Department of Statistics

University of Wisconsin
Madison, WI 53706

TECHNICAL REPORT NO. 344

October 1973

GERGONNE'S 1815 PAPER ON THE DESIGN AND ANALYSIS OF POLYNOMIAL REGRESSION EXPERIMENTS

by

Stephen M. Stigler

University of Wisconsin, Madison

Typist: Candy Smith

Gergonne's 1815 Paper on the Design and Analysis of Polynomial Regression Experiments

by

Stephen M. Stigler[1]

University of Wisconsin, Madison

Summary.

Following some remarks on the early history of the design of experiments, a seemingly unknown 1815 paper of Gergonne's on polynomial regression is discussed, and a translation of the paper presented.

1. Introduction.

The design of experiments may well be oldest of all the fields of statistics. Examples of planned, controlled experiments considerably predate the first attempts at formal analysis of random data, going back at least to the old testament, perhaps further. Some thoughts on sample size can be found in work of Galen the Physician dating from 150 AD (see Galen (150), pp. 96-119). In the eleventh century, many modern principles of design were spelled out by the famous

---

[1] This work was partially supported by the Wisconsin Alumni Research Foundation, and, while the author was on leave in the the Department of Statistics at the University of Chicago, by NSF Research Grant GP 32037x.

Arabic doctor, scientist, and philosopher Avicenna, in the second volume of his Canon of Medicine, the leading medical text for nearly eight centuries. Avicenna listed seven rules for medical experimentation, stressing the need for controls and replication, the danger of confounding effects, the necessity of varying one factor at a time, and the wisdom of observing the effects for many differing factor levels. (See Crombie (1952), pp. 89-104).

It is therefore not surprising that work on the design of regression experiments preceded the introduction of the method of least squares. One early example of this can be found in the first volume (1799) of Laplace's Mécanique Céleste (see sections 28-29, chapter 4, book 2), where it is recommended that if random errors are present, a polynomial regression function will be best estimated at a point by spreading the observations in a wide interval about that point. The purpose of this present note is to introduce a paper which was inspired by these sections of the Mécanique Céleste.

The paper we present was written by Joseph-Diez Gergonne of the University of Montpellier, France, editor and founder of the journal Annales de Mathématiques Pures et Appliqueés, and it appeared in his own journal in 1815 under the title "Application de la méthode des moindrés quarrés à l'interpolation des suites". It appeared in the midst of a twenty-year period in which it might be claimed that mathematical statistics advanced further than in any similar period in history:

Legendre's first publication of least squares in 1805, Gauss's linking of least squares to probability in 1809, Laplace's <u>Theorie Analytique des Probabilités</u> in 1812 and its supplements of 1814, 1818, 1820, and Gauss's papers of 1821 and 1823 presenting the so-called Gauss-Markov theorem.

Gergonne's paper was not a landmark of this era; indeed it seems to have completely escaped the attention of all bibliographers of the statistical literature of that time,[1] although Gergonne's journal was widely circulated at the time and must have been read by most European mathematicians. While the paper itself contains no really startling results, it nonetheless is an extremely interesting document in the history of statistics, both as one of the earliest attempts to discuss some of the problems of design and analysis which are inspiring so much research today, and for the insight it gives us into the spread and development of statistical thought in the early years of the nineteenth century.

2. Gergonne's Paper

The problem Gergonne considered is one we would now describe as follows: given a situation where one observes a response which depends upon a single independent variable, and

---

[1] It is not mentioned by Merriman (1877), Gore (1902), or Kendall and Doig (1968), nor is Gergonne listed in Lancaster (1968). The paper <u>is</u> listed in the Royal Society of London's Catalogue of Scientific Papers 1800-1900, Subject Index (Vol. I, Pure Mathematics) under "Interpolation".

where one wishes to estimate the value of the response function and its derivatives at a single point, how should one select the values of the independent variable at which the experiment will be performed, when random errors in the observed responses are expected. Gergonne's treatment of this problem is interesting but not profound. He began with a general discussion of the problem of interpolation, viewed both geometrically (in terms of points and curves) and algebraically (in terms of variables and functions). He observed that even with no errors present, the problem is somewhat indeterminate, but that with sufficiently many observations this would not cause serious difficulty, and one could conveniently fit a simple polynomial model to the data.

The first method of fitting he discussed is the one which was most prevalent at the time: fit a polynomial consisting of as many terms as there are data points. Gergonne was aware of the difficulties this method presented when the number of observations was large, but he went on to extend an argument of Legendre's analyzing the effect an error in a single observation would have on the derivatives of the interpolating polynomial, concluding that Laplace's advice was sound: within the class of equally spaced designs, accuracy increases with increasing spread and more distant spacing.

Gergonne then noted that the only way a widely spaced experiment could be achieved over a narrow range would be by discarding (or declining to take) observations, and suggested that a much more sensible procedure would be to use least squares. He developed the normal equations for polynomial regression, discussed the numerical simplification which came with an equally spaced design, and showed how any design may, by the appropriate transformation on the independent variable, be transformed to an equal spacing design to simplify calculations. The paper closes by posing a problem which cannot be said to be well solved today: "we know that a number of points, however many, are located <u>near</u> a parabolic curve of unknown fixed degree, and we wish to know the most likely ['plus probablement'] value of the degree of this curve."

In many respects, the paper belongs more to data analysis than statistics. By not explicitly introducing any probability structure, Gergonne was following the example of Legendre rather than that of Gauss or Laplace, thus illustrating that true scientific innovation is often very slow in catching on: the <u>technique</u> of least squares was not in principle greatly different from many of the techniques which preceded it, and it was its computational simplicity coupled with the authority of Gauss and Laplace which led to its early widespread adoption. The truly innovative work of Gauss and Laplace, incorporating probability models as a foundation and justification for the adoption of this technique, was not well understood by Gergonne and many others at this time, but was only very slowly spread as the work was extended and improved over the following century.

Gergonne's paper was, however, novel in a number of respects. It presents what may be the first explicit application of the principle of least squares to a general polynomial regression model. More significantly, it is one of the earliest attempts to deal mathematically with a design problem in a regression framework, showing that the planning of experiments was already being considered in mathematical terms in 1815. The paper also describes the use of coding as a device for simplifying computations, and it displays a surprisingly modern feel for the problems of statistical analysis and model fitting, including a realization that polynomial models are ill-suited for extrapolation and an understanding that no single method of analysis gives a uniquely best answer.

It is likely that Gergonne's paper, written in the south of France by an educated man who followed work in all the major intellectual centers of Europe, is more representive of the general level of statistical thought in Europe than is the work of giants such as Laplace and Gauss. While the paper seems to have never been cited in the statistical literature, it would be a mistake to conclude that it must then necessarily have had no influence on the statistical practice of the time. To see how this may be, we turn to Gergonne's journal and its role in the development of applied mathematics.

3. Gergonne and his Journal.

Joseph-Diez Gergonne (1771-1859) is best known as the co-founder and editor of the Annales de Mathématiques Pures et Appliquées. Gergonne founded his Annales in 1810, at which time it was the first and only journal devoted entirely to mathematics and its applications. It remained the only such journal until 1826 and the first appearance of Crelle's journal.

Gergonne's Annales was a remarkably lively and broad journal. By the time he became rector of the University of Montpellier and ceased publication of the journal in 1831, articles had been published on nearly every branch of pure mathematics, and on a wide range of applications including optics, circulation of the blood, sundials, economics, political science, celestial mechanics, gambling, and law. The list of contributers includes some of the foremost mathematicians of the time: Cauchy, Poisson, Ampère, Abel, Poncelet, and Galois. Gergonne himself contributed over 200 papers, a majority in geometry, the field Gergonne was most interested in and the one in which he is most recognized. Many of his papers were published anonymously, attributed to "un abonné" ("A subscriber"); these included his only other effort in statistics, a paper on the estimation of means which appeared in 1821.

Of Gergonne's own work, the only portion which receives recognition in most histories of mathematics is his work in geometry, where he became embroiled in a bitter priority fight with Poncelet over the discovery of reciprocal polars and the principal of duality. In many respects his achievements as editor were greater than those as author; his journal was widely read and had a lasting influence on the development of mathematics far beyond that of the individual articles.

Bibliographical Note:

The most accessible treatment of Gergonne's life and work is the article by D.J. Struik in the Dictionary of Scientific Biography (Struik (1972)) (with references), although Struik has overlooked a number of important sources, including Bouisson (1859) and Henry (1881), making his account incomplete and incorrect in some minor respects, such as the date of Gergonne's death and the spelling of his middle name.

In the following translation of Gergonne's paper, an effort has been made not to introduce any modern statistical terminology and to accurately reflect Gergonne's thinking. To ease the way for modern readers, however, some of the mathematical terminology has been updated (examples: "polynomial function" for "fonction complète, rationelle et entière" and "derivatives" for "coefficiens différentiels". All italics, including those in the quotation from Laplace, are Gergonne's, as are the footnotes unless otherwise noted. Some readers may be unfamiliar with the term "osculating circle", which is a geometric measure of curvature at a point - a geometric analogue of a second derivative.

References

[1] Bouisson (1859). Notice Biographique sur Joseph-Diez Gergonne. Academie des Sciences et Lettres de Montpellier, Memoires de la Section de Medecine. Tome 3 (1858-1862), 191-202.

[2] Crombie, A.C. (1952). Avicenna on medieval scientific tradition. In Avicenna: Scientist and Philosopher, A Millenary Symposium, Ed. by G.M. Wickens, London: Luzac & Co.

[3] Galen (150). Galen on Medical Experience Ed. by R. Walzer, London: Oxford University Press, 1944.

[4] Gergonne, J.D. (1815). Application de la méthode des moindre quarres a l'interpolation des suites. Annales des Math. Pures et Appl. Tome 6, 242-252.

[5] Gore, J.H. (1902). A Bibliography of Geodesy, (2nd Ed.) Appendix No. 8 (pp. 427-787) to Report of the Superintendent of the Coast and Geodetic Survey for 1902, Washington: U.S. Coast and Geodetic Survey, 1903.

[6] Henry, C. (1881). Supplément a la bibliographie de Gergonne. Bullettino di Bibliografia e di Storia delle Scienze Matematiche e Fisiche (pub. by Boncompagni). Tomo 14, 211-218.

[7] Kendall, M.G. and Doig, A.G. (1968). Bibliography of Statistical Literature Pre-1940. Edinburgh and London: Oliver and Boyd.

[8] Lancaster, H.O. (1968). Bibliography of Statistical Bibliographies. Edinburgy and London: Oliver and Boyd.

[9] Laplace, P.S. (1799). Mecanique Celeste, Tome I, Translated as Celestial Mechanics by N. Bowditch, New York: Chelsea Pub. Co., 1966. (Sections 28-29 of Book II are found on pp. 407-417 of Vol. I of the translation.)

[10] Merriman, M. (1877). A list of writings relating to the method of least squares, with historical and critical notes. Transactions of the Connecticut Academy of Arts and Sciences 4, 151-232.

[11] Struik, D.J. (1972). Gergonne, Joseph Diaz [sic]. Article in Volume 5, p. 367-369, of Dictionary of Scientific Biography (Ed. by C.C. Gillispie) New York: Charles Scribner's Sons.

The Application of the Method of Least Squares to the Interpolation of Sequences

by

J.D. Gergonne

(Translated by Ralph St. John, Bowling Green State University)

When a function of a single variable is known, we can always determine rigorously and directly the values of the function and of its various derivatives at a given value of the independent variable. Similarly, given a curve we can always, for any abscissa, obtain the ordinate, the tangent, the osculating circle, etc.

Just as instead of giving a curve we can give only a certain number of its points, we can similarly instead of giving a function of a variable give only the values this function takes for a certain number of values of the independent variable, and subsequently ask what are the values of the function and its various derivatives for any other value of this variable. Similarly we could ask for a given abscissa what are the ordinate, the tangent, the osculating circle, etc. of a curve about which we know only a certain number of points. This constitutes the problem of the <u>interpolation of sequences</u>.

This problem obviously reduces to recovering from the given values, the function from which they were obtained, or from the given points, the plot of the curve on which we assume they are located. However, the problem is indeterminate for, given non-consecutive points, even an infinite number of them, we can always pass through them an infinity of different curves.[1]

These curves could very well differ notably from one another in certain parts of their range; the same difference will be observed in the ordinates, tangents, osculating circles, etc. for a given abscissa. However, we note that if the given points are close enough to each other, the curves which include them will not differ greatly over this interval, at least if none of the curves has within this interval an asymptote parallel to the axis of the ordinates. We also note that these given points can always be numerous enough, and, at the same time, sufficiently close to each other, that the differences between these curves become almost indistinguishable. The ordinates which result from a single abscissa within this range will therefore be essentially equal; however, the difference between the tangents can be more sensitive, that between the osculating circles even more so, and so forth.

We conclude from this that, if functions of diverse form have the same value for certain known neighboring values of the independent variable without becoming infinite for any value included between these, then these functions will take on values

---

[1] We can consult on this subject a dissertation on page 252 of volume 5 of this journal. [Trans. note: The article referred to, "Considerations philosophiques sur l'interpolation", is by Gergonne but contains no material relevant to statistics.]

scarcely different for other values of the independent variable included within the above interval. However, this will not be the case for the derivatives of the various functions, which can differ more and more as the corresponding order increases.

We can therefore, without noticeable error, arbitrarily adopt one of thse functions as the desired function; similarly when many curves which pass through the same points have only slight differences, we can assume that any one of these is really the curve on which these points lie.

Since the curve or the function can be selected in an infinity of different ways, it it convenient to select the simplest way, that is, the parabolic curve or the polynomial function that graphically represents it. This choice is well founded since it is known that all finite functions of a finite variable can always be expressed in a series of increasing powers of this variable.

The procedure we have just arrived at is also that which is commonly followed; we assume that the ordinate of the desired curve is a polynomial function of the abscissa, into which we allow as many terms as there are sets of given values; the coefficients of these terms are unknown, and we determine them by assuming that the curve passes through the given points. Once these coefficients are determined, it is a simple matter to calculate the ordinate and the derivatives for any abscissa. However, we can rely on the values obtained from this formula only when it is applied to an abscissa within the interval containing the given points, and also not too close to the largest or the smallest.

This method, which was employed by Mr. Laplace in his memoir Recherche des orbites des comètes[1], contains a source of error in the supposition that the curve is a parabolic curve. Nevertheless, if we could rigorously believe in the given values of the function, and if these values were very numerous and very close to each other, then what we have said abouve shows that the error resulting from this supposition would never be very large.

However, this is not always the case. The discrete values of the function, which we have used to construct our formula, are often deduced from experience or from observations subject to limited precision. Thus, as Mr. Legendre has observed[2], it often happens that the errors which affect these observations can have more and more influence on the final solution and on the results we deduce from this solution, as more and more values are obtained.

Assume that we have plotted an arbitrary curve, and that we have obtained from it many ordinates very close to each other. Suppose we have subjected these ordinates to very small changes, sometimes positive and sometimes negative, and subsequently we attempt to pass a continuous curve through these altered ordinates. We will easily see that, even if these alterations have had only a

---

1 See the Mémoires de l'Académie des Sciences, Paris, for 1780.

2 See his Nouvelles méthodes pour la détermination des orbites des comètes, Paris, 1806, p. iv.

very small influence on the size of intermediate ordinates, that is not the case with regard to the tangent, which may have undergone a notable change for the same abscissa, and this change may be even more noticeable with regard to the osculating circle.

These graphical observations can easily be confirmed by calculations. Suppose we have an odd number of given ordinates corresponding to equidistant abscissas, and assume that this common distance is one. Let zero be the abscissa and b the ordinate at the middle; 1, 2, 3 ... the abscissas and b^1, b^2, b^3 ... the ordinates which follow; -1, -2, -3, the abscissas and b_1, b_2, b_3,... the ordinates which precede. We wish to obtain the various derivatives at zero. We obtain for the case of three ordinates

$$\frac{dy}{dx} = \frac{b^1 - b_1}{2}, \quad \frac{d^2y}{dx^2} = (b^1 + b_1) - 2b;$$

for the case of five ordinates

$$\frac{dy}{dx} = \frac{8(b^1 - b_1) - (b^2 - b_2)}{12}, \quad \frac{d^2y}{dx^2} = - \frac{30b - 16(b^1 + b_1) + (b^2 + b_2)}{12};$$

for the case of seven ordinates

$$\frac{dy}{dx} = \frac{45(b^1 - b_1) - 9(b^2 - b_2) + (b^3 - b_3)}{60},$$

$$\frac{d^2y}{dx^2} = - \frac{490b - 270(b^1 + b_1) + 27(b^2 + b_2) - 2(b^3 + b_3)}{180};$$

and so forth.

Suppose that the other ordinates are exact and that the ordinate b^1 is in error by the quantity β. Let $E\frac{dy}{dx}$, $E\frac{d^2y}{dx^2}$ denote the resulting errors in the derivatives at zero. It is easy to see that, in the case of three ordinates

$$E\,\frac{dy}{dx} = \frac{1}{2}\beta,\ E\,\frac{d^2y}{dx^2} = \frac{2}{2}\beta;$$

in the case of five ordinates

$$E\,\frac{dy}{dx} = \frac{2}{3}\beta,\ E\,\frac{d^2y}{dx^2} = \frac{4}{3}\beta;$$

in the case of seven ordinates

$$E\,\frac{dy}{dx} = \frac{3}{4}\beta,\ E\,\frac{d^2y}{dx^2} = \frac{6}{4}\beta.$$

Therefore the errors in the first order derivative increase as do the numbers $\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \ldots$ and thus tend monotonically to the actual error in the ordinate b^1. Similarly, the error in the second order derivative is double that of the first order derivative.

Mr. Legendre was therefore justified in saying that in increasing the number of values, we exposed ourselves to an increase in the errors in the same proportion. This result assumes that there is only one incorrect ordinate, which excludes all possibility of compensating errors. Moreover, this assumes that the incorrect ordinate is precisely that whose value, exact or not, exerts the most influence on our two derivatives.

Whatever the case, this source of error did not escape the attention of Mr. Laplace. Here are his comments (<u>Mécanique</u>

celeste, Tome I, p. 201)[1]: "These expressions are more precise as there are more observations, and as the interval separating them is smaller. We could therefore use all the neighboring observations for the chosen period, if they were exact, but the errors to which they are subject would lead us to a false result. Therefore, to reduct the influence of these errors, we must increase the interval of the extreme observations as we employ more observations."

It would probably be more correct to say that we must employ observations more and more distant from each other as we employ more observations. We shall see, in effect, that with this procedure we can reduce these errors. Let a be the interval, assumed constant, which separates consecutive values of x, an interval which we previously assumed to be one. Our previous results then become, for three observations

$$E\,\frac{dy}{dx} = \frac{1}{2}\,\frac{\beta}{a},\quad E\,\frac{d^2y}{dx^2} = \frac{2}{2}\,\frac{\beta}{a^2}\,;$$

for five observations

$$E\,\frac{dy}{dx} = \frac{2}{3}\,\frac{\beta}{a},\quad E\,\frac{d^2y}{dx^2} = \frac{4}{3}\,\frac{\beta}{a^2};$$

for seven observations

$$E\,\frac{dy}{dx} = \frac{3}{4}\,\frac{\beta}{a},\quad E\,\frac{d^2y}{dx^2} = \frac{6}{4}\,\frac{\beta}{a^2}\,.$$

---

[1] This passage may be found on p. 411 of volume I of Bowditch's translation. [Trans.]

Therefore, as long as a takes on values which increase more rapidly than does the sequence $\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots$ our errors will continually decrease as we have more and more observations. Suppose, for example, that we increase the value of a according to the positive integers. Let this value be one for the case of three observations. We thus have for three observations

$$E\,\frac{dy}{dx} = \frac{1}{2}\beta,\quad E\,\frac{d^2y}{dx^2} = \frac{2}{2}\beta;$$

for five observations

$$E\,\frac{dy}{dx} = \frac{1}{3}\beta,\quad E\,\frac{d^2y}{dx^2} = \frac{1}{3}\beta;$$

for seven observations

$$E\,\frac{dy}{dx} = \frac{1}{4}\beta,\quad E\,\frac{d^2y}{dx^2} = \frac{1}{6}\beta$$

Thus we see that the errors in the first order derivatives decrease as do the inverse of the positive integers, and that the errors which affect the second order derivatives decrease according to the progression, even more rapid, of the inverse of the triangular numbers. The method of Mr. Laplace is therefore, from this point of view, entirely beyond reproach.

However, suppose we have between two fixed known limits sufficient observations to reduce to a very small value the difference between successive values of x. Following what we have just said, we must discard as many observations as we will use in our search for $\frac{dy}{dx}$ and $\frac{d^2y}{dx^2}$. Thus, this is a serious inconvenience, especially if we have no reason to suspect that the values we discard are worse than those we use. In this manner we deprive ourselves of the compensation of errors upon which we may rely if we use all the values.

While reflecting on this subject, it seemed to me that it was possible, using the method of least squares[1], to reconcile things and to obtain by this method all the precision one can possibly hope for in this situation. Here is the method I believe we should use.

Let a, a_1, a_2... be the values of x, however many, and let b, b_1, b_2,... be the observed corresponding values of y. Let

$$y = A + bx + Cx^2 + Dx^3 + \ldots$$

allowing as many terms in this function as we would have employed using the previously described procedure of discarding observations. We wish to determine the value of the coefficients A, B, C, D,... . If the number of coefficients were equal to the number of observations, we could assign the coefficients values giving zero errors. But this is impossible in this case and we shall be content to minimize the sum of their squares.

---

[1] We know that the method of least squares is based on the principle that the mean value (which is most probable to be nearly exact) of many values near a quantity, is that which, assuming it were correct, would minimize the sum of squares of the errors which affect the other observations. The first printed work in which this method was mentioned is the memoir of Mr. Legendre already cited in a preceding note (1806). In a work published in 1809, Mr. Gauss declared that he has been using a similar method since 1795. Mr. Laplace subsequently showed that this method conforms rigorously to the theory of probability.

These errors are respectively

$$A + Ba + Ca^2 + Da^3 + \dots -b \quad ;$$

$$A + Ba_1 + Ca_1^2 + Da_1^3 + \dots - b_1;$$

$$A + Ba_2 + Ca_2^2 + Da_2^3 + \dots - b_2.$$

We wish to obtain

$$\begin{aligned} &(A + Ba + Ca^2 + Da^3 + \dots -b)^2 \\ +\ &(A + Ba_1 + Ca_1^2 + Da_1^3 + \dots -b_1)^2 \qquad = \text{minimum} \\ +\ &(A + Ba_2 + Ca_2^2 + Da_2^3 + \dots -b_2)^2 \\ +\ &\dots\dots\dots\dots \end{aligned}$$

That is, in differentiating with respect to A, B, C, D, ...

$$\begin{aligned} &(A + Ba + Ca^2 + \dots - b)(dA + adB + a^2dC + \dots) \\ +\ &(A + Ba_1 + Ca_1^2 + \dots - b_1)(dA + a_1dB + a_1^2dC + \dots) \qquad = 0. \\ +\ &(A + Ba_2 + Ca_2^2 + \dots - b_2)(dA + a_2dB + a_2^2dC + \dots) \end{aligned}$$

Because of the independence between A, B, C, ... the multipliers of dA, dB, dC, ... must separately be zero. We abbreviate in general

$$\Sigma a^m = a^m + a_1^m + a_2^m + \dots ;$$

$$\Sigma a^m b = a^m b + a_1^m b_1 + a_2^m b_2 + \dots ;$$

and we obtain these equations

$$\Sigma a^0 A + \Sigma aB + \Sigma a^2 C + \Sigma a^3 D + \ldots = \Sigma a^0 b,$$

$$\Sigma aA + \Sigma a^2 B + \Sigma a^3 C + \Sigma a^4 D + \ldots = \Sigma aD, \qquad (1)$$

$$\Sigma a^2 A + \Sigma a^3 B + \Sigma a^4 C + \Sigma a^5 D + \ldots = \Sigma a^2 D,$$

$$\cdots\cdots\cdots\cdots$$

There are exactly as many equations as there are unknown coefficients A, B, C, D, ... Although the methods previously discussed give values for y and its derivatives of a precision slightly inferior to that of the observations from which they were calculated, we can often hope with this new procedure to improve on the precision of the observations themselves.

The simplest case, and the most frequent, is that in which the values of x increase by a constant difference. Thus we can substitute the natural numbers for this progression. Let there be 2n + 1 known corresponding values of x and y. Let zero be the middle value of x, such that the numerical sequence is

$$-n, -(n-1), \ldots -3, -2, -1, \mp 0, +1, +2, +3, \ldots + (n-1), n .$$

Let Σn^m denote the sum of the m^{th} powers of these integers. We obtain

$$\Sigma a^0 = 2n + 1, \ \Sigma a = 0, \ \Sigma a^2 = 2\Sigma n^2, \ \Sigma a^3 = 0, \ \Sigma a^4 = 2\Sigma n^4, \ldots .$$

Thus equations (1) become

$$(2n+1)A + 2\Sigma n^2 C + \ldots = \Sigma b \quad , \quad 2\Sigma n^2 B + 2\Sigma n^4 D + \ldots = \Sigma ab \ ,$$

$$2\Sigma n^2 A + 2\Sigma n^4 C + \ldots = \Sigma a^2 b, \quad 2\Sigma n^4 B + 2\Sigma n^6 D + \ldots = \Sigma a^3 b,$$

$$2\Sigma n^4 A + 2\Sigma n^6 C + \ldots = \Sigma a^4 b, \quad 2\Sigma n^6 B + 2\Sigma n^8 D + \ldots = \Sigma a^5 b,$$

. . . .

In addition to that fact that the sums of powers of the integers are given by known formulas, we also gain the advantage of being able to calculate separately the coefficients of even terms and those of odd terms, which will considerably simplify the amount of work.

Even in the case where neither the values of x nor the values of y occur in an arithmetical progression, we can profit from these simplifications by proceeding as follows. Suppose that x and y are both functions of a third variable z, whose values are completely arbitrary, but are equally spaced, as with x above. We would seek by our procedure the values of $\frac{dx}{dz}, \frac{dy}{dz}, \frac{d^2x}{dz^2}, \frac{d^2y}{dz^2}, \ldots$. We would then obtain, using known formulas

$$\frac{dy}{dx} = \frac{dy/dz}{\frac{dx}{dz}}, \quad \frac{d^2y}{dx^2} = \frac{\frac{dx}{dz}\frac{d^2y}{dz^2} - \frac{dy}{dz}\frac{d^2x}{dz^2}}{\left(\frac{dx}{dz}\right)^2} \ .$$

This method seems to me preferable to that which consists of interpolation between observations in order to render them equidistant. It is understood, of course, that it may be dangerous, in a problem of a rather delicate nature, to change the values of the observations before using them.

It seems to us that the introduction of the method which we have described into the method of Mr. Laplace, for the determination of the orbit of comets, will greatly increase its precision, at least in the case where we have a large number of observations. However, this method, as is true of many other methods, will basically be nothing more than well-directed groping.

There remains another problem to be resolved, which can be stated as follows; we know that a number of points, however many, are located near a parabolic curve of unknown fixed degree, and we wish to know the most likely value of the degree of this curve. The solution to this problem would eliminate the uncertainty of the analyst who, wishing to apply the method of Mr. Laplace, is able to employ a large number of observations.