245 | +395 | +7931 |
| | 5/2 | +5 | −59 | −185 | +655 | +6701 |
| | 3/2 | +3 | −63 | −115 | +840 | +4456 |
| | 1/2 | +1 | −65 | −39 | +936 | +1560 |
| | $N_t=$ | 7308 | 95,004 | 2,103,660 | 19,634,160 | 1,354,757,040 |
| 29 | 14 | +14 | +126 | +819 | +4095 | +8190 |
| | 13 | +13 | +99 | +468 | +1170 | −585 |
| | 12 | +12 | +74 | +182 | −780 | −4810 |
| | 11 | +11 | +51 | −44 | −1930 | −5885 |
| | 10 | +10 | +30 | −215 | −2441 | −4958 |
| | 9 | +9 | +11 | −336 | −2460 | −2946 |
| | 8 | +8 | −6 | −412 | −2120 | −556 |
| | 7 | +7 | −21 | −448 | −1540 | +1694 |
| | 6 | +6 | −34 | −449 | −825 | +3454 |
| | 5 | +5 | −45 | −420 | −66 | +4521 |
| | 4 | +4 | −54 | −366 | +660 | +4818 |

TABLE XIII.—*Continued.*

| $n$ | $r$ | $V_1$ | $V_2$ | $V_3$ | $V_4$ | $V_5$ |
|---|---|---|---|---|---|---|
| | 3 | +3 | −61 | −292 | +1290 | +4373 |
| | 2 | +2 | −66 | −203 | +1775 | +3298 |
| | 1 | +1 | −69 | −104 | +2080 | +1768 |
| | 0 | 0 | −70 | 0 | +2184 | 0 |
| | $N_t=$ | 2030 | 113,274 | 4,207,320 | 107,987,880 | 500,671,080 |
| 30 | 29/2 | +29 | +203 | +1827 | +23751 | +16965 |
| | 27/2 | +27 | +161 | +1071 | +7371 | −585 |
| | 25/2 | +25 | +122 | +450 | −3744 | −9360 |
| | 23/2 | +23 | +86 | −46 | −10504 | −11960 |
| | 21/2 | +21 | +53 | −427 | −13749 | −10535 |
| | 19/2 | +19 | +23 | −703 | −14249 | −6821 |
| | 17/2 | +17 | −4 | −884 | −12704 | −2176 |
| | 15/2 | +15 | −28 | −980 | −9744 | +2384 |
| | 13/2 | +13 | −49 | −1001 | −5929 | +6149 |
| | 11/2 | +11 | −67 | −957 | −1749 | +8679 |
| | 9/2 | +9 | −82 | −858 | +2376 | +9768 |
| | 7/2 | +7 | −94 | −714 | +6096 | +9408 |
| | 5/2 | +5 | −103 | −535 | +9131 | +7753 |
| | 3/2 | +3 | −109 | −331 | +11271 | +5083 |
| | 1/2 | +1 | −112 | −112 | +12376 | +1768 |
| | $N_t=$ | 8990 | 302,064 | 21,360,240 | 3,671,587,920 | 2,145,733,200 |

TABLE XIV. Values of $W_t$ and $L_t$ for Eq. (54). $W_t$ are the least integer coefficients used in expressing $t!a_t$ as a weighted sum of the $t$th order data-differences.

| $n$ | $r$ | $W_1$ | $W_2$ | $W_3$ | $W_4$ | $W_5$ |
|---|---|---|---|---|---|---|
| 2 | 0 | 1 | | | | |
| | $L_t=$ | 1 | | | | |
| 3 | 0 | | 1 | | | |
| | 1/2 | 1 | | | | |
| | $L_t=$ | 2 | 1 | | | |
| 4 | 0 | 4 | | 1 | | |
| | 1/2 | | 1 | | | |
| | 1 | 3 | | | | |
| | $L_t=$ | 10 | 2 | 1 | | |
| 5 | 0 | | 3 | | 1 | |
| | 1/2 | 3 | | 1 | | |
| | 1 | | 2 | | | |
| | 3/2 | 2 | | | | |
| | $L_t=$ | 10 | 7 | 2 | 1 | |
| 6 | 0 | 9 | | 8 | | 1 |
| | 1/2 | | 9 | | 1 | |
| | 1 | 8 | | 5 | | |
| | 3/2 | | 5 | | | |
| | 2 | 5 | | | | |
| | $L_t=$ | 35 | 28 | 18 | 2 | 1 |
| 7 | 0 | | 12 | | 5 | |
| | 1/2 | 6 | | 2 | | 1 |
| | 1 | | 10 | | 3 | |
| | 3/2 | 5 | | 1 | | |

TABLE XIV.—*Continued.*

| $n$ | $r$ | $W_1$ | $W_2$ | $W_3$ | $W_4$ | $W_5$ |
|---|---|---|---|---|---|---|
| | 2 | | 5 | | | |
| | 5/2 | 3 | | | | |
| | $L_t=$ | 28 | 42 | 6 | 11 | 2 |
| 8 | 0 | 16 | | 20 | | 12 |
| | 1/2 | | 20 | | 15 | |
| | 1 | 15 | | 16 | | 7 |
| | 3/2 | | 15 | | 7 | |
| | 2 | 12 | | 7 | | |
| | 5/2 | | 7 | | | |
| | 3 | 7 | | | | |
| | $L_t=$ | 84 | 84 | 66 | 44 | 26 |
| 9 | 0 | | 100 | | 45 | |
| | 1/2 | 10 | | 50 | | 9 |
| | 1 | | 90 | | 35 | |
| | 3/2 | 9 | | 35 | | 4 |
| | 2 | | 63 | | 14 | |
| | 5/2 | 7 | | 14 | | |
| | 3 | | 28 | | | |
| | 7/2 | 4 | | | | |
| | $L_t=$ | 60 | 462 | 198 | 143 | 26 |
| 10 | 0 | 25 | | 200 | | 21 |
| | 1/2 | | 25 | | 75 | |
| | 1 | 24 | | 175 | | 16 |
| | 3/2 | | 21 | | 50 | |
| | 2 | 21 | | 112 | | 6 |
| | 5/2 | | 14 | | 18 | |
| | 3 | 16 | | 42 | | |
| | 7/2 | | 6 | | | |
| | 4 | 9 | | | | |
| | $L_t=$ | 165 | 132 | 858 | 286 | 65 |
| 11 | 0 | | 75 | | 35 | |
| | 1/2 | 15 | | 175 | | 14 |
| | 1 | | 70 | | 30 | |
| | 3/2 | 14 | | 140 | | 9 |
| | 2 | | 56 | | 18 | |
| | 5/2 | 12 | | 84 | | 3 |
| | 3 | | 36 | | 6 | |
| | 7/2 | 9 | | 30 | | |
| | 4 | | 15 | | | |
| | 9/2 | 5 | | | | |
| | $L_t=$ | 110 | 429 | 858 | 143 | 52 |
| 12 | 0 | 36 | | 245 | | 224 |
| | 1/2 | | 315 | | 245 | |
| | 1 | 35 | | 224 | | 189 |
| | 3/2 | | 280 | | 189 | |
| | 2 | 32 | | 168 | | 108 |
| | 5/2 | | 216 | | 105 | |
| | 3 | 27 | | 96 | | 33 |
| | 7/2 | | 135 | | 33 | |
| | 4 | 20 | | 33 | | |
| | 9/2 | | 55 | | | |
| | 5 | 11 | | | | |
| | $L_t=$ | 286 | 2002 | 1287 | 1144 | 884 |
| 13 | 0 | | 147 | | 980 | |
| | 1/2 | 21 | | 98 | | 196 |
| | 1 | | 140 | | 882 | |
| | 3/2 | 20 | | 84 | | 147 |
| | 2 | | 120 | | 630 | |
| | 5/2 | 18 | | 60 | | 77 |
| | 3 | | 90 | | 330 | |
| | 7/2 | 15 | | 33 | | 22 |

TABLE XIV.—*Continued.*

| $n$ | $r$ | $W_1$ | $W_2$ | $W_3$ | $W_4$ | $W_5$ |
|---|---|---|---|---|---|---|
| | 4 | | 55 | | 99 | |
| | 9/2 | 11 | | 11 | | |
| | 5 | | 22 | | | |
| | 11/2 | 6 | | | | |
| | $L_t =$ | 182 | 1001 | 572 | 4862 | 884 |
| 14 | 0 | 49 | | 1568 | | 1764 |
| | 1/2 | | 98 | | 1764 | |
| | 1 | 48 | | 1470 | | 1568 |
| | 3/2 | | 90 | | 1470 | |
| | 2 | 45 | | 1200 | | 1078 |
| | 5/2 | | 75 | | 990 | |
| | 3 | 40 | | 825 | | 528 |
| | 7/2 | | 55 | | 495 | |
| | 4 | 33 | | 440 | | 143 |
| | 9/2 | | 33 | | 143 | |
| | 5 | 24 | | 143 | | |
| | 11/2 | | 13 | | | |
| | 6 | 13 | | | | |
| | $L_t =$ | 455 | 728 | 9724 | 9724 | 8398 |
| 15 | 0 | | 784 | | 15876 | |
| | 1/2 | 28 | | 1176 | | 15876 |
| | 1 | | 756 | | 14700 | |
| | 3/2 | 27 | | 1050 | | 12936 |
| | 2 | | 675 | | 11550 | |
| | 5/2 | 25 | | 825 | | 8316 |
| | 3 | | 550 | | 7425 | |
| | 7/2 | 22 | | 550 | | 3861 |
| | 4 | | 396 | | 3575 | |
| | 9/2 | 18 | | 286 | | 1001 |
| | 5 | | 234 | | 1400 | |
| | 11/2 | 13 | | 91 | | |
| | 6 | | 91 | | | |
| | 13/2 | 7 | | | | |
| | $L_t =$ | 280 | 6188 | 7956 | 92,378 | 83,980 |
| 16 | 0 | 64 | | 7056 | | 3024 |
| | 1/2 | | 336 | | 5292 | |
| | 1 | 63 | | 6720 | | 2772 |
| | 3/2 | | 315 | | 4620 | |
| | 2 | 60 | | 5775 | | 2112 |
| | 5/2 | | 275 | | 3465 | |
| | 3 | 55 | | 4400 | | 1287 |
| | 7/2 | | 220 | | 2145 | |
| | 4 | 48 | | 2860 | | 572 |
| | 9/2 | | 156 | | 1001 | |
| | 5 | 39 | | 1456 | | 143 |
| | 11/2 | | 91 | | 273 | |
| | 6 | 28 | | 455 | | |
| | 13/2 | | 35 | | | |
| | 7 | 15 | | | | |
| | $L_t =$ | 680 | 2856 | 50,388 | 33,592 | 16,796 |
| 17 | 0 | | 432 | | 1260 | |
| | 1/2 | 36 | | 504 | | 2772 |
| | 1 | | 420 | | 1188 | |
| | 3/2 | 35 | | 462 | | 2376 |
| | 2 | | 385 | | 990 | |
| | 5/2 | 33 | | 385 | | 1716 |
| | 3 | | 330 | | 715 | |
| | 7/2 | 30 | | 286 | | 1001 |
| | 4 | | 260 | | 429 | |
| | 9/2 | 26 | | 182 | | 429 |
| | 5 | | 182 | | 195 | |
| | 11/2 | 21 | | 91 | | 104 |
| | 6 | | 105 | | 52 | |
| | 13/2 | 15 | | 28 | | |
| | 7 | | 40 | | | |
| | 15/2 | 8 | | | | |
| | $L_t =$ | 408 | 3876 | 3876 | 8398 | 16,796 |
| 18 | 0 | 81 | | 1440 | | 30492 |
| | 1/2 | | 810 | | 1980 | |
| | 1 | 80 | | 1386 | | 28512 |
| | 3/2 | | 770 | | 1782 | |
| | 2 | 77 | | 1232 | | 23166 |
| | 5/2 | | 693 | | 1430 | |
| | 3 | 72 | | 1001 | | 16016 |
| | 7/2 | | 585 | | 1001 | |
| | 4 | 65 | | 728 | | 9009 |
| | 9/2 | | 455 | | 585 | |
| | 5 | 56 | | 455 | | 3744 |
| | 11/2 | | 315 | | 260 | |
| | 6 | 45 | | 224 | | 884 |
| | 13/2 | | 180 | | 68 | |
| | 7 | 32 | | 68 | | |
| | 15/2 | | 68 | | | |
| | 8 | 17 | | | | |
| | $L_t =$ | 969 | 7752 | 11,628 | 14,212 | 193,154 |
| 19 | 0 | | 675 | | 21780 | |
| | 1/2 | 45 | | 4950 | | 4356 |
| | 1 | | 660 | | 20790 | |
| | 3/2 | 44 | | 4620 | | 3861 |
| | 2 | | 616 | | 18018 | |
| | 5/2 | 42 | | 4004 | | 3003 |
| | 3 | | 546 | | 14014 | |
| | 7/2 | 39 | | 3185 | | 2002 |
| | 4 | | 455 | | 9555 | |
| | 9/2 | 35 | | 2275 | | 1092 |
| | 5 | | 350 | | 5460 | |
| | 11/2 | 30 | | 1400 | | 442 |
| | 6 | | 240 | | 2380 | |
| | 13/2 | 24 | | 680 | | 102 |
| | 7 | | 136 | | 612 | |
| | 15/2 | 17 | | 204 | | |
| | 8 | | 51 | | | |
| | 17/2 | 9 | | | | |
| | $L_t =$ | 570 | 6783 | 42,636 | 163,438 | 29,716 |
| 20 | 0 | 100 | | 27225 | | 104544 |
| | 1/2 | | 825 | | 81675 | |
| | 1 | 99 | | 26400 | | 99099 |
| | 3/2 | | 792 | | 75075 | |
| | 2 | 96 | | 24024 | | 84084 |
| | 5/2 | | 728 | | 63063 | |
| | 3 | 91 | | 20384 | | 63063 |
| | 7/2 | | 637 | | 47775 | |
| | 4 | 84 | | 15925 | | 40768 |
| | 9/2 | | 525 | | 31850 | |
| | 5 | 75 | | 11200 | | 21658 |
| | 11/2 | | 400 | | 17850 | |
| | 6 | 64 | | 6800 | | 8568 |
| | 13/2 | | 272 | | 7650 | |
| | 7 | 51 | | 3264 | | 1938 |
| | 15/2 | | 153 | | 1938 | |
| | 8 | 36 | | 969 | | |
| | 17/2 | | 57 | | | |
| | 9 | 19 | | | | |
| | $L_t =$ | 1330 | 8778 | 245,157 | 653,752 | 742,900 |
| 21 | 0 | | 3025 | | 49005 | |
| | 1/2 | 55 | | 9075 | | 254826 |
| | 1 | | 2970 | | 47190 | |
| | 3/2 | 54 | | 8580 | | 231231 |

TABLE XIV.—*Continued.*

| $n$ | $r$ | $W_1$ | $W_2$ | $W_3$ | $W_4$ | $W_5$ |
|---|---|---|---|---|---|---|
| | 2 | | 2808 | | 42042 | |
| | 5/2 | 52 | | 7644 | | 189189 |
| | 3 | | 2548 | | 34398 | |
| | 7/2 | 49 | | 6370 | | 137592 |
| | 4 | | 2205 | | 25480 | |
| | 9/2 | 45 | | 4900 | | 86632 |
| | 5 | | 1800 | | 16660 | |
| | 11/2 | 40 | | 3400 | | 44982 |
| | 6 | | 1360 | | 9180 | |
| | 13/2 | 34 | | 2040 | | 17442 |
| | 7 | | 918 | | 3876 | |
| | 15/2 | 27 | | 969 | | 3876 |
| | 8 | | 513 | | 969 | |
| | 17/2 | 19 | | 285 | | |
| | 9 | | 190 | | | |
| | 19/2 | 10 | | | | |
| | $L_t =$ | 770 | 33,649 | 86,526 | 408,595 | 1,931,540 |
| 22 | 0 | 121 | | 4840 | | 184041 |
| | 1/2 | | 605 | | 70785 | |
| | 1 | 120 | | 4719 | | 176176 |
| | 3/2 | | 585 | | 66066 | |
| | 2 | 117 | | 4368 | | 154154 |
| | 5/2 | | 546 | | 57330 | |
| | 3 | 112 | | 3822 | | 122304 |
| | 7/2 | | 490 | | 45864 | |
| | 4 | 105 | | 3136 | | 86632 |
| | 9/2 | | 420 | | 33320 | |
| | 5 | 96 | | 2380 | | 53312 |
| | 11/2 | | 340 | | 21420 | |
| | 6 | 85 | | 1632 | | 27132 |
| | 13/2 | | 255 | | 11628 | |
| | 7 | 72 | | 969 | | 10336 |
| | 15/2 | | 171 | | 4845 | |
| | 8 | 57 | | 456 | | 2261 |
| | 17/2 | | 95 | | 1197 | |
| | 9 | 40 | | 133 | | |
| | 19/2 | | 35 | | | |
| | 10 | 21 | | | | |
| | $L_t =$ | 1771 | 7084 | 48,070 | 624,910 | 1,448,655 |
| 23 | 0 | | 1452 | | 102245 | |
| | 1/2 | 66 | | 3146 | | 20449 |
| | 1 | | 1430 | | 99099 | |
| | 3/2 | 65 | | 3003 | | 18876 |
| | 2 | | 1365 | | 90090 | |
| | 5/2 | 63 | | 2730 | | 16016 |
| | 3 | | 1260 | | 76440 | |
| | 7/2 | 60 | | 2352 | | 12376 |
| | 4 | | 1120 | | 59976 | |
| | 9/2 | 56 | | 1904 | | 8568 |
| | 5 | | 952 | | 42840 | |
| | 11/2 | 51 | | 1428 | | 5168 |
| | 6 | | 765 | | 27132 | |
| | 13/2 | 45 | | 969 | | 2584 |
| | 7 | | 570 | | 14535 | |
| | 15/2 | 38 | | 570 | | 969 |
| | 8 | | 380 | | 5985 | |
| | 17/2 | 30 | | 266 | | 209 |
| | 9 | | 210 | | 1463 | |
| | 19/2 | 21 | | 77 | | |
| | 10 | | 77 | | | |
| | 21/2 | 11 | | | | |
| | $L_t =$ | 1012 | 17,710 | 32,890 | 937,365 | 170,430 |
| 24 | 0 | 144 | | 81796 | | 572572 |
| | 1/2 | | 5148 | | 20449 | |
| | 1 | 143 | | 80080 | | 552123 |
| | 3/2 | | 5005 | | 19305 | |

TABLE XIV.—*Continued.*

| $n$ | $r$ | $W_1$ | $W_2$ | $W_3$ | $W_4$ | $W_5$ |
|---|---|---|---|---|---|---|
| | 2 | 140 | | 75075 | | 494208 |
| | 5/2 | | 4725 | | 17160 | |
| | 3 | 135 | | 67200 | | 408408 |
| | 7/2 | | 4320 | | 14280 | |
| | 4 | 128 | | 57120 | | 308448 |
| | 9/2 | | 3808 | | 11016 | |
| | 5 | 119 | | 45696 | | 209304 |
| | 11/2 | | 3213 | | 7752 | |
| | 6 | 108 | | 33915 | | 124032 |
| | 13/2 | | 2565 | | 4845 | |
| | 7 | 95 | | 22800 | | 61047 |
| | 15/2 | | 1900 | | 2565 | |
| | 8 | 80 | | 13300 | | 22572 |
| | 17/2 | | 1260 | | 1045 | |
| | 9 | 63 | | 6160 | | 4807 |
| | 19/2 | | 693 | | 253 | |
| | 10 | 44 | | 1771 | | |
| | 21/2 | | 253 | | | |
| | 11 | 23 | | | | |
| | $L_t =$ | 2300 | 65,780 | 880,030 | 197,340 | 4,942,470 |
| 25 | 0 | | 2028 | | 143143 | |
| | 1/2 | 78 | | 26026 | | 143143 |
| | 1 | | 2002 | | 139425 | |
| | 3/2 | 77 | | 25025 | | 133848 |
| | 2 | | 1925 | | 128700 | |
| | 5/2 | 75 | | 23100 | | 116688 |
| | 3 | | 1800 | | 112200 | |
| | 7/2 | 72 | | 20400 | | 94248 |
| | 4 | | 1632 | | 91800 | |
| | 9/2 | 68 | | 17136 | | 69768 |
| | 5 | | 1428 | | 69768 | |
| | 11/2 | 63 | | 13566 | | 46512 |
| | 6 | | 1197 | | 48450 | |
| | 13/2 | 57 | | 9975 | | 27132 |
| | 7 | | 950 | | 29925 | |
| | 15/2 | 50 | | 6650 | | 13167 |
| | 8 | | 700 | | 15675 | |
| | 17/2 | 42 | | 3850 | | 4807 |
| | 9 | | 462 | | 6325 | |
| | 19/2 | 33 | | 1771 | | 1012 |
| | 10 | | 253 | | 1518 | |
| | 21/2 | 23 | | 506 | | |
| | 11 | | 92 | | | |
| | 23/2 | 12 | | | | |
| | $L_t =$ | 1300 | 26,910 | 296,010 | 1,430,715 | 1,300,650 |
| 26 | 0 | 169 | | 66248 | | 429429 |
| | 1/2 | | 1183 | | 273273 | |
| | 1 | 168 | | 65065 | | 416416 |
| | 3/2 | | 1155 | | 260260 | |
| | 2 | 165 | | 61600 | | 379236 |
| | 5/2 | | 1100 | | 235620 | |
| | 3 | 160 | | 56100 | | 323136 |
| | 7/2 | | 1020 | | 201960 | |
| | 4 | 153 | | 48960 | | 255816 |
| | 9/2 | | 918 | | 162792 | |
| | 5 | 144 | | 40698 | | 186048 |
| | 11/2 | | 798 | | 122094 | |
| | 6 | 133 | | 31920 | | 122094 |
| | 13/2 | | 665 | | 83790 | |
| | 7 | 120 | | 23275 | | 70224 |
| | 15/2 | | 525 | | 51205 | |
| | 8 | 105 | | 15400 | | 33649 |
| | 17/2 | | 385 | | 26565 | |
| | 9 | 88 | | 8855 | | 12144 |
| | 19/2 | | 253 | | 10626 | |
| | 10 | 69 | | 4048 | | 2530 |
| | 21/2 | | 138 | | 2530 | |

TABLE XIV.—*Continued.*

| $n$ | $r$ | $W_1$ | $W_2$ | $W_3$ | $W_4$ | $W_5$ |
|---|---|---|---|---|---|---|
| | 11 | 48 | | 1150 | | |
| | 23/2 | | 50 | | | |
| | 12 | 25 | | | | |
| | $L_t =$ | 2925 | 16,380 | 780,390 | 2,861,430 | 4,032,015 |
| 27 | 0 | | 8281 | | 372645 | |
| | 1/2 | 91 | | 8281 | | 3279276 |
| | 1 | | 8190 | | 364364 | |
| | 3/2 | 90 | | 8008 | | 3097094 |
| | 2 | | 7920 | | 340340 | |
| | 5/2 | 88 | | 7480 | | 2756754 |
| | 3 | | 7480 | | 302940 | |
| | 7/2 | 85 | | 6732 | | 2302344 |
| | 4 | | 6885 | | 255816 | |
| | 9/2 | 81 | | 5814 | | 1790712 |
| | 5 | | 6156 | | 203490 | |
| | 11/2 | 76 | | 4788 | | 1281987 |
| | 6 | | 5320 | | 150822 | |
| | 13/2 | 70 | | 3724 | | 829521 |
| | 7 | | 4410 | | 102410 | |
| | 15/2 | 63 | | 2695 | | 471086 |
| | 8 | | 3465 | | 61985 | |
| | 17/2 | 55 | | 1771 | | 223146 |
| | 9 | | 2530 | | 31878 | |
| | 19/2 | 46 | | 1012 | | 79695 |
| | 10 | | 1656 | | 12650 | |
| | 21/2 | 36 | | 460 | | 16445 |
| | 11 | | 900 | | 2990 | |
| | 23/2 | 25 | | 130 | | |
| | 12 | | 325 | | | |
| | 25/2 | 13 | | | | |
| | $L_t =$ | 1638 | 118,755 | 101,790 | 4,032,015 | 32,256,120 |
| 28 | 0 | 196 | | 41405 | | 3179904 |
| | 1/2 | | 3185 | | 248430 | |
| | 1 | 195 | | 40768 | | 3097094 |
| | 3/2 | | 3120 | | 238238 | |
| | 2 | 192 | | 38896 | | 2858856 |
| | 5/2 | | 2992 | | 218790 | |
| | 3 | 187 | | 35904 | | 2494206 |
| | 7/2 | | 2805 | | 191862 | |
| | 4 | 180 | | 31977 | | 2046528 |
| | 9/2 | | 2565 | | 159885 | |
| | 5 | 171 | | 27360 | | 1566873 |
| | 11/2 | | 2280 | | 125685 | |
| | 6 | 160 | | 22344 | | 1106028 |
| | 13/2 | | 1960 | | 92169 | |
| | 7 | 147 | | 17248 | | 706629 |
| | 15/2 | | 1617 | | 61985 | |
| | 8 | 132 | | 12397 | | 396704 |
| | 17/2 | | 1265 | | 37191 | |
| | 9 | 115 | | 8096 | | 185955 |
| | 19/2 | | 920 | | 18975 | |
| | 10 | 96 | | 4600 | | 65780 |
| | 21/2 | | 600 | | 7475 | |
| | 11 | 75 | | 2080 | | 13455 |
| | 23/2 | | 325 | | 1755 | |
| | 12 | 52 | | 585 | | |
| | 25/2 | | 117 | | | |
| | 13 | 27 | | | | |
| | $L_t =$ | 3654 | 47,502 | 525,915 | 2,804,880 | 32,256,120 |

TABLE XIV.—*Continued.*

| $n$ | $r$ | $W_1$ | $W_2$ | $W_3$ | $W_4$ | $W_5$ |
|---|---|---|---|---|---|---|
| 29 | 0 | | 3675 | | 662480 | |
| | 1/2 | 105 | | 63700 | | 2252432 |
| | 1 | | 3640 | | 649740 | |
| | 3/2 | 104 | | 61880 | | 2144142 |
| | 2 | | 3536 | | 612612 | |
| | 5/2 | 102 | | 58344 | | 1939938 |
| | 3 | | 3366 | | 554268 | |
| | 7/2 | 99 | | 53295 | | 1662804 |
| | 4 | | 3135 | | 479655 | |
| | 9/2 | 95 | | 47025 | | 1343034 |
| | 5 | | 2850 | | 395010 | |
| | 11/2 | 90 | | 39900 | | 1013859 |
| | 6 | | 2520 | | 307230 | |
| | 13/2 | 84 | | 32340 | | 706629 |
| | 7 | | 2156 | | 223146 | |
| | 15/2 | 77 | | 24794 | | 446292 |
| | 8 | | 1771 | | 148764 | |
| | 17/2 | 69 | | 17710 | | 247940 |
| | 9 | | 1380 | | 88550 | |
| | 19/2 | 60 | | 11500 | | 115115 |
| | 10 | | 1000 | | 44850 | |
| | 21/2 | 50 | | 6500 | | 40365 |
| | 11 | | 650 | | 17550 | |
| | 23/2 | 39 | | 2925 | | 8190 |
| | 12 | | 351 | | 4095 | |
| | 25/2 | 27 | | 819 | | |
| | 13 | | 126 | | | |
| | 27/2 | 14 | | | | |
| | $L_t =$ | 2030 | 56,637 | 841,464 | 7,713,420 | 23,841,480 |
| 30 | 0 | 225 | | 156800 | | 5470192 |
| | 1/2 | | 6300 | | 4331600 | |
| | 1 | 224 | | 154700 | | 5346432 |
| | 3/2 | | 6188 | | 4176900 | |
| | 2 | 221 | | 148512 | | 4988412 |
| | 5/2 | | 5967 | | 3879876 | |
| | 3 | 216 | | 138567 | | 4434144 |
| | 7/2 | | 5643 | | 3464175 | |
| | 4 | 209 | | 125400 | | 3741309 |
| | 9/2 | | 5225 | | 2962575 | |
| | 5 | 200 | | 109725 | | 2979504 |
| | 11/2 | | 4725 | | 2413950 | |
| | 6 | 189 | | 92400 | | 2220834 |
| | 13/2 | | 4158 | | 1859550 | |
| | 7 | 176 | | 74382 | | 1530144 |
| | 15/2 | | 3542 | | 1338876 | |
| | 8 | 161 | | 56672 | | 956340 |
| | 17/2 | | 2898 | | 885500 | |
| | 9 | 144 | | 40250 | | 526240 |
| | 19/2 | | 2250 | | 523250 | |
| | 10 | 125 | | 26000 | | 242190 |
| | 21/2 | | 1625 | | 263250 | |
| | 11 | 104 | | 14625 | | 84240 |
| | 23/2 | | 1053 | | 102375 | |
| | 12 | 81 | | 6552 | | 16965 |
| | 25/2 | | 567 | | 23751 | |
| | 13 | 56 | | 1827 | | |
| | 27/2 | | 203 | | | |
| | 14 | 29 | | | | |
| | $L_t =$ | 4495 | 100,688 | 2,136,024 | 52,451,256 | 59,603,700 |

# J. Mathematical Appendix

J. W. WEINBERG

## J1. Introduction

THE task of fitting a polynomial to a series of data by the method of least squares is essentially the solution of a set of simultaneous, linear, "normal" equations. The problem is so much further simplified in the case of equally weighted and uniformly spaced data, that explicit analytical formulas for most quantities of interest can readily be found by elementary algebra. Ninety years ago P. L. Tchebycheff[9] thoroughly expounded the general problem, and explicit results for the simplified case were obtained twenty-five years ago by K. Jordan.[20] Since that time, many more or less independent presentations have appeared in diverse publications (see footnotes 3, 12, 15, 19, 23, 25, 45, 46, 47), often accompanied by extensive tables and directions for use. At this late date, it would seem impossible to contribute to the literature of this subject anything fundamentally new or important.

And yet, here is another independent exposition. The reason for its existence rests on three differences it bears from other work in this field.

Past treatments have been content to obtain the coefficients and adjusted values of least-squares' polynomials and to set aside the problem of their probable errors. But the advance in accurate measurement has given new importance to questions of probable error, and it was found necessary to extend the theory to these questions. Because of the relative inaccessibility of the general theory—a consequence both of the scattered journals in which it has been reported and the specialized learning lavished upon it—it seems difficult to explain the foundation of these new developments without a thorough elementary treatment of the whole subject.

Wherever analysis and numerical application have been developed in a unified way,[15, 20, 45, 46] they have been directed toward use in connection with adding-tabulating machines. The convenience and prevalence of modern automatic calculators have brought to the fore, however, the method of Birge and Shea.[3] It is to the application of their rapid and accurate technique, which makes use to the fullest extent of the general results of theory, that the following analysis is oriented.

In the past, expositions have derived their cogency from learned and powerful mathematical methods—Tchebycheff's from orthogonal polynomials as the successive convergents of a continued fraction, Jordan's from elaborate analysis in the calculus of finite differences. It seems unlikely that many workers with least squares have the opportunity to follow such arguments. The result has been that many conclusions of Tchebycheff and Jordan have been rediscovered accidentally or empirically, and incorporated in various methods and tables without recognition of their origin or full significance.[3, 12, 25] An even more striking result has been the publication of methods making no use whatever of the great practical simplifications that theory can bring. There appears to be a real need for a presentation of the theory, appropriate to the obviously elementary character of the problem—a presentation on the basis of simple algebra, such as the one that follows.

Although this section is designed to provide a mathematical basis for the methods set forth in preceding sections, and a fixed point of principle by which the preceding critique has been oriented, it is convenient to present it as an analytical unit with its own equation numbers in square brackets. Where this procedure involves repetition of an equation quoted in an earlier section, its number there will be repeated in parentheses.

## J2. Tchebycheff Polynomials

Tchebycheff's significant contribution to this field was the introduction of orthogonal polynomials, which render the least-squares' problem of degree $j$ essentially independent of that for

[47] E. U. Condon, Univ. of Calif. Pub. in Math. 2, 55 (1927). This article is mentioned also in footnote 38, in connection with the work of H. T. Davis.

degree $(j-1)$. The orthogonal polynomials $T_t(x)$ of degree $t$ in $x$ are defined over a range of $n$ discrete values of $x$ at which the data $y(x)$ appear. The property of orthogonality, in this case, requires that

$$\sum_x T_k(x)T_t(x)=0, \quad (t\neq k), \qquad [1], (4)$$

and if $T_t(x)$ is further limited by requiring its leading coefficient to be unity, i.e., its leading term to be just $x^t$, a definite numerical value as a function of $n$ and $t$ is further assigned to

$$M_t=\sum_x[T_t(x)]^2. \qquad [2], (4)$$

*Any* polynomial of degree $j$ can, of course, be expressed as a linear combination of Tchebycheff polynomials $T_t(x)$ with $t\leqslant j$, in accordance with

$$u_j(x)=\sum_{t=0}^{j} a_tT_t(x), \qquad [3], (5)$$

the coefficients $a_t$ being obtained explicitly by multiplying both sides with $T_t(x)$, summing over $x$, and applying Eqs. [1] and [2]:

$$\sum_x u_j(x)T_t(x)=a_tM_t. \qquad [4]$$

In addition to this fundamental property of orthogonal polynomials—their simplification of the determination of the $a_t$ coefficients from the function $u_j(x)$—there is a special simplicity that they bring to least-squares' problems. Suppose, for example, that $y(x)$ is a set of $n$ equally weighted data, and that $u_j(x)$ is the least-squares' polynomial of degree $j$ expanded in the manner of Eq. [3]. Then $\sum_x[y(x)-u_j(x)]^2$ is the sum of squared residuals which must be minimized by suitable adjustment of the values of the several $a_t$. Upon differentiation with respect to $a_t$, therefore, and observation that $T_t(x)=\partial u_j/\partial a_t$, one obtains, by means of Eq. [4], the minimizing condition

$$\sum_x y(x)T_t(x)=\sum_x u_j(x)T_t(x)=M_ta_t.$$

This is a direct expression in terms of the data for the value of $a_t$ that makes $u_j(x)$ the least-squares' solution of degree $j$:

$$a_t=\sum_x y(x)[T_t(x)/M_t]. \qquad [5], (7)$$

Because this result is independent of the degree $j$, of $u_j(x)$, the suppression of any $j$ dependence in the symbol $a_t$ is justified. (The possible $n$ dependence of any quantity is generally not explicitly indicated.) On proceeding from $u_{j-1}(x)$ to the more detailed representation of the data afforded by $u_j(x)$, one has merely to compute a single new quantity $a_j$ from Eq. [5] in order to obtain explicitly the value

$$u_j(x)=u_{j-1}(x)+a_jT_j(x). \qquad [6]$$

It is this circumstance that makes $a_jT_j(x)$ itself the least-squares' fit of degree $j$ to the residuals $y(x)-u_{j-1}(x)$, as discussed in connection with Eq. (76).

The minimum value of the sum of squared residuals is needed for the calculation of the probable error $r_j'$ of an individual datum, Eq. (34); and it is readily found with the aid of Eqs. [4] and [5]:

$$\sum_x y(x)u_j(x)=\sum_x\sum_{t=0}^{j} y(x)a_tT_t(x)=\sum_{t=0}^{j} M_ta_t^2,$$

$$\sum_x[u_j(x)]^2=\sum_x\sum_{t=0}^{j} u_j(x)a_tT_t(x)=\sum_{t=0}^{j} M_ta_t^2.$$

As a consequence,

$$\sum_x[y(x)-u_j(x)]^2$$
$$=\sum_x[y(x)]^2-2\sum_x y(x)u_j(x)+\sum_x[u_j(x)]^2$$
$$=\sum_x[y(x)]^2-\sum_{t=0}^{j} M_ta_t^2, \qquad [7], (29)$$

and the sum of squared residuals is thus diminished by precisely the amount $M_ja_j^2$ on proceeding from $u_{j-1}(x)$ to the fit of next higher degree, $u_j(x)$. This suggests that $M_t$ is the statistical weight of $a_t$, and that the probable error of $a_t$ is therefore $r'/(M_t)^{\frac{1}{2}}$, a fact that can be demonstrated by reference to Eq. [5]. Since in that equation the coefficients of $y(x)$ are $[T_t(x)/M_t]$, the reciprocal of the weight of $a_t$ relative to the data must be given by the sum of the squares of those coefficients, i.e., $\sum_x[T_t(x)/M_t]^2=1/M_t$, according to Eq. [2].

It is quite important that the coefficients $a_t$ of the Tchebycheff polynomials are, in effect, statistically independent combinations of the data, so that they act like *independently observed quantities* of appropriate weight $M_t$. This fact is seen by studying a linear combination of the $a_t$ with arbitrary coefficients $q_t$,

$$Q=\sum_{t=0}^{j} q_t a_t=\sum_{t=0}^{j} q_t \sum_x y(x)[T_t(x)/M_t]$$
$$=\sum_x y(x)\left[\sum_{t=0}^{j} q_t T_t(x)/M_t\right],$$

upon application of Eq. [5]. The weight of $Q$, denoted $p_Q$, may be found by summing the squares of the coefficients of the data $y(x)$ to form

$$1/p_Q=\sum_x\left[\sum_{t=0}^{j} q_t T_t(x)/M_t\right]^2$$
$$=\sum_x \sum_{t=0}^{j} \sum_{k=0}^{j}[q_k q_t/M_k M_t]T_t(x)T_k(x),$$

which, by means of Eqs. [1] and [2], may be expressed as

$$1/p_Q=\sum_{t=0}^{j} q_t^2/M_t. \qquad [8]$$

This is precisely the relation that would result from applying the law of propagation of errors directly to the coefficients $a_t$ considered as *independently observed* quantities of weight $M_t$.

An immediate consequence of this new theorem may be used in the problem of the probable error of the adjusted or smoothed data, the values of $u_j(x)$. If $p_j(x)$ denotes the weight of $u_j(x)$ relative to the weight of the data, then by Eqs. [3] and [8]

$$1/p_j(x)=\sum_{t=0}^{j}[T_t(x)]^2/M_t. \qquad [9], (35)$$

Further application of the theorem [8] may be made in finding the weights of the coefficients of $u_j(x)$ expressed as a sum of polynomials other than $T_t(x)$, i.e., the monomials $x^t$, as in Eq. [16] ahead.

### J3. Symmetry Properties

Up to this point there has been no need to limit arguments to the case of equally spaced data, but to make further progress by elementary methods, it is necessary to assume that situation. By introducing a new variable, $\epsilon$, the data may be placed at points symmetric about the origin $\epsilon=0$ and ranging by integer steps from $-(n-1)/2$ to $+(n-1)/2$. In this notation $a_j$ is the coefficient of $\epsilon^j$ in $u_j(\epsilon)$, and $y(-\epsilon)$ is the same set of data as $y(\epsilon)$ but enumerated in the opposite order as $\epsilon$ runs through its values. The function $u_j(-\epsilon)$ must certainly be the least-squares' fit to the data $y(-\epsilon)$, ordered in reverse, with $(-)^j a_j$ as the coefficient of $\epsilon^j$. Application of Eq. [5] to this case yields

$$(-)^j a_j=\sum_\epsilon y(-\epsilon)[T_j(\epsilon)/M_j],$$

which, on reversing the order of summation becomes

$$(-)^j a_j=\sum_\epsilon y(\epsilon)[T_j(-\epsilon)/M_j].$$

Comparing this result with the application of Eq. [5] to the data ordered as usual, namely,

$$a_j=\sum_\epsilon y(\epsilon)[T_j(\epsilon)/M_j],$$

one must conclude that

$$T_j(-\epsilon)=(-)^j T_j(\epsilon), \qquad [10]$$

because $y(\epsilon)$ may be chosen arbitrarily. This last result means that $T_j(\epsilon)$ contains only odd or only even powers of $\epsilon$ as $j$ is odd or even, respectively, and the number of parameters determining $T_j$ as a function of $\epsilon$ is thereby reduced by half. The summation in Eq. [5] that expresses $a_t$ in terms of the data, is correspondingly shortened:

$$M_t a_t=\sum_{r>0}[y(r)+(-)^t y(-r)]T_t(r)$$
$$+y(0)T_t(0), \qquad [11]$$

with $r$ defined as $|\epsilon|$, and the term in $y(0)$ to be included only when the number of data is odd and $\epsilon$ is capable of assuming the value zero. Equations [10] and [11] express the "pair-factor" type of symmetry, to the systematic exploitation of which is due some of the special convenience of the method of Birge and Shea.[3]

These results may be embodied in the properties of the coefficients $R_{kt}$ of $\epsilon^k$ in the explicit form of $T_t$:

$$T_t(\epsilon)=\sum_k R_{kt}\epsilon^k. \qquad [12]$$

(Cf. Table III, Section C2.)

The fact that $T_t(\epsilon)$ is a polynomial of degree $t$ in $\epsilon$ with leading coefficient unity can be expressed as

$$R_{kt}=0 \text{ for } k<0 \text{ or } k>t, \quad R_{tt}=1, \qquad [13]$$

and the pair-factor symmetry appears as

$$R_{kt}=0 \text{ for } (t-k) \text{ odd}. \qquad [14]$$

One may use the values of $R_{kt}$ to determine directly the coefficients $a_{kj}$ of $\epsilon^k$ in $u_j(\epsilon)$ by means of Eqs. [3] and [12]:

$$a_{kj}=\sum_{t=k}^{j} R_{kt}a_t. \qquad [15], (10)$$

According to this relation one needs to use the data directly only in finding the quantities $a_t$, while the coefficients $a_{kj}$ are specific linear combinations of these $a_t$. A special use of this is, by Eq. [15],

$$a_{jj}=R_{jj}a_j=a_j.$$

On proceeding from $u_{j-1}(\epsilon)$ to $u_j(\epsilon)$, $a_{kj}$ may be found from $a_{k,j-1}$ by addition of $R_{kj}a_j$. Furthermore, the theorem concerning the effective statistical independence of $a_t$ as expressed by Eq. [8], may be employed to obtain the weight $p_{kj}$ of $a_{kj}$:

$$1/p_{kj}=\sum_{t=k}^{j}[R_{kt}]^2/M_t, \qquad [16], (37)$$

which reduces, by Eq. [13], to the required $1/p_{jj}=1/M_j$.

### J4. Recursive Relations

To realize numerically the foregoing results one must evaluate the quantities $R_{kt}$, $M_t$, and eventually $T_t(x)$ as functions of $n$ and of their explicit arguments. Although analytical formulas for $T_t(x)$ and $M_t$ will be developed later, similar results for $R_{kt}$ are not possible because of the intervention of the Stirling numbers, which connect $(x+k)(x+k-1)\cdots(x+1)$ and $x^k$, $x^{k-1}$, etc., and for which there is no general formula. The quantities $R_{kt}$ are nevertheless connected by a recursive relation which enables them to be found when $R_{k-1,t-1}$, $R_{k,t-2}$, $M_{t-1}$, and $M_{t-2}$ are known. As a matter of fact, tables of $R_{kt}$ values are more readily constructed by the recursive algorithm than by any (necessarily quite complicated) analytical formula. In similar fashion, one may also tabulate the numbers $T_t(\epsilon)$ and $M_t$ by simple recursive methods, but the versatility and power of the theory is severely limited without the analysis which leads to closed formulas for $M_t$ and $T_t(x)$.

The recursive relations in question stem from a formula of Tchebycheff, readily derived by expanding $\epsilon T_t(\epsilon)$ in orthogonal polynomials. Now $\epsilon T_t(\epsilon)$ is a polynomial of degree $(t+1)$, containing only odd powers of $\epsilon$ when $t$ is even, and only even powers when $t$ is odd. It may therefore be expressed as a sum of orthogonal polynomials of degree $(t+1)$, $(t-1)$, $(t-3)$, etc., and since $\epsilon T_t(\epsilon)$ and $T_{t+1}(\epsilon)$ both have the leading term $\epsilon^{t+1}$, it is evident that the first term in the expansion must have unit coefficient:

$$\epsilon T_t(\epsilon)=T_{t+1}(\epsilon)+c_tT_{t-1}(\epsilon)+\cdots,$$

the remaining terms being of degree $\leqslant(t-3)$. One may next evaluate $\sum_\epsilon T_k(\epsilon)\cdot\epsilon\cdot T_t(\epsilon)$ which certainly vanishes, by Eq. [1], when $k>(t+1)$ because the expansion of $\epsilon T_t(\epsilon)$ contains no Tchebycheff polynomials of degree higher than $(t+1)$. Since the expression is, furthermore, quite symmetric in $k$ and $t$, one may apply the same argument, interchanging $k$ and $t$, with the conclusion that the expression vanishes whenever $t>(k+1)$. It has thus a finite value only for $k=t\pm1$, and, therefore, according to Eq. [4], only the first two terms in the expansion of $\epsilon T_t(\epsilon)$ have coefficients not identically zero, i.e.,

$$\epsilon T_t(\epsilon)=T_{t+1}+c_tT_{t-1}(\epsilon). \qquad [17]$$

To obtain the value of $c_t$ in this simple recursion formula, one must consider $\sum_\epsilon T_{t-1}(\epsilon)\cdot\epsilon\cdot T_t(\epsilon)$. This expression can be evaluated by multiplying Eq. [17] by $T_{t-1}(\epsilon)$ and summing over all $\epsilon$, in which case the result is $c_tM_{t-1}$ according to Eq. [2]. It can be found from the form which Eq. [17] takes when $(t-1)$ replaces $t$, by multiplying both sides by $T_t(\epsilon)$ and summing over $\epsilon$, in which case the result is evidently $M_t$. Comparing these two results for $\sum_\epsilon T_{t-1}(\epsilon)\cdot\epsilon\cdot T_t(\epsilon)$, one obtains

$$c_t=M_t/M_{t-1}. \qquad [18]$$

The combination of Eqs. [17] and [18] is known as Tchebycheff's recursion formula for obtaining $T_{t+1}(\epsilon)$ in terms of $T_t(\epsilon)$ and $T_{t-1}(\epsilon)$:

$$T_{t+1}(\epsilon)=\epsilon T_t(\epsilon)-(M_t/M_{t-1})T_{t-1}(\epsilon). \qquad [19]$$

For example, one may use the values of $T_0$ and $T_1$ obtained from the restriction that they have unit leading coefficient to calculate the first few Tchebycheff polynomials from Eq. [17]:

$T_0(\epsilon)=1,\quad T_1(\epsilon)=\epsilon,\quad T_2(\epsilon)=\epsilon^2-c_1,$
$T_3(\epsilon)=\epsilon^3-(c_1+c_2)\epsilon,$

$$T_4(\epsilon)=\epsilon^4-(c_1+c_2+c_3)\epsilon^2+c_1c_3,$$
$$T_5(\epsilon)=\epsilon^5-(c_1+c_2+c_3+c_4)\epsilon^3+[c_1c_3+(c_1+c_2)c_4]\epsilon. \quad [20]$$

Explicit values of $M_t$, and thence of $c_t$, may be found in succession by means of Eqs. [2] and [18]; e.g.,

$$M_0=n,\quad M_1=n(n^2-1)/12,$$
$$M_2=n(n^2-1)(n^2-4)/180,$$
$$M_3=\frac{n(n^2-1)(n^2-4)(n^2-9)}{2800}, \quad [21]$$
$$c_1=(n^2-1)/12,\quad c_2=(n^2-4)/15,$$
$$c_3=9(n^2-9)/140,\quad \text{etc.}$$

It will shortly be proved (see Eq. [44]) that, as might be guessed from the values in Eq. [21], $c_t=[t^2(n^2-t^2)]/[4(4t^2-1)]$; this formula alone is clearly sufficient to place the entire theory on a quantitative basis.

Substituting Eq. [12] into Eq. [19] and equating coefficients of $\epsilon^{k+1}$ yields at once the recursion on which the Birge and Shea $R_{kt}$ tables may be built:

$$R_{k+1,\,t+1}=R_{kt}-c_tR_{k+1,\,t-1}. \quad [22], (6)$$

The mode of formation of $R_{kt}$ from $c_t$ may be summarized by

$$R_{t-2,\,t}=-(c_1+c_2+\cdots+c_{t-1}),$$
$$R_{t-4,\,t}=c_1c_3+(c_1+c_2)c_4+\cdots+(c_1+c_2+\cdots+c_{t-3})c_{t-1}, \quad [23]$$
$$\vdots \qquad\qquad \vdots$$
$$R_{0t}=(-)^{t-2}c_1c_3\cdots c_{t-1}.$$

Further progress requires the development of the explicit formulas for $c_t$ as well as for $T_t$ as a function of $\epsilon$.

### J5. Legendre Polynomials

The method by which $T_t(x)$ and $M_t$ will be obtained involves the use of $k$th degree polynomials in $x$, such as $x(x-1)\cdots(x-k+1)$, where $x$ is a variable ranging by integer steps from $x=0$ to $x=n-1$. In order, however, to expose the pattern of reasoning through which the final result is reached, it is desirable to free the argument of those complications introduced by factorial polynomials by first treating the special case in which the number of data in any finite range is allowed to increase without limit. This process eliminates $n$ from all equations and brings forth a striking analogy between the polynomials of Tchebycheff and Legendre, independently noted by many authors,[19, 20, 47] and it leads, incidentally, to the solution of the most difficult algebraic problem connected with the determination of $T_t(x)$ and $M_t$.

It will be convenient, therefore, to express all functions of $x$ first as functions of $\xi=x/n$, and to define

$$\mathcal{T}_k(\xi)=\lim_{n\to\infty}T_k(x)/n^k,$$
$$\int_0^1 d\xi\cdots=\lim_{n\to\infty}[1/n]\sum_{x=0}^{n-1}\cdots, \quad [24]$$
$$\mathfrak{M}_k=\lim_{n\to\infty}M_k/n^{2k+1},$$

in terms of which the fundamental Eqs. [1] and [2] become the continuous orthogonality relations over the interval $(0\leqslant\xi\leqslant1)$ for the $k$th degree polynomials $\mathcal{T}_k(\xi)$ with leading term $\xi^k$:

$$\int_0^1 d\xi\cdot\mathcal{T}_k(\xi)\mathcal{T}_t(\xi)=0,\quad (k\neq t),$$
$$\int_0^1 d\xi[\mathcal{T}_t(\xi)]^2=\mathfrak{M}_t. \quad [25]$$

Apart from the change of interval from $(-1,+1)$ to $(0,1)$, and the change of normalization to make the leading coefficient unity, Eq. [25] is a characteristic equation for the Legendre polynomials.

The process by which $\mathcal{T}_t(\xi)$ is found explicitly begins with the observation that any power $\xi^k$ may be expressed as a sum of $\mathcal{T}_i(\xi)$ with $i\leqslant k$. From this it follows that $\int_0^1 d\xi\cdot\xi^k\mathcal{T}_t(\xi)$ must vanish for all $k<t$, and, since $\xi^t$ contains a term in $\mathcal{T}_t(\xi)$ with unit coefficient, one may derive, from Eq. [25], the equations

$$\int_0^1 d\xi\cdot\xi^k\mathcal{T}_t(\xi)=0,\quad (0\leqslant k<t),$$
$$\int_0^1 d\xi\cdot\xi^t\mathcal{T}_t(\xi)=\mathfrak{M}_t. \quad [26]$$

Writing the Legendre polynomials explicitly in powers of $\xi$, one obtains

$$\mathcal{T}_t(\xi) = \mathfrak{M}_t \sum_{i=0}^{t} w_{it}\xi^i, \qquad [27]$$

and it thus appears that Eq. [26] is a set of $(t+1)$ linear equations to determine the $(t+1)$ unknowns $w_{0t}, w_{1t}, \cdots w_{tt}$. The equations for $w_{it}$ are constructed by substituting Eq. [27] into Eq. [26] and by making use of

$$\int_0^1 d\xi \cdot \xi^{i+k} = 1/(i+k+1):$$

$$\sum_{i=0}^{t} w_{it}/(i+k+1) = 0, \quad (0 \leqslant k < t),$$
$$\sum_{i=0}^{t} w_{it}/(i+k+1) = 1, \quad (k=t). \qquad [28]$$

Since the $w_{it}$ will recur in the determination of $T_t(x)$ and $M_t$, it is useful to solve Eq. [28] at this point. So simple are the coefficients of this set of linear equations that it would be quite awkward[12] to forego the knowledge of their special construction and to attempt to apply mathematical induction to the general solution in terms of determinants. It is appropriate, instead, to recall a device in use before the invention of determinants, for equations which showed the "persymmetric" property.

Consider the function of $z$ defined in terms of $w_{it}$, by

$$w(z) = \sum_{i=0}^{t} \frac{w_{it}}{i+z} = \frac{w_{0t}}{z} + \frac{w_{1t}}{1+z} + \cdots \frac{w_{tt}}{t+z}.$$

By combining the $(t+1)$ terms in the sum over the common denominator $z(z+1)\cdots(z+t)$, it is clear that the numerator of the resulting fraction must be a polynomial of degree $t$ in $z$. Since $w(z)$ is required by Eq. [28] to vanish at the $t$ points $z=1, 2, \cdots t$, its numerator must be proportional to $(z-1)(z-2)\cdots(z-t)$; i.e.,

$$w(z) \propto [(z-1)(z-2)\cdots(z-t)]/[z(z+1)\cdots(z+t)].$$

At $z=t+1$ the right side of this relation assumes the value $(t!)^2/(2t+1)!$, while $w(z)$ is known from Eq. [28] to assume the value $w(t+1)=1$. These relations fix the proportionality factor at $(2t+1)!/(t!)^2$ and the value of $w(z)$ at

$$\frac{w_{0t}}{z} + \cdots + \frac{w_{tt}}{z+t} = \frac{(2t+1)!(z-1)\cdots(z-t)}{(t!)^2 z(z+1)\cdots(z+t)}.$$

On multiplying both sides by $(z+k)$, with $k$ an integer between zero and $t$, one observes that a corresponding factor is cancelled in the denominator on the right and in the term $w_{kt}/(z+k)$ on the left. If $z$ is then made to approach $-k$, every term on the left but the latter vanishes, so that

$$\lim_{z\to -k} (z+k)w(z) = w_{kt} = (-)^{k+t}[(2t+1)!(k+t)!]/[(t!)^2(k!)^2(t-k)!]. \qquad [29]$$

This relation enables one to determine $\mathfrak{M}_t$ at once, for, in Eq. [27], the leading coefficient in $\mathcal{T}_t(\xi)$, $\mathfrak{M}_t w_{tt}$, has been defined as unity. Taking this fact into account, one may summarize the results of Eqs. [27] and [29] as

$$\mathfrak{M}_t = 1/w_{tt} = (t!)^4/[(2t)!(2t+1)!],$$
$$(-)^t \binom{2t}{t} \mathcal{T}_t(\xi) = \sum_{k=0}^{t} (-)^k \binom{2k}{k}\binom{t+k}{2k}\xi^k. \qquad [30]$$

As an application of these formulas one may consider the limit of the Tchebycheff recursion Eq. [19], on division by $n^{t+1}$ and use of Eqs. [24], [30], and $\epsilon = x-(n-1)/2$

$$\mathcal{T}_{t+1}(\xi) = (\xi - \tfrac{1}{2})\mathcal{T}_t(\xi) - [t^2/4(2t^2-1)]\mathcal{T}_{t-1}(\xi),$$

which is a form of the well-known Gaussian recursion for Legendre polynomials.

### J6. Explicit Formulas

A preliminary suggestion for determining $T_t(x)$ might be to attempt to carry through the same process used for $\mathcal{T}_t(\xi)$ in terms of powers of $x$. But then it would become necessary to effect sums of the form

$$(1/n) \sum_{x=0}^{n-1} (x^{k+i}/n^{k+i})$$

in analogy with $\int_0^1 d\xi \cdot \xi^{k+i}$, and the remainder of the problem would turn upon the evaluation of persymmetric determinants of Bernouilli numbers, which is not entirely a simple matter. It seems appropriate, therefore, to try to express $T_t(x)$ in terms of polynomials that can readily be summed over integer values of $x$ from 0 to

$(n-1)$. One thinks, at once, of binomial coefficients, which have a role in problems of summation and differencing analogous to that of the powers in integration and differentiation.

The binomial coefficient $\binom{x}{t}$ may be considered as the polynomial of degree $t$ in $x$ with leading coefficient $1/t!$ given by

$$\binom{x}{t}=x(x-1)\cdots(x-t+1)/t!. \quad [31]$$

It is thus defined for *all* values of $x$ but only for positive integer values of $t$. It reduces to zero for $x=0,\ 1,\ \cdots(t-1)$ and to positive integer values for integer values of $x$ greater than or equal to $t$. Its importance here comes from the readily verified binomial recursion formula

$$\Delta\binom{x}{t+1}=\binom{x+1}{t+1}-\binom{x}{t+1}=\binom{x}{t}. \quad [32]$$

It is convenient to be able to extend the definition of $\binom{x}{t}$ to negative integer values of $t$, and this result is accomplished through Eq. [32];

$$\binom{x}{0}=\binom{x+1}{1}-\binom{x}{1}=(x+1)-x=1,$$

$$\binom{x}{-1}=\binom{x+1}{0}-\binom{x}{0}=0,$$

$$\binom{x}{-2}=0, \quad \text{etc.}$$

The analogy of $\binom{x}{t}$ to the function $x^t/t!$ is suggested, for example, by the binomial theorem for the latter, and is proved by induction on the basis of Eq. [32], exactly as in the usual form of the theorem:

$$\sum_{t=0}^{j}\binom{x}{t}\binom{z}{j-t}=\binom{x+z}{j}. \quad [33]$$

With the aid of the elementary identity

$$\binom{-x}{t}=(-)^t\binom{x+t-1}{t} \quad [34]$$

the binomial theorem may also take the form

$$\sum_{t=0}^{j}(-)^t\binom{x}{t}\binom{z-t}{j-t}=\binom{z-x}{j}.$$

The most significant property of the binomial coefficient polynomials is the ease with which they may be summed by repeated application of Eq. [32]:

$$\binom{x+1}{t+1}=\binom{x}{t+1}+\binom{x}{t}$$
$$=\binom{x-1}{t+1}+\left[\binom{x-1}{t}+\binom{x}{t}\right]=\cdots$$
$$=\binom{x-k}{t+1}+\left[\binom{x-k}{t}+\binom{x-k+1}{t}+\cdots+\binom{x}{t}\right].$$

In particular, if $x$ ranges over integer values from 0 to $(n-1)$,

$$\binom{n}{t+1}=\sum_{x=0}^{n-1}\binom{x}{t};\quad \binom{n+t}{t+1}=\sum_{x=0}^{n-1}\binom{x+t}{t},$$

which may be rewritten to bring out their resemblance to the corresponding continuous formula as

$$(1/n)\sum_{x=0}^{n-1}\left[\binom{x}{t}\Big/\binom{n-1}{t}\right]=1/(t+1)$$
$$=(1/n)\sum_{x=0}^{n-1}\left[\binom{x+t}{t}\Big/\binom{n+t}{t}\right]. \quad [35]$$

In fact, as $n$ increases without limit, and $\binom{x}{t}\Big/\binom{n-1}{t}$ as well as $\binom{x+t}{t}\Big/\binom{n+t}{t}$ approach $(x/n)^t=\xi^t$, Eq. [35] approaches

$$\int_0^1 d\xi\cdot\xi^t=1/(t+1).$$

Furthermore, by means of the identities

$$\binom{x}{i}\binom{x+k}{k}=\binom{i+k}{k}\binom{x+k}{i+k},$$

and

$$\binom{i+k}{k}\binom{n+k}{i+k+1}$$
$$=n\binom{n-1}{i}\binom{n+k}{k}\Big/(i+k+1), \quad [36]$$

which follow directly from the definition [31], it is easy to derive from Eq. [36] the important result

$$(n-1)\sum_{x=0}^{n-1}\left[\binom{x}{i}\Big/\binom{n-1}{i}\right]\left[\binom{x+k}{k}\Big/\binom{n+k}{k}\right] = 1/(i+k+1). \quad [37]$$

This equation approaches

$$\int_0^1 d\xi\cdot\xi^i\xi^k = 1/(i+k+1)$$

in the limit of large $n$.

One is now equipped to retrace the argument that led to Eq. [28] in the case of $\mathcal{T}_t(\xi)$. Just as in the derivation of Eq. [26], one may here expand the polynomials

$$\left[\binom{x+k}{k}\Big/\binom{n+k}{k}\right] \quad \text{and} \quad \left[\binom{x}{i}\Big/\binom{n-1}{i}\right]$$

in sums of Tchebycheff polynomials; Eqs. [1] and [2] will yield, just as did their analog Eq. [25], the results

$$(1/n)\sum_{x=0}^{n-1} T_t(x)\left[\binom{x}{i}\Big/\binom{n-1}{i}\right]=0, \quad (i<j),$$
$$(1/n)\sum_{x=0}^{n-1} T_t(x)\left[\binom{x+k}{k}\Big/\binom{n+k}{k}\right]=0, \quad (k<j), \quad [38]$$

and

$$(1/n)\sum_{x=0}^{n-1} T_t(x)\left[\binom{x}{t}\Big/\binom{n-1}{t}\right] = M_t\Big/\left[t!n\binom{n-1}{t}\right],$$
$$(1/n)\sum_{x=0}^{n-1} T_t(x)\left[\binom{x+t}{t}\Big/\binom{n+t}{t}\right] = M_t\Big/\left[t!n\binom{n+t}{t}\right]. \quad [39]$$

These equations form the analog of Eq. [26]. Following the lead of Eq. [27] one may convert Eqs. [38] and [39] into linear equations for $f_{kt}$ and $g_{it}$, the coefficients of $\binom{x+k}{k}$ and $\binom{x}{i}$ in the appropriate expansions of $T_t(x)$,

$$T_t(x) = \left\{M_t\Big/\left[t!n\binom{n-1}{t}\right]\right\} \times\sum_{k=0}^{t} f_{kt}\binom{x+k}{k}\Big/\binom{n+k}{k}$$
$$= \left\{M_t\Big/\left[t!n\binom{n+t}{t}\right]\right\} \times\sum_{i=0}^{t} g_{it}\binom{x}{i}\Big/\binom{n-1}{i}. \quad [40]$$

The outer factor has been chosen to cancel the right side of Eq. [39]. One must then substitute these relations into Eqs. [38] and [39] in the order in which they are written, to obtain the simultaneous linear equations for $f_{kt}$ and $g_{it}$ after effecting the sum over $x$ on the left by means of Eq. [37]:

$$\sum_{k=0}^{t} f_{kt}/(i+k+1) = 0 = \sum_{i=0}^{t} g_{it}/(i+k+1), \quad (i<t) \quad (k<t)$$
$$\sum_{k=0}^{t} f_{kt}/(t+k+1) = 1 = \sum_{i=0}^{t} g_{it}/(t+i+1). \quad [41]$$

These equations for $f_{kt}$, $g_{it}$ are not merely *analogous* to those for $w_{it}$ in Eq. [28], they are actually *identical* to them. This special simplification was brought about by throwing the summation formula in Eq. [37] into a form where the result of summing terms with $n$-dependent factors was quite independent of $n$; this, in turn, was secured by the use of the *pair* of polynomials

$$\binom{x+k}{k}\Big/\binom{n+k}{k} \quad \text{and} \quad \binom{x}{i}\Big/\binom{n-1}{i}$$

which approach the monomials $\xi^k$ and $\xi^i$ in the limit of large $n$.

The solution of the problem of the explicit form of $\mathcal{T}_t(\xi)$ contained in Eq. [29] may now be directly taken over into the problem at hand, through the identity

$$w_{it} = f_{it} = g_{it}. \quad [42]$$

The first application of Eqs. [29] and [42] appears in the determination of $M_t$ by an argument similar to that which led to Eq. [30]; namely, the leading coefficient of $x^t$ in Eq. [40] must be unity:

$$w_{tt}M_t \Big/ \left[t!n\binom{n-1}{t}\binom{n+t}{t}\right] = 1,$$

or

$$M_t = (t!)^2\binom{n+t}{2t+1}\Big/\binom{2t}{t}. \quad [43], (16)$$

At once, the quantities $c_t$ may be found and applied to Eqs. [17]–[23],

$$c_t = M_t/M_{t-1} = [t^2(n^2-t^2)]/[4(4t^2-1)], \quad [44]$$

as exemplified by the values in Eq. [21].

For the final step in the explicit formula for $T_t(x)$, the $w_{it}$ are substituted from Eqs. [42] and [29] into Eq. [40] in terms of binomial coefficients:

$$T_t(x)/M_t = \left\{1\Big/\left[t!\binom{n+t}{2t+1}\right]\right\} \times \sum_{i=0}^{t}\binom{t-n}{t-i}\binom{t+i}{i}\binom{x}{i}$$

$$= \left\{(-)^t\Big/\left[t!\binom{n+t}{2t+1}\right]\right\} \times \sum_{k=0}^{t}\binom{t+n}{t-k}\binom{t+k}{t}(-)^k\binom{x+k}{k}. \quad [45]$$

This expression gives the weighting factor for the data in Eq. [5] in terms of a small number of integer factors. The ratio of Eqs. (15) and (16) in Section C3 may be shown to be identical with Eq. [45].

As emphasized in the introduction to this section, results equivalent to Eqs. [43] and [45] have been independently derived by many methods[3, 15, 20, 25, 45] ranging from ingenious generalization of numerical tables to polished analysis. The elementary argument above may, nevertheless, be useful through its accessibility to the general reader. It was independently developed[4] in substantially the form presented here.

### J7. Solution in Data Differences

Since successive differences of the data $y(x)$ are often formed in preliminary work to see if polynomial fitting is justified, it is worth transforming the results obtained for $a_k$ into a form in which the latter appear as linear combinations of data differences. In particular, $a_k$ must be linearly constructed from the $k$th differences of $y(x)$ because in the special case that the $k$th differences are actually constant one has the well-known result

$$\Delta^k y(x) = k!a_k, \quad \text{with} \quad \Delta y(x) = y(x+1) - y(x).$$

It will be convenient to introduce fictitious data which vanish identically at integer points outside the actual data range, $0 \leqslant x \leqslant n-1$. In this way one avoids having to specialize the limits of summation in the following treatment, e.g.,

$$\sum_{x=0}^{n-1} y(x)f(x) = \sum_x y(x)f(x),$$

where the summation extends over all integer values of $x$. Under these conditions it is easy to apply "summation by parts" to sums of the form

$$\sum_x y(x)\binom{x}{i} = \sum_x y(x)\Delta\binom{x}{i+1}$$

according to Eq. [32], by merely shifting the index of summation by one unit in the second term of

$$\sum_x y(x)\Delta\binom{x}{i+1} = \sum_x y(x)\left[\binom{x+1}{i+1} - \binom{x}{i+1}\right]$$

$$= \sum_x [y(x) - y(x+1)]\binom{x+1}{i+1}.$$

Thus

$$\sum_{x=0}^{n-1} y(x)\binom{x}{i} = -\sum_x \Delta y(x)\binom{x+1}{i+1}. \quad [46]$$

The disappearance of an explicit "summed part" has been secured by the introduction of the fictitious data, and this method may be iterated as many times as desired since the limits of summation never enter explicitly. On $t$-fold iteration, for example,

$$\sum_{x=0}^{n-1} y(x)\binom{x}{i} = (-)^t \sum_x \Delta^t y(x)\binom{x+t}{i+t}. \quad [47]$$

This result may be used in connection with the explicit formula for $a_t$ obtained by combining Eqs. [5] and [45]:

$$a_t=\sum_{x=0}^{n-1} y(x)\sum_{i=0}^{t}\binom{t-n}{t-i}\binom{t+i}{i}\binom{x}{i}\Big/\left[t!\binom{n+t}{2t+1}\right]$$

$$=\sum_{i=0}^{t}\binom{t-n}{t-i}\binom{t+i}{i}(-)^t\sum_x \Delta^t y(x)\binom{x+t}{i+t}\Big/\left[t!\binom{n+t}{2t+1}\right]$$

$$=\sum_x \Delta^t y(x)(-)^t\sum_{i=0}^{t}\binom{t-n}{t-i}\binom{t+i}{i}\binom{x+t}{i+t}\Big/\left[t!\binom{n+t}{2t+1}\right]$$

by rearrangement of the order of the summations over $i$ and $x$. On reference to the identity in Eq. [36], $\binom{x+t}{i+t}\binom{i+t}{i}$ may be written as $\binom{x+t}{t}\binom{x}{i}$, and the $i$-dependent factors may be separately summed with the aid of the binomial theorem Eq. [33] and the identity Eq. [34]:

$$\sum_{i=0}^{t}\binom{t-n}{t-i}\binom{x}{i}=\binom{x+t-n}{t}=(-)^t\binom{n-x-1}{t}.$$

Substitution into the result obtained above for $a_t$ then yields

$$a_t=\sum_x \Delta^t y(x)\binom{x+t}{t}\binom{n-x-1}{t}\Big/\left[t!\binom{n+t}{2t+1}\right]. \qquad [48]$$

The weighting factor in this sum of data differences is proportional to the polynomial of degree $t$, $\binom{x+t}{t}$, which vanishes at $x=-1, -2, \cdots, -t$ and also to the polynomial of degree $t$, $\binom{n-x-1}{t}$, which vanishes at $x=n-1, n-2, \cdots n-t$. The advancing difference $\Delta^t y(x)$ is formed from data $y(x)y(x+1)\cdots y(x+t)$. Therefore $\Delta^t y(x)$ vanishes identically for $x<(-t)$ or $x>(n-1)$, for only fictitious evanescent quantities contribute to its value. Because, furthermore, the weighting factors of $\Delta^t y(x)$ in Eq. [48] vanish for $(-t)\leqslant x<0$ and for $(n-t-1)<x\leqslant(n-1)$, the limits of summation over $x$ in Eq. [48] may be fixed at $x=0$ and $x=n-t-1$. No fictitious data then appear in the final formula:

$$t!a_t=\sum_{x=0}^{n-t-1}\Delta^t y(x)\binom{x+t}{t}\binom{n-x-1}{t}\Big/\binom{n+t}{2t+1}. \qquad [49], (57)$$

This result is given in Section D as Eq. (57) with symbols defined by Eqs. (46), (53), and (58).

Equation [49] has the special advantage that all weighting factors of data differences are positive; the weights sum to unity, as may be seen by inserting a set of data with exactly equal differences of degree $t$. The weights are symmetrical about the point $x=(n-t-1)/2=\nu$ in a manner evident upon introduction of the variable $\epsilon=x-\nu$, and the central differences $\delta^t y(\epsilon)=\Delta^t y(x)$:

$$t!a_t=\sum_{\epsilon=-\nu}^{+\nu}\delta^t y(\epsilon)\binom{\nu+t+\epsilon}{t}\binom{\nu+t-\epsilon}{t}\Big/\binom{2\nu+2t+1}{2t+1}. \qquad [50], (52)$$

This expression appears in Section D as Eq. (52), with the symbols defined by Eqs. (51) and (53). Here is illustrated the analog of the pair-factor symmetry of Eq. [11] which enables the data differences to be grouped in pairs and which halves the number of weighting factors that need be tabulated. Considered as a function of $\epsilon$, the weighting factor is a polynomial of degree $2t$, symmetrical about the origin where it has its maximum and diminishing to zero just beyond the interval of summation. There it has its roots at the points $\epsilon=\pm(\nu+1), \pm(\nu+2), \cdots \pm(\nu+t)$.

One may therefore extend the summation in Eq. [50] from $\epsilon=-(\nu+t)$ to $\epsilon=(\nu+t)$ without altering its value; application of the same technique of summation by parts which led from Eq. [46] to Eq. [49] transforms Eqs. [49] and [50] into

$$\begin{aligned} t!a_t &= \sum_{x=0}^{n-1} y(x)(-)^t\Delta^t\binom{x}{t}\binom{n-1-x+t}{t}\Big/\binom{n+t}{2t+1} \\ &= \sum_{\epsilon} y(\epsilon)(-)^t\delta^t\binom{\nu+t+\epsilon}{t}\binom{\nu+t-\epsilon}{t}\Big/\binom{2\nu+2t+1}{2t+1}. \end{aligned} \quad [51]$$

Comparison with Eq. [5] yields a new form for the explicit Tchebycheff polynomials:

$$\begin{aligned} T_t(x)/M_t &= (-)^t\Delta^t\binom{x}{t}\binom{n-1-x+t}{t}\Big/\left[t!\binom{n+t}{2t+1}\right] \\ &= (-)^t\delta^t\binom{\nu+t+\epsilon}{t}\binom{\nu+t-\epsilon}{t}\Big/\left[t!\binom{2\nu+2t+1}{2t+1}\right]. \end{aligned} \quad [52]$$

Some insight into these expressions, first derived by Jordan,[20] may be afforded by passing to the limit of large $n$ according to Eq. [24]. Observing that $\Delta^t \to n^{-t}(d/d\xi)^t$ and cancelling a factor of $n^{-t-1}$ on both sides, one obtains the result

$$\mathcal{T}_t(\xi)/\mathfrak{M}_t = [(-)^t(2t+1)!/(t!)^3] \times (d/d\xi)^t\xi^t(1-\xi)^t,$$

or

$$\mathcal{T}_t(\xi) = [(-)^t t!/(2t)!] \times (d/d\xi)^t[\xi^t(1-\xi)^t]. \quad [53]$$

Apart from range and from normalizing factor, this is just the well-known formula of Rodrigues for the Legendre polynomials. Jordan's Eq. [52] is thus disclosed as the analog of Rodrigues' theorem for the Tchebycheff polynomials.

Just as in the continuous case, the process of differencing produces roots near the maxima and minima of the differenced function; because the roots of the functions differenced in Eq. [52] lie between $+(\nu+t)$ and $-(\nu+t)$, their maxima and minima must be similarly located. $T_t(\epsilon)$ thus has all its roots in the range of the data and can only increase or decrease steadily and without limit outside the data range. To this circumstance must be ascribed the extreme unreliability of an extrapolated point; for, according to Eq. [9], the corresponding probable error increases essentially with the $t$th power of the distance from the range of data.

The simple construction of the weighting factors in Eqs. [49] and [50], as well as their positive sign, recommend these formulas for the determination of $a_t$. For the same reasons, Eq. [52] appears to be a practical means to evaluate $T_t(\epsilon)$.

### J8. Computation of Tables

A fundamental consideration in making tables for numerical realization of the theory is the representation of the rational numbers involved as the quotient of the smallest possible integers. The simplest problem of this type occurs in connection with Eq. [49], where the weighting factors $\binom{x+t}{t}\binom{n-x-1}{t}\Big/\binom{n+t}{2t+1}$ might well appear unreduced to lowest terms. Much of the reduction, however, has already been accomplished through merely having written this expression in terms of the binomial coefficient function. For the latter is known to be an integer for all integer values of its argument, although when explicitly written as a quotient of factorials it might not appear so at first glance.

The effect of the binomial coefficients has been to cancel in numerator and denominator any common factors independent of $n$. All that now remains is to examine whether for certain values of $n$ common factors may be extracted from $\binom{x+t}{t}\binom{n-x-1}{t}$ for all integer values of $x$ in the range $0 \leqslant x \leqslant n-t-1$. If such factors exist, they must surely be cancelled by the denominator, since the latter is the sum over all $x$ in the range $0 \leqslant x \leqslant n-t-1$ of the values of the numerator, i.e.,

$$\binom{n+t}{2t+1} = \sum_{x=0}^{n-t-1} \binom{x+t}{t}\binom{n-x-1}{t},$$

a consequence of the remark following Eq. [49].

Before carrying out the factorization it is worth observing that such a process will achieve the same effect in $T_t(x)/M_t$ in Eqs. [5] and [45]. This is a consequence of Eqs. [52] and [45], according to which

$$\sum_{i=0}^{t} \binom{t-n}{t-i}\binom{t+i}{i}\binom{x}{i} = (-)^t \Delta^t \binom{x}{t}\binom{n-x+t-1}{t}. \quad [54]$$

On the left is a function proportional to $T_t(x)$, and on the right occur the $t$th differences of the numbers $\binom{x+t}{t}\binom{n-x-1}{t}$, which give the weighting factors for the data differences in Eq. [49]. Since the $t$ initial and the $t$ terminal values of the latter sequence are null, the sequence can be obtained entirely from the $t$th differences appearing in Eq. [54] by $t$-fold iteration of the process of forming, in order, the partial sums and by adjoining a null value to the beginning and end. If the sequence $\binom{x+t}{t}\binom{n-x-1}{t}$ for $0 \leqslant x \leqslant (n-1)$ has all its common factors divided out, its $t$th differences will then also be freed of common factors. This is, of course, not generally true, but is a consequence of the $t$ initial and final null values; for if the $t$th differences had a common factor, their partial sums, and hence the primary sequence, would have to possess the same common factors.

One concludes, therefore, that the process of reduction to lowest terms of the complicated expressions for $T_t(x)/M_t$ is accomplished by cancelling from numerator and denominator exactly the same common factors as those that occur in the relatively simple weighting factors $\binom{x+t}{t}\binom{n-x-1}{t}\Big/\binom{n+t}{2t+1}$ occurring in Eq. [49.]

The reduction of these weighting factors, in turn, rests on the periodic occurrence of multiples of a given prime among the successive integers that must be multiplied together to form the numerators of $\binom{x+t}{t}$ and $\binom{n-x-1}{t}$. It is not difficult to discover the condition on $n$ in order that there exist at least one value of $x$ in the range $0 \leqslant x \leqslant n-t-1$ for which *neither* $\binom{x+t}{t}$ *nor* $\binom{n-x-1}{t}$ can have a prime factor $P$ in their numerators *not* cancelled by a corresponding factor in their denominator $t!$. The result, which formalizes the procedure of the "Sieve of Eratosthenes," may be stated upon writing both $n$ and $t$ in the form of a multiple of $P^q$, with a positive or zero remainder less than $P^q$, i.e.,

$$n = NP^q + n^*, \quad t = TP^q + t^*$$

with

$$0 \leqslant n^* < P^q, \quad 0 \leqslant t^* < P^q.$$

*There exists a common factor of $P$ for every value of $n$ such that $P^q - t^* \leqslant n^* \leqslant t^*$; and if these conditions can be satisfied for some $P$, $t$, $n$ with more than one value of $q$, there is one distinct factor of $P$ for every possible choice of $q$.*

An immediate corollary is that one need consider neither prime factors $P > 2t$ nor powers of such primes $P^q > 2t$. In reducing to lowest terms one might proceed by first setting down all $P^q \leqslant 2t$, and then dividing these quantities into $t$ to obtain the corresponding remainders $t^*$. One must reject all cases for which $P^q - t^* > t^*$, and for those that remain one must permit $n^*$ to assume all values with $P^q - t^* \leqslant n^* \leqslant t^*$. The values of $n$ for which a common factor of $P$ is present are then given by $n = NP^q + n^*$, where $N$ is any integer which makes $n > t$. As an example, consider the most complicated case in the accompanying table, that for $t = 5$, $2t = 10$, and $P(\leqslant 2t) = 7, 5, 3, 2$. In this way there is fashioned a kind of Sieve of Eratosthenes which sifts from the values of $n$, set down in succession, the periodically recurring groups for which the

Model Form 9. Scheme for determining common factors.

| $P$ | $P^q$ | $t^*$ | $P^q-t^*$ | $n^*$ | $n(5)$ | $P$ |
|---|---|---|---|---|---|---|
| 7 | 7 | 5 | 2 | 2, 3, 4, 5 | $7N\pm2$, $7N\pm3$ | 7 |
| 5 | 5 | 0 | 5 | — | — | 5 |
| 3 | 3 | 2 | 1 | 1, 2 | $3N\pm1$ | 3 |
| | 9 | 5 | 4 | 4, 5 | $9N\pm4$ | |
| 2 | 2 | 1 | 1 | 1 | $2N+1$ | 2 |
| | 4 | 1 | 3 | — | — | |
| | 8 | 5 | 3 | 3, 4, 5 | $8N\pm3$, $8N+4$ | |

weighting factors possess in common a given prime divisor. The labor of computation of pair factors for data and for data differences is thereby greatly reduced.

How to apply this process in order to represent $T_t(x)$ as proportional, by a fixed factor, to a sequence of relatively prime integers, and how further to exploit this numerical simplicity throughout least-squares' problems, has been thoroughly explained in preceding sections.

In conclusion, let the author of this section gratefully acknowledge the initiative and the practical evaluation of results of the foregoing research as coming from Professor Birge. Many of the formulas derived here for the first time were, in fact, predicted by him in advance of analytical proof, and without his distinguished guidance, this work would not have been possible.