THE EXPERIMENTAL EVALUATION OF DEFINITE INTEGRALS

BY
GEORGE WILLIAM TYLER
B.S. (EMORY AND HENRY COLLEGE), M.A. (DUKE UNIVERSITY)

A THESIS PRESENTED IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY AT
VIRGINIA POLYTECHNIC INSTITUTE
1949

THE EXPERIMENTAL EVALUATION OF DEFINITE INTEGRALS

George W. Tyler

TABLE OF CONTENTS

I. INTRODUCTION

Two types of statistical problems, frequently met with in practice, are not always distinguished between sharply. These types may conveniently be described and contrasted in terms of an agricultural experiment in which, in order to keep the argument one-dimensional, we assume the experimental area is long and narrow, so that fertility variations across the area may be ignored.

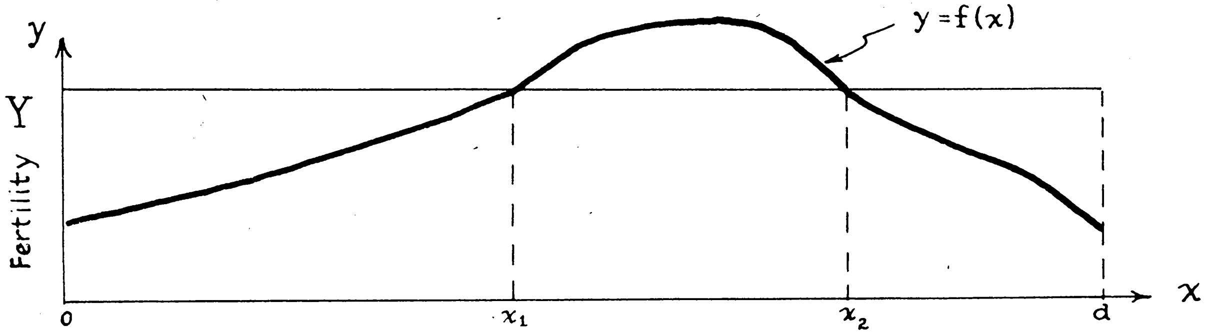
Type I.

Let us suppose that the experimental area is to be used to compare several varieties. It is feared that a part and perhaps a large part of the observed differences, among varieties, comes from differences among the fertility levels of the various parts of the experimental area. Two questions then arise:

1. What can be done to minimize the effects of fertility variations?
2. How can variations in fertility be prevented from imposing biases on the experimental comparisons?

Neither of these questions would have any substance if the fertility levels at all parts of the experimental area were known exactly. However, in practice, these levels are entirely unknown or are at best only vaguely mapped out.

The second of these questions is usually answered as follows: allot the varieties to their experimental plots by a random process. It is perhaps not immediately clear that randomness has anything to do with the question. After all, *there is no underlying frequency distribution of fertility levels.* There is simply variation in fertility level in some more or less continuous fashion, as shown in Figure 1. Thus, the theory of errors is not directly applicable to this situation. If probability methods are to become available here, *a frequency distribution must be generated by the sampling process.*



position in experimental area.

FIGURE 1.

To this end, consider the following method of transforming the fertility curve of Figure 1 into a frequency distribution.

Let y denote the fertility level, x the position in the experimental area and let $y = f(x)$ be the equation of the fertility curve. We assume, further, that $f(x)$ is a single-valued function of x, which is well behaved mathematically. Let $F(Y)$ be the measure of the x-values for which $y < Y$ (in Figure 1, $x_1 + d - x_2$). Then $\frac{F(Y)}{d}$ satisfies all the requirements for a probability distribution function and therefore defines a probability distribution of y. If, therefore, we choose a random sample of y's from this distribution (presumably by selecting randomly on x), those differences among observed yields which result from differences in fertility levels may be treated according to probability methods, provided either that the form of $F(Y)$ is reasonably well known or that methods are used which do not depend on the underlying distribution.

It is customary to assume that $\frac{F(Y)}{d}$ is the distribution function of a normally-distributed variable. This assumption may be seriously invalid, especially if the fertility curve contains substantial trends. One might, therefore, undertake to study the actual distribution and devise probability methods appropriate to it. This approach, while theoretically possible, appears to possess little practical merit, inasmuch as fertility trends, over a considerable range, may assume any form whatever and one is sure to encounter probability distributions of almost every conceivable type. A further serious disadvantage, associated with this approach, stems from the magnitude of the errors to which the experiment is exposed. Large trends in the fertility curve are reflected in a large variance of the associated probability distribution.

The alternative to this approach brings us to Question 1. Since large trends are the source of most of the trouble, let us restrict the range within which randomization is practiced sufficiently that no trends of sizable magnitude are (or are likely to be) encountered within these ranges. This leads directly to the randomized block and other stratified designs. Shrewdly applied, these designs have been used with remarkable success in controlling the error which results from randomization and in furnishing accurate estimates of variety effects. The probability distributions generated in employing these methods, have usually, it appears, not been seriously different from normality. Thus, the experimenter has, without loss of confidence in his results, been able to test for the significance of observed differences by referring to probability levels of tabulated functions. The chi-square and F distributions have been very useful in this connection.

<u>Type II</u>.

If, now, we suppose that our area contains, not a set of experimental units which are to be contrasted, but a stand (presumably, though not necessarily, all of the same variety) whose total yield we wish to estimate, we have a problem of somewhat different type. The question of bias entering into a comparison no longer exists, so that randomization need not be introduced on this account.

It is reasonable to assume that the stand can be depicted by a curve like that in Figure 1, where y represents now the rate of yield (density of stand). The whole problem here is to estimate the total yield, that is, to estimate the value of $\int_0^d ydx$, or what amounts to the same thing, the average yield, defined by $\frac{1}{d}\int_0^d ydx$. Clearly this *could* be done by converting the yield curve into a frequency distribution and, on the basis of a random sample from this distribution, estimating the mean of the distribution, which is, by definition, $\frac{1}{d}\int Y\, d\, F(Y)$. This is seen to be numerically equal to $\frac{1}{d}\int_0^d ydx$. The reasons why this procedure might not be effective have already been mentioned; the probability distribution specified by F(Y) may well be of such a kind that ordinary methods of estimating the mean are not too efficient and the dispersion of the distribution may be so great that the error to which the estimate of total yield is subject, may be unsatisfactorily large. Both of these defects would result from the presence of large trends and could conceivably arise from other causes.

In an attempt to decrease the effects of these sources of difficulty, one might resort here also to the use of stratification, and in some cases, depending upon the behavior of the function f(x), this device would furnish all the control that is needed. In other cases, however, the degree of stratification required would entail more observations than would be considered feasible. The situation here is likely to be somewhat different from that encountered in the planning of an experiment. In this latter case, one is not obliged to accept any function $f(x)$, however badly behaved it may be, since he can choose an experimental area which is known to be reasonably uniform and free from abrupt trends. On the other hand, most of the cases in which it is desired to estimate a total yield or stand, permit no choice whatever in this respect. For examples, one may list the following few illustrations:

(a) The stand of timber on a given area.

(b) The number of oysters in an oyster bed.

(c) The amount of gold or other metal or mineral under a specified area.

(d) The total volume of a mountain peak, or the amount of water in a lake.

(e) The electrical power which is directed by an antenna through a given solid angle or the amount of energy being released in a given time and space.

It is a matter of experience that, in cases such as these, the frequency distributions generated by a random sampling scheme are usually grotesque and one feels the need of some alternative procedure.

The problem is, essentially, very simple. We have a curve, unknown except for a number of experimentally determined points, and we require an estimate of the area under this curve. The possibility of using numerical integration methods thus arises naturally. Strictly speaking, there is only one source of statistical error, that which enters through errors of measurement of the ordinates of the curve. The magnitude of this error may, in some instances, be estimated from the differences among duplicate measurements. When the taking of duplicates is not feasible, it may be possible to form an estimate indirectly, perhaps by means of the variate difference method or some adaptation of it. The question needs further study.

When a numerical integration formula is applied to obtain an estimate of the area under the curve, another source of error, non-statistical in character, is introduced, owing to the failure of the fitted polynomial, on which the integration formula is based, to depict accurately the actual curve. Errors of this sort appear to be troublesome to discuss theoretically, but in practice they should cause little difficulty. The fitting and polynomial errors are defined and discussed briefly in section 3.

The choice of numerical integration formula would depend chiefly on individual circumstances. One general consideration to be kept in mind when making the choice has to do with minimizing the effects of errors of measurement on the final estimate. The weighting factors of the ordinates should be kept as nearly equal as possible. It is shown in section 6 that, under reasonable assumptions for many problems, the variance due to measurement errors is a minimum, when the weighting factors are equal.

An obvious objection to this procedure is that the errors cannot be dealt with entirely on the basis of probability theory. While this objection is valid, it should not be given undue weight. The primary objective in all these problems is to form as accurate an estimate as possible of a mean or other statistic. An accurate statement of the error to which the estimate is liable is also unquestionably desirable and if this can be given in terms of classical statistical theory, so much the better. Undoubtedly, this is accomplished when randomization is employed. On the other hand, there are many situations in which one finds it hard to justify a procedure which inflates the error to which the estimate is liable in order that it may be treated according to standard statistical methods.

It is true that if ordinates are chosen randomly, numerical integration methods may still be used. In general, however, the numerical labor would be heavy, and, as is well known, the accuracy would be lower than would be obtained by taking ordinates at carefully chosen points. This matter will be discussed more fully and specific examples given in section 12. It would seem, too, that in many problems of this sort, double sampling methods might be advantageously employed. This question merits some investigation.

When variation in more than one dimension is involved, each of the two methods discussed above may be extended to cope with it. If, for example, we envisage fertility variation in two dimensions, we may think of a function $z = f(x,y)$ which furnishes the measure of fertility at each point (x,y) within a domain of area A. Then, $\frac{F(Z)}{A}$, where $F(Z)$ is the measure of the totality of (x,y) points for which $z < Z$, defines a one-dimensional probability distribution which plays the same role as $\frac{F(Y)}{d}$ in the earlier discussion.

This thesis deals with methods for handling problems of Type II. Numerical integration formulae enable one to estimate the value of a definite integral of a function by measuring or otherwise evaluating the function at certain points. It is intuitively evident that increased accuracy should generally accompany an increased number of measurements, though certainly in as far as the polynomial error is concerned this is not necessarily true and the formulae are developed to help show where to measure and how to average these measurements for a "best" estimate. If the measurements must be made subject to restrictions which preclude the application of otherwise suitable formulae, one must expect a larger polynomial error. In such cases, perhaps the methods which have been used to locate the measurement points can be employed to develop special formulae which will increase the efficiency of the estimating procedure.

Certain of the formulae for evaluating a single integral were developed soon after the discovery of the calculus itself, Newton's name being associated with one of these which is discussed in section 4. A review of the literature reveals the existence of a large variety of formulae for estimating the values of a single integral. Ingenious mathematical devices have been employed in deriving some of these formulae, notably Gauss' and Tchebichef's which will be discussed in some detail in sections 5 and 6. The development of analogous devices for solving the systems of equations encountered in the sections of this paper on multiple integrals would enable one, at least from the pure mathematicians' point of view, to treat the material of these sections more satisfactorily.

The general problem of numerical evaluation of multiple integrals apparently has received very little attention. The 1923 paper by Aitken and Frewin [14] appears to be the most systematic study of this general problem that has been made. The method employed by Aitken and Frewin is essentially the double or multiple application of single integral formulae. An attempt is made to assess the merits of certain formulae for double integrals in terms of the accuracy given for two or three special integrals of known value.

The eight point, fifth degree accuracy formula, which is derived in section 10, was given by W. Burnside [12] in 1908, and has since been known as Burnside's formula. Burnside simply wrote this formula for integrating over a square and briefly discussed its use. He gave no other formulae nor did he give the derivation of this one.

In 1940, Sadowsky [16] gave a 42 point fifth degree accuracy formula for approximating the value of a triple integral over a cube. The method of derivation and result is of special interest because the location of the points were restricted to the surface of the cube. He shows that no greater than fifth degree accuracy can be achieved under this restriction.

The classical works of Walsh [6] and Saks [8] are of fundamental mathematical importance in a study of this kind. However, they are generally concerned with more delicate mathematical considerations than are of immediate intrest to one attempting to write out explicit formulae and rules for practical estimating purposes. Throughout the following pages we confine our attention to the Riemann integral.

2. ELEMENTS OF NUMERICAL INTEGRATION.

If f(x) is a linear function, it is geometrically evident that the value of the integral

(1) $$\int_a^b f(x)\, dx = \tfrac{1}{2}[b-a][f(b) + f(a)]$$

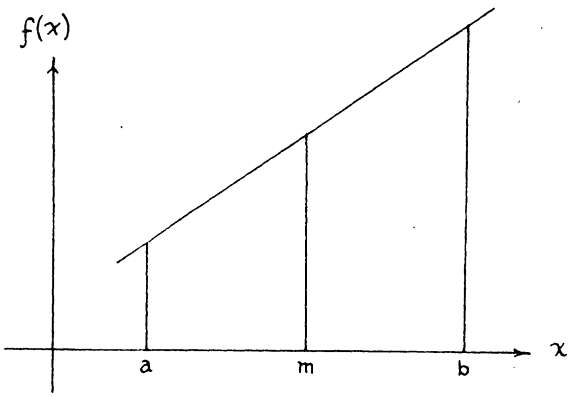


Figure 2.

This is the basis for the trapezoidal rule for approximate integration. Also it is evident from elementary geometry that if $m = \frac{b+a}{2}$ then

(2) $$\int_a^b f(x)\, dx = f(m)\, [b-a] \; .$$

This makes it clear that in this case we can calculate the value of the integral by determining the value of the function at a single properly chosen point. If the value of the function is to be found experimentally, that is by measurement, it may be more convenient to apply (2) rather than (1) to estimate the value of the integral. The polynomial error would be no greater when applying (2) but we would expect the statistical error to be larger.

A translation and scale change, given by $x' = \frac{2x - b - a}{b - a}$, reduces

$$\int_a^b f(x)\, dx \qquad \text{to} \qquad \int_{-1}^{1} F(x')\, dx',$$

where $f(x)$ and $F(x')$ are of the same degree.

Certain advantages of this transformation are immediately evident and this matter will be discussed in more detail a little later. Let

(3) $\qquad F(x) = A + Bx + Cx^2$, $\qquad$ Then by direct integration

(4) $$I = \int_{-1}^{+1} F(x)\, dx = 2A + \frac{2C}{3}$$

The basis of the parabolic (Simpson's) rule is the fact that the area under a parabolic segment equals the length of the base multiplied by one-sixth the sum of the heights of the segment at the two extremities, and four times the height at the center of the base. This fact is readily verified in this form, since from (3)

$$2\left[\frac{F(-1) + 4F(0) + F(1)}{6}\right] = 2A + \frac{2C}{3}.$$

It is also evident that we can express the value of I in terms of only two properly chosen ordinates, and again this may be more convenient experimentally.

From (3), $$2\left[\frac{F(-x) + F(x)}{2}\right] = 2A + 2CX^2$$

and if this is to represent the value of I in (4), we see that

$$2A + 2CX^2 \equiv 2A + \frac{2C}{3},$$

and this identity is satisfied if $X = \pm \frac{1}{\sqrt{3}}$. Therefore

(5) $$I = F(-\frac{1}{\sqrt{3}}) + F(\frac{1}{\sqrt{3}}).$$

Considering F(x) as a cubic, the third degree term cancels in the integration, and we see that the value of

(6) $$\int_{-1}^{+1} (A + Bx + Cx^2 + Dx^3)dx$$

is expressible in the same terms as in the quadratic case. The effect of any odd degree term neutralizes itself and it follows that a general expression for $\int_{-\lambda}^{\lambda} F(x)dx$, where F(x) is of degree 2n, also holds for F(x) of degree 2n + 1.

If F(x) is a fourth degree polynomial, we can express

$$\int_a^b F(x)dx$$

as (b-a) times the weighted average of three ordinates at properly chosen abscissa values.

Again let:

$$(7) \qquad I = \int_{-1}^{1} F(x)dx = \int_{-1}^{1}(A+Bx+Cx^2+Dx^3+Ex^4)dx = 2\left[A+\frac{C}{3}+\frac{E}{5}\right]$$

Intuitively one might suspect that the ordinates should be $F(o)$, $F(x)$, and $F(-x)$ and further that $F(x)$ and $F(-x)$ should have equal weights. This is a good guess for

$$2\left[\frac{F(-x) + KF(O) + F(x)}{K+2}\right] = 2\left[A + \frac{C}{3} + \frac{E}{5}\right],$$

provided, $\quad (K+2)\ A+2Cx^2 + 2Ex^4 \equiv (K+2)\left[A + \frac{C}{3} + \frac{E}{5}\right]$.

This identity in A, C and E is satisfied if

$$X^2 = \frac{K+2}{6} \text{ and } X^4 = \frac{K+2}{10} .$$

The useful solution for these equations is; $K = 8/5$ and $X = \frac{\sqrt{15}}{5}$

It follows that

$$(8) \qquad I = 2\left[\frac{5F\left(\frac{\sqrt{15}}{5}\right) + 8F(o) + 5F\left(\frac{\sqrt{15}}{5}\right)}{18}\right]$$

The foregoing development constitutes an approach to the problem of numerical integration which has the advantage of using only relatively widely understood mathematical principles, and is perhaps the simplest approach that can be made. However, this approach rapidly develops into a trial and error method, in that no provision has been made for assigning weights to the ordinates, in terms of which one may wish to express the value of the integral. Further, this approach does not lend itself well to developing methods for the experimental evaluation of multiple integrals.

Before proceeding to attack this problem from a slightly different point of view, a discussion of possible benefits from a change in the placement of the coordinate axes may be helpful. Given a curve, which is the graph of a polynomial of a given degree, we can by placing the origin at the center of the integration interval, eliminate from consideration in the actual integration, all odd degree terms. This is evident from the work of the preceding paragraphs, and from a strictly mathematical point of view is of distinct advantage. This advantage exists for any function, a portion of which is an odd function , since in the integration, we could disregard that portion. The engineer or research scientist who may have occasion to use some of the results or methods of this study, usually will be concerned with approximating an integral for an experimentally determined function for which he will have no exact analytical expression. Under these circumstances, the net gain from symmetrization of the integral is not as great as may at first appear. To obtain a clearer understanding of this, consider a polynomial, $f(x)$, of degree $2n$. Let $f(x) = g(x) + h(x)$, where $g(x)$ is the even function and $h(x)$ is the odd function. Given the coefficients of $g(x)$, the value of $\int_{-\lambda}^{\lambda} f(x)dx$ is independent of $h(x)$. However, the determination of the coefficients of $g(x)$, which frequently will involve a large portion of the work, is not independent of $h(x)$. In particular, if all terms of $h(x)$ were disregarded, the coefficients of $g(x)$ could be determined by solving a set of $n + 1$ linear equations, but in general $f(x)$ would then contain no more than $n + 1$ points of the experimental curve. If, on the other hand, all terms of $h(x)$ were used in the fitting, $f(x)$ could then be made to contain $2n + 1$ points of the experimental curve. The coefficients of $h(x)$ would then be determined from a set of $2n + 1$ linear equations, and would be expected to be different from those determined from the smaller set of equations. This latter procedure would generally give a better fit. Unfortunately, we could not, without a loss of precision, ignore the coefficients of $h(x)$ from the beginning of the fitting procedure, even though they need not explicitly appear in the integration. Employing methods which are usually used in solving sets of linear equations, it would be nearly as difficult to obtain the $n + 1$ needed coefficients from the $2n + 1$ equations as to obtain the whole set of $2n + 1$ coefficients. In view of this situation, and from the point of view of the applied mathematician, one is tempted to speak of the pseudo benefits from a symmetrical integral. One can argue that it is more realistic to use a non-symmetrical form for the integral, and in the next section we will consider the integration interval from zero to one. In general there would be at least a small net gain in taking the integration limits as equal but opposite in sign, and we shall presently return to the symmetrical form. The use of this form leaves open the possibility of developing special methods for solving a set of simultaneous equations for only certain of the variables, and this might substantially increase the efficiency of the approximating methods which are under consideration here.

It is to be expected that in many engineering and scientific research problems where numerical integration methods may be employed, it will be unnecessary to make actual curve fitting calculations. After a suitable formula has been selected, values of the function to be integrated may be determined experimentally for values of the independent variable, or variables, which are prescribed by the formula. This would appear to be an appropriate procedure when it is possible to determine the function values by the use of a mechanical, electrical, or electronic measuring instrument. In other problems, the formal fitting of polynomials to experimental data may lend valuable assistance to the engineer or scientist in deciding upon the most suitable formula for the problem at hand.

If the function determinations are subject to large error, then the polynomial error will be of secondary importance and higher degree accuracy should be sacrificed in order to keep the weights of the function values more nearly equal. In problems where an analytic expression is available which defies exact integration but which can be evaluated as accurately as one wishes at specified points, the higher degree accuracy formulae will usually offer the best means of increasing the precision of an estimate. This condition will exist frequently when one is attempting to make an estimate from theoretical considerations alone.

3. THE ERRORS IN APPROXIMATING THE INTEGRAL OF A FUNCTION OF A SINGLE VARIABLE.

Let us now attack the problem of finding the approximate value of

$$\int_a^b F(x)dx$$

where $F(x)$ represents an experimentally obtained function for which no exact analytical expression is known. By a proper choice of origin and scale, we can reduce this integral to the form $\int_0^1 f(x)dx$

A polynomial in x of the form, $A_0 + A_1 x + A_2 x^2 + - - - + A_n x^n = \sum_{i=0}^{n} A_i x^i$,

can be determined so as to pass through n + 1 points of, or satisfy a total of n + 1 linear restrictions which one might wish to impose with respect to, the experimental curve. We assume that f(x) and its derivatives possess the necessary continuity properties so that if e represents the failure of the polynomial to faithfully describe f(x); that is if,

(9) $$f(x) = \sum_{i=0}^{n} A_i x^i + e,$$

then the fitting error e can be made as small as we wish by taking n sufficiently large.

Let

(10) $$I = \int_0^1 (A_0 + A_1 x + A_2 x^2 + \cdots + A_n x^n)\,dx.$$

By direct integration we have

(11) $$I = A + \frac{A_1}{2} + \frac{A_2}{3} + \cdots + \frac{A_n}{n+1} = \sum_{i=0}^{n} \frac{A_i}{i+1}$$

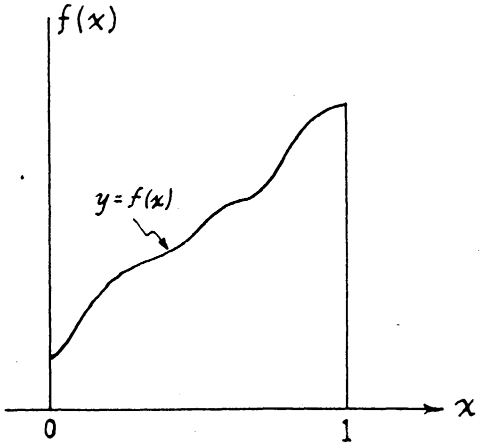


Figure 3.

The value of the integral I can be approximated by,

(12) $$I_1 = R_0 y_0 + R_1 y_1 + \cdots R_m y_m; \qquad m \leqq n:$$

where the R's are constants yet to be determined and the y's are ordinates which may be experimentally obtainable, or calculated from the functional relation:

$$y_\alpha = \sum_{i=0}^{n} A_i x_\alpha^i$$ when it exists.

The polynomial error, E, of this approximation is given by

(13) $$E = I_1 - I$$

$$= \sum_{\alpha=0}^{m} R_\alpha y_\alpha - \sum_{i=0}^{n} \frac{A_i}{i+1}$$

$$= \sum_{\alpha=0}^{m} R_\alpha \sum_{i=0}^{n} A_i x_\alpha^i - \sum_{i=0}^{n} \frac{A_i}{i+1}$$

If these expressions are expanded for a few terms, it is readily seen that,

$$
\begin{aligned}
(14)\quad E = {} & A_0(R_0 + R_1 + R_2 + - - + R_m - 1) + A_1(R_0x_0 + R_1x_1 + R_2x_2 + - - - + R_mx_m - \tfrac{1}{2}) \\
& + A_2(R_0x_0^2 + R_1x_1^2 + R_2x_2^2 + - - - + R_mx_m^2 - \tfrac{1}{3}) + - - + A_n(R_0x_0^n + R_1x_1^n + R_2x_2^n + - - - + R_mx_m^n - \tfrac{1}{n+1}).
\end{aligned}
$$

E will be zero if the coefficients of the A_i in (14) are each zero. This condition is met if the following system of equations is satisfied.

$$
(15)\quad \left\{
\begin{array}{llllll}
R_0 & + R_1 & + R_2 & + - - - & + R_m & = 1 \\
R_0x_0 & + R_1x_1 & + R_2x_2 & + - - - & + R_mx_m & = \frac{1}{2} \\
R_0x_0^2 & + R_1x_1^2 & + R_2x_2^2 & + - - - & + R_mx_m^2 & = \frac{1}{3} \\
- - - & - - - & - - - & - - - & - - - & - - \\
- - - & - - - & - - - & - - - & - - - & - - \\
R_0x_0^n & + R_1x_1^n & + R_2x_2^n & + - - - & + R_mx_m^n & = \frac{1}{n+1}
\end{array}
\right.
\qquad \text{or} \qquad \sum_{\alpha=0}^{m} R_\alpha x_\alpha^i = \frac{1}{i+1} \qquad i = 0, 1, - - -, n.
$$

This system of equations possesses greater utility than may be apparent at first. For arbitrarily assigned values of the m + 1 abscissa values, (15) reduces to a system of n + 1 linear equations in the m + 1 R's. We generally would be able to determine the R's so that m + 1 of the equations would be satisfied.

4. THE NEWTON-COTES FORMULA

If in (15) values are assigned to the x_α so that they are equally spaced, and the end points are included, we are led to the Newton-Cotes formula for numerical integration. As an example of this for $m = 3$ and $n = 3$, we would choose $x_0 = 0, x_1 = \frac{1}{3}, x_2 = \frac{2}{3}, x_3 = 1$. The system (15) then becomes

$$(16)\begin{cases} R_0 + R_1 + R_2 + R_3 = 1 \\ 2R_1 + 4R_2 + 6R_3 = 3 \\ R_1 + 4R_2 + 9R_3 = 3 \\ 4R_1 + 32R_2 + 108R_3 = 27 \end{cases}$$

and the solution is

$$R_0 = R_3 = \frac{1}{8}$$

$$R_1 = R_2 = \frac{3}{8}$$

This, of course, means that neglecting errors associated with terms of the fourth and higher degree the value of

$$I = \int_0^1 f(x)\, dx = \frac{1}{8}\left[f(0) + 3\, f(\tfrac{1}{3}) + 3\, f(\tfrac{2}{3}) + f(1)\right].$$

More generally the Newton-Cotes formula can be written

$$\int_a^{a+nw} f(x)\, dx = R_0 f(a) + R_1 f(a+w) + R_2 f(a+2w) + - - - + R_n f(a+nw).$$

The values of R_α have been called the Cotes Numbers and may be found for $n \leq 10$ in reference [2], page 168 and in other of the references for $n \leq 6$. It is of interest to note that these values can be obtained from the formula,

$$R_\alpha = \frac{(-1)^{n-\alpha}\, w}{\alpha!\,(n-\alpha)!}\int_0^n t(t-1) - - - (t-\alpha+1)(t-\alpha-1) - - - (t-\alpha)dt ,$$

which can be developed from LaGranges formula of interpolation. [1] .

Most of the better known numerical integration formulae can be obtained from (15) but before discussing other of these, a remarkable property of the system (15) will be demonstrated. A knowledge of this property logically returns us to the symmetrical form of the integral which both simplifies and decreases the amount of mathematical work necessary to obtain these formulae.

The system (15) has the property that if it is satisfied by a set of values x_α ($\alpha = 0, 1, --, m$) it is satisfied also by the set $(1-x_\alpha)$. To prove this statement let us observe that the second equation of (15), upon replacing x_α by $(1-x_\alpha)$, becomes

$$R_0(1-x_0) + R_1(1-x_1) + --- + R_m(1-x_m) = \tfrac{1}{2},$$

or

$$\sum_{\alpha=0}^{m} R_\alpha - \sum_{\alpha=0}^{m} R_\alpha x_\alpha = \tfrac{1}{2}$$

But this is precisely equation one minus equation two of (15), and this relation is satisfied if the first two equations of (15) are satisfied. Continuing this reasoning we see that each equation of (15) with $(1 - x_\alpha)$ replacing x_α is simply a linear combination of the equations in x_α up to and including the one being considered. For the $(K+1)^{st}$ equation of (15) we have

$$\sum_{\alpha=0}^{m} R_\alpha (1-x_\alpha)^K = \frac{1}{K+1}$$

or

(17) $$\sum_{i=0}^{K} \sum_{\alpha=0}^{m} (-1)^i {}_K C_i R_\alpha x_\alpha^i = \frac{1}{K+1}, \text{ where } {}_K C_i = \frac{K!}{i!\,(K-i)!}$$

This is to say, the left hand side of equation (17) is that particular linear combination of the left hand sides of the original equations in which the j^{th} equation is multiplied by the j^{th} coefficient in the expansion of $(1-x)^K$. Now the relation (17) will be satisfied, provided the right hand side, which is $\frac{1}{K+1}$, can be expressed as the same linear combination of the right hand sides of the original equations. That is, if

$$(18) \qquad \frac{1}{K+1} = 1 - K\tfrac{1}{2} + \frac{K(K-1)}{1.2}\tfrac{1}{3} - - - (-1)^K \frac{1}{K+1}.$$

$$= \sum_{i=0}^{K} (-1)^i \, {}_K C_i \frac{1}{i+1}.$$

But (18) is true, for each side gives the value of the first K^{th} difference which can be formed from the terms of the harmonic series which make up the right hand sides of (15). $\frac{1}{K+1}$ can be established as the value of this difference by mathematical induction, whereas the right hand side of (18) can be recognized as the usual method, in the theory of finite differences, of calculating the K^{th} difference from a series of $K + 1$ terms.

The property of the system (15), that if a set of values x_α is a solution then the set $(1-x_\alpha)$ is also a solution, means that if we can determine the values of the abscissae for which the function is to be found in one half of the integration interval, then the values of the abscissae in the other half can be written immediately. To take advantage of this condition let us consider,

(19) $$I = \int_{-1}^{1} F(x)\,dx$$

where, $F(x) = A_0 + A_1 x + A_2 x^2 + --- + A_{2n} x^{2n}$.

If we define: $\varphi(x) = \frac{1}{2}[F(x) + F(-x)]$,

and $\psi(x) = \frac{1}{2}[F(x) - F(-x)]$, then

(20) $$F(x) = \phi(x) + \psi(x)\ ,\ \text{and}$$

(21) $$I = \int_{-1}^{1} \phi(x)\,dx \quad \text{since} \quad \psi(x) \quad \text{is odd.}$$

$$= 2\int_{0}^{1} \varphi(x)\,dx.$$

If now we proceed as we did in obtaining (15) and let the approximation to I be

(22) $$I_1 = R_0[F(x_0) + F(-x_0)] + R_1[F(x_1) + F(-x_1)] + --- + R_m[F(x_m) + F(-x_m)],$$

we are led to the following system of equations, where $n = 2m + 1$

(23) $$\begin{cases} R_0 + R_1 + ----- + R_m = 1 \\ R_0 x_0^2 + R_1 x_1^2 + ----- + R_m x_m^2 = \frac{1}{3} \\ ---------------------- \\ R_0 x_0^{4m+2} + R_1 x_1^{4m+2} + ----- + R_m x_m^{4m+2} = \frac{1}{4m+3} \end{cases}$$

or

$$\sum_{\alpha=0}^{m} R_\alpha x_\alpha^{2p} = \frac{1}{2p+1}$$

$p = 0, 1, ---, (2m+1)$.

This system (23) contains 2m+2 parameters: the m+1 X_a^2 are the square of the abscissa values which we need to determine the sums in (21), and the m+1 R_a are the corresponding weights by which these sums are multiplied to obtain I_1. It will be observed that (23) is (15) with alternate equations deleted. If we make an arbitrary choice of K parameters in (23), we cannot expect to solve for the remaining parameters so as to satisfy more than 2m+2-K equations of (23).

The information obtained from solving the four equations in (16) in as far as the polynomial error is concerned, can now be obtained from the first two equations in (23) for m = 1. If we divide the interval on the x axis from -1 to 1 into 3 equal parts as was done in arriving at (16), it is clear that $X_0 = 1/3$, and $X_1 = 1$, so that the first two equations of (23) reduce to;

$$(24)\quad \begin{cases} R_0 + R_1 = 1 \\ R_0 + 9R_1 = 3 \end{cases} \qquad \text{with solution} \qquad \begin{aligned} R_0 &= \tfrac{3}{4} \\ R_1 &= \tfrac{1}{4}. \end{aligned}$$

These values when put in (22) give the result arrived at earlier.

5. <u>Gauss' Formula</u>

If we determine the R_a and the X_a in (23) so that all the 2m+2 equations are satisfied, we obtain the greatest possible polynomial accuracy and this leads us to Gauss' Formula for numerical integration. The simplest case is for m = 0 for which (23) becomes

$$(25)\quad \begin{cases} R_0 = 1 \\ R_0 X_0^2 = 1/3 \end{cases} \qquad \text{with solution} \qquad \begin{matrix} R_0 = 1 \\ X_0 = \pm \frac{\sqrt{3}}{3} \end{matrix}$$

These values placed in (22) yield

$$(26)\qquad I_1 = F\left(\frac{\sqrt{3}}{3}\right) + F\left(-\frac{\sqrt{3}}{3}\right).$$

In the sense that there is no error associated with terms up to and including the third degree, (25) gives the same accuracy with two determinations of the value of the function as (16) and (24) yield with four determinations. Even for small m, (23) is troublesome to solve by ordinary algebraic methods. Moors, [13] has developed a scheme for finding the solution of a general system of equations, of which (15) and (23) are special cases.

The most convenient method of solving (23) appears to be by the use of Legendre polynomials. It can be shown [3] that the X_α of (23) are given by the positive roots of the Legendre polynomial of degree 2m + 2, $P_{2m+2}(Z)$, if the total number of ordinates in the interval ± 1 is even. If the total number of ordinates to be used in the approximation is odd, then $X_0 = 0$ and the remaining X_α are the positive roots of $P_{2m+1}(Z)$. P(Z) is readily obtained from Rodriques' formula which is

$$(27)\qquad P_n(Z) = \frac{1}{2^n\, n!}\,\frac{d^n}{dZ^n}(Z^2-1)^n;$$

or the following recurrence formula can be used.

$$(28) \qquad P_{n+1}(Z) = \frac{2n+1}{n+1} Z P_n(Z) - \frac{n}{n+1} P_{n-1}(Z).$$

As an example of the use of Legendre polynomials for solving the system (23), we have for $m = 1$:

$$P_4(Z) = 35Z^4 - 30Z^2 + 3 = 0$$

and the positive roots are,

$$Z_0 = \sqrt{\frac{15 - 2\sqrt{30}}{35}}, \quad Z_1 = \sqrt{\frac{15 + 2\sqrt{30}}{35}}.$$

If the first four equations of (23) for $m = 1$ are examined, it will be found that the above values for X_0 and X_1 with $H_0 = \frac{1}{2} + \frac{\sqrt{30}}{36}$ and $H_1 = \frac{1}{2} - \frac{\sqrt{30}}{36}$ satisfy them.

It can be shown by applying Rolle's theorem to Rodrique's formula that $P\ (Z) = 0$ has n real roots which lie between -1 and 1 and it follows that (23) always has a real and usable solution.

From the above example we see the X_a satisfying (23) are irrational and this apparently is to be expected generally. It is of interest to note that Gauss [2] wrote that these X values should always be expressed in sixteen decimals to insure no error for the first 2m terms of (19). This statement is not taken to imply that any special magic is attached to sixteen decimals, but rather to sound a warning that the determination of the abscissa values is critical in achieving accuracy. This becomes apparent also in comparing special cases of Gauss' formula with a less accurate formula. The values for

X_α to sixteen decimal places for $n \leq 10$ can be found in Table C of [11], and for a smaller number of decimals and smaller values of n, in others of the references.

6. TCHEBICHEF'S FORMULA

Errors in the experimental determination of $F(x_\alpha)$ and $\varphi(x_\alpha)$ of (19) and (20) would usually be expected even if the X_α could be determined without error. Frequently one can reasonably make the assumption that the error in the determination of $F(x_\alpha)$ is independent of X_α and has constant variance, say σ^2. Under such an assumption the variance of the estimate of the value of the integral is given by,

$$(29) \qquad V(R_\alpha) = \sum_{\alpha=0}^{n} R_\alpha^2 \sigma^2.$$

The R_α are restricted according to the first equation of (15) or (23), and in order to minimize the variance given by (29) and subject to this restriction, we can find the unrestricted minimum of,

$$(30) \qquad V(R_\alpha, \lambda) = \sum_{\alpha=0}^{n} R_\alpha^2\sigma^2 + \lambda\left(\sum_{\alpha=0}^{n} R_\alpha - 1\right)$$

where λ is the Lagrangian multiplier.

Equating to zero the partial derivatives of (30), it can be seen that a necessary condition for a minimum variance is that the $n + 1$ R_α shall be equal.

When there is a relatively large error in the determination of values of $F(X_\alpha)$, minimization of the above variance becomes an important means of increasing the precision of approximation. Tchebichef's formula was developed to take advantage of this condition and can be obtained from (15) or (23) by setting $R_\alpha = \frac{1}{m+1}$. As is the case with Gauss' formula, the resulting equations are troublesome to solve for the abscissa values. Tchebichef, [9], showed that the abscissa values are the roots of polynomials which are obtained from

$$(31) \qquad X^n e^{-\sum_{k=1}^{\infty} \frac{n}{2k(2k+1)X^{2k}}} = 0.$$

This expression can be written after expanding the exponential part, as

$$(32) \qquad X^n[1 + (-\frac{n}{2\cdot 3X^2} - \frac{n}{4\cdot 5X^4} - \ldots) + \frac{1}{2!}(-\frac{n}{2\cdot 3X^2} - \frac{n}{4\cdot 5X^4} - - -)^2 + - - -] = 0$$

Using only the part of (32) for which X has a positive exponent, for $n = 2$ we have

$$X^2 - 1/3 = 0.$$

which shows that in this case Tchebichef's formula is identical with Gauss' formula as shown by (26). For $n = 5$ we obtain from (32),

$$(33) \qquad X^5 - \frac{5}{6} X^3 + \frac{7}{72} X = 0.$$

The roots of (33) which are $0, \pm \frac{1}{6}\sqrt{15 \pm \sqrt{99}}$ give the abscissae in the interval

$\pm$ 1 where the values of the function should be determined and given equal weight to obtain 5th degree accuracy. The first three equations of (23) for $m = 1$ are,

$$(34)\quad \begin{cases} R_0 + R_1 = 1 \\ R_0 X_0^2 + R_1 X_1^2 = 1/3 \\ R_0 X_0^4 + R_1 X_1^4 = 1/5 \end{cases}$$

The roots of (23) will not satisfy (34) directly, since the total number of abscissae is odd. If we adjust the first equation of (34) to account for the five equal R's, remembering that only four of them enter into this equation, it then becomes $R_0 = R_1 = 2/5$. The roots will then be found to be identical with the four non-zero roots of (33). Tchebichef's formula appears to have been applied extensively in problems of Naval architecture.

The classical criticism of both Gauss' and Tchebichef's formula has been that their application requires a great number of extended calculations, with many chances for arithmetic errors. With present achievements and continuing improvements in high speed computing machines, much of the practical importance of this criticism has vanished. Both these formulae should receive greater acceptance in the future than they appear to have received during the years they have been known.

It is clear that formulae like Gauss' and Tchebicheff's would have little value in many cases, when the accurate determination of abscissa values are meaningless; for example, when the ordinates represent densities and must be approximated by finding the *average* density over a considerable range of x-values. The timber cruise is a case in point and there are many others.

7. THE INTEGRAL OVER A RECTANGLE OF A FUNCTION OF TWO VARIABLES.

As in the case of a function of a single variable, a translation and scale change will convert

$$\int_k^l \int_m^n f(x',y')\,dx'\,dy' \qquad \text{into} \qquad \int_{-a}^{a}\int_{-b}^{b} F(x,y)\,dx\,dy$$

where $f(x',y')$ and $F(x,y)$ are of the same degree.

We assume that the function $F(x,y)$ can be represented satisfactorily by a polynomial of sufficiently high degree and write

(35) $$F(x,y) = A_{00} + A_{10}x + A_{01}y + A_{20}x^2 + A_{11}xy + A_{02}y^2 + - - - + A_{0,2n}y^{2n}.$$

$$= \sum_{\alpha=0}^{2n}\sum_{\beta=0}^{2n} A_{\alpha\beta}x^{\alpha}y^{\beta} \qquad \text{for } \alpha + \beta \leqq 2n.$$

(36) Let

$$I = \int_{-a}^{a}\int_{-b}^{b} F(x,y)\,dx\,dy.$$

and,

(37) $$I_1 = 4ab[R_1Z_1 + R_2Z_2 + - - + R_mZ_m)$$

$$= 4ab\sum_{i=1}^{m} R_iZ_i$$

Where $Z_i = F(x_i, y_i)$ and the R's are constants to be determined. I_1 is an approximation to I and is a weighted average of the values of F(x,y) at the m points multiplied by the area of the integration space.

Replacing F(x,y) in (36) by its series form as given by (35) and integrating, we have

$$I = \int_{-a}^{a}\int_{-b}^{b} (A_{00} + A_{10}x + A_{01}y + - - - + A_{2i,2j}x^{2i}y^{2j} + - - - + A_{0,2n}y^{2n})dx\,dy$$

$$= \int_{-a}^{a}\left[A_{00}y + A_{10}xy + \frac{A_{01}y^2}{2} + - - - + \frac{A_{2i,2j}x^{2i}y^{2j+1}}{2j+1} + - - - + \frac{A_{0,2n}y^{2n+1}}{2n+1}\right]_{-b}^{b} dx$$

$$= 2b\int_{-a}^{a} (A_{00} + A_{10}x + 0 + - - - + \frac{A_{2i,2j}x^{2i}b^{2j}}{2j+1} + - - - + \frac{A_{0,2n}b^{2n}}{2n+1})dx$$

$$= 4ab\left[A_{00} + - - - - - - - - - - - - - + \frac{A_{2i,2j}a^{2i}b^{2j}}{(2i+1)(2j+1)} + - - - + \frac{A_{0,2n}b^{2n}}{2n+1}\right]$$

Now let E be the error of this approximation, so that

$E = I_1 - I.$

Replacing I by its value as obtained in (38) and expressing I_1 as shown in (37) but in terms of (x_i, y_i) we see that we can write:

$$E = 4ab[\{R_1(A_{00} + A_{10}x_1 + - - - + A_{0,2n}y^{2n}) + R_2(A_{00} + A_{10}x_2 + - - - + A_{0,2n}y_2^{2n}) + - - -$$
$$+ R_m(A_{00} + A_{10}x_m - - - + A_{0,2n}y_m^{2n})\} - (A_{00} + \frac{A_{20}a^2}{3} + \frac{A_{02}b^2}{3} + - - - + \frac{A_{0,2n}b^{2n}}{2n+1})]$$

$$= 4ab[A_{00}(R_1 + R_2 + - - - + R_m - 1) + A_{10}(R_1x_1 + R_2x_2 + - - - + R_mx_m) + A_{01}(R_1y_1 + R_2y_2 + - - + R_my_m)$$
$$+ A_{20}(R_1x_1^2 + R_2x_2^2 + - - - + R_mx_m^2 - \frac{a^2}{3}) + - - - + A_{0,2n}(R_1y_1^{2n} + R_2y_2^{2n} + - - - + R_my_m - \frac{b^{2n}}{2n+1})]$$

If in (40), the coefficients of $A_{\alpha\beta}$ are equated to zero, E will be zero and we are thus led to the following system of equations.

$$
41)\quad \left\{
\begin{array}{lllll}
R_1 & + R_2 & + - - - + R_m & = 1 \\
R_1 x_1 & + R_2 x_2 & + - - - + R_m x_m & = 0 \\
R_1 y_1 & + R_2 y_2 & + - - - + R_m y_m & = 0 \\
R_1 x_1^2 & + R_2 x_2^2 & + - - - + R_m x_m^2 & = \frac{a^2}{3} \\
R_1 x_1 y_1 & + R_2 x_2 y_2 & + - - - + R_m x_m y_m & = 0 \\
- - - - & - - - - & - - - - - - - & - - \\
- - - - & - - - - & - - - - - - - & - - \\
R_1 y_1^{2n} & + R_2 y_2^{2n} & + - - - + R_m y_m^{2n} & = \frac{b^{2n}}{2n+1}
\end{array}
\right.
\quad \text{or} \quad
\left\{
\begin{array}{l}
\sum_{\alpha=1}^{m} R_\alpha x_\alpha^i y_\alpha^j = c_{ij} \\
\text{for all } i,j \text{ for which } i+j \leq 2n \\
\text{and where} \\
c_{ij} = \frac{a^i \; b^j}{(i+1)(j+1)} \text{ for both } i \text{ and } j \text{ even.} \\
\quad = 0 \text{ for } i \text{ or } j \text{ odd.}
\end{array}
\right.
$$

For $m = 1$, the solution of the first three equations of (41) is $R_1 = 1$, $x_1 = y_1 = 0$ which asserts that if $F(x,y)$ is linear,

$$(42)\quad \int_{-a}^{a}\int_{-b}^{b} F(x,y)\,dx\,dy = 4ab\,F(0,0).$$

When the coefficients $A_{\alpha\beta}$ are known or can be calculated or estimated, it may be useful to have an expression for the polynomial error in terms of the coefficients of the lowest degree terms which contribute a non-zero error when a particular formula is applied. We shall call this the remainder error and define it as the true value of the integral when the function is written to the next higher degree than the formula gives accuracy, subtracted from the value which the formula gives. Hence, *subtracting* the remainder error from the results of applying a formula should give a more accurate estimate. We would not, in general, expect even the major portion of the error sustained by using a given formula to be shown by what we are defining as the remainder error. This will depend upon the rapidity of convergence of the series into which the function has been expanded, or which, for some other reason represents the function to be integrated. Indeed, one can construct functions in which the remainder error will be arbitrarily small and the error due to some higher degree term, arbitrarily large. These facts notwithstanding, the remainder error can reasonably be expected to be a better measure of the precision one is attaining from a given formula than any other quantity of comparable simplicity.

The formula (42) gives $I = 4ab(A_{00})$, while the true value for the integral when $F(x,y)$ is second degree is $I = 4ab\left[A_{00} + \frac{A_{20}a^2}{3} + \frac{A_{02}b^2}{3}\right]$ hence the remainder error for the formula (42) is given by,

$$E = -\frac{4ab}{3}\left(A_{20}a^2 + A_{02}b^2\right).$$

For $m = 4$, a solution of the first three equations is given by, $R_1 = R_2 = R_3 = R_4 = \frac{1}{4}$

$$\begin{array}{llll} x_1 = a & x_2 = a & x_3 = -a & x_4 = -a \\ y_1 = b & y_2 = -b & y_3 = b & y_4 = -b \end{array}$$

Again if $F(x,y)$ is linear this gives,

$$(43) \qquad \int_{-a}^{a}\int_{-b}^{b} F(x,y)\,dx\,dy = ab\left[F(a,b) + F(a,-b) + F(-a,b) + F(-a,-b)\right].$$

and $\qquad E = \frac{8ab}{3}\left(A_{20}a^2 + A_{02}b^2\right)$

Both (42) and (43) are readily verified geometrically, since the value of the integral in this case is the volume of a truncated prism.

8. THE FIVE POINT FORMULAE FOR THIRD DEGREE ACCURACY.

Considering the system (41) for $m = 5$, we discover an interesting and potentially useful formula. If the first ten equations of (41), for $m = 5$ are written, it is found that they are satisfied by the following values,

$R_1 = 1/6$	$R_2 = 1/6$	$R_3 = 1/6$	$R_4 = 1/6$	$R_5 = 1/3$
$x_1 = a$	$x_2 = 0$	$x_3 = -a$	$x_4 = 0$	$x_5 = 0$
$y_1 = 0$	$y_2 = b$	$y_3 = 0$	$y_4 = -b$	$y_5 = 0$

This result gives formula (44) which we shall call the first five point formula,

$$(44) \quad \int_{-a}^{a}\int_{-b}^{b} F(x,y)\,dx\,dy = \frac{2ab}{3}\left[F(a,0) + F(0,b) + F(-a,0) + F(0,-b) + 2F(0,0)\right],$$

and

$$E = \frac{4ab}{45}\left(6\,A_{40}a^4 - 5\,A_{22}a^2b^2 + 6\,A_{04}b^4\right)$$

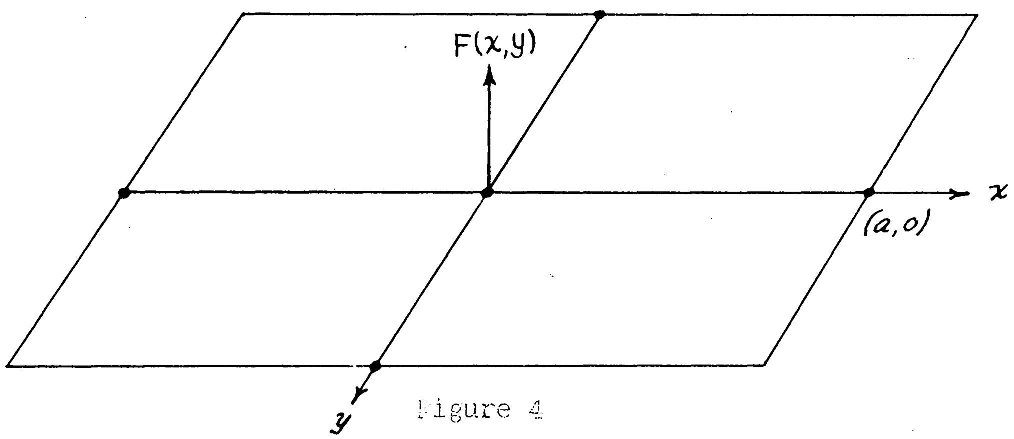


Figure 4

Formula (44) enables us to approximate the value of the integral (36) by calculating or measuring the value of the function at the five points shown in Figure 4.

The second five point formula, (45) can be developed by considering the points at the center and corners of the rectangle as shown in Figure 5.

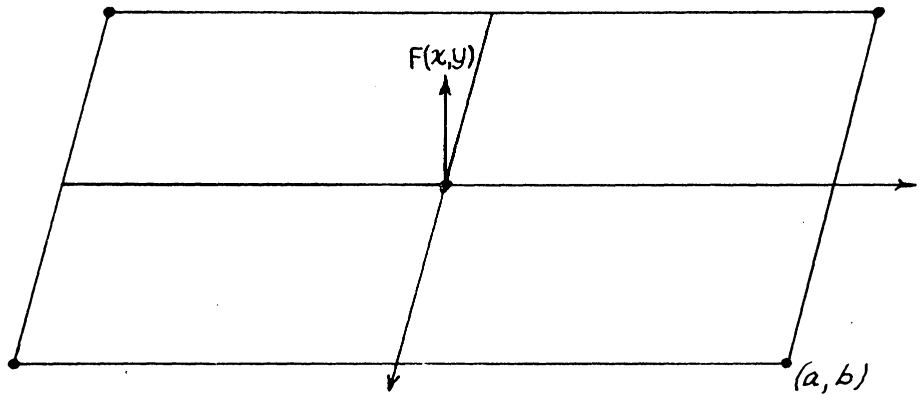


Figure 5

It will be found that the ten equations referred to in the previous paragraph are satisfied by the following values.

$R_1 = 1/12$	$R_2 = 1/12$	$R_3 = 1/12$	$R_4 = 1/12$	$R_5 = 8/12$
$x_1 = a$	$x_2 = a$	$x_3 = -a$	$x_4 = -a$	$x_5 = 0$
$y_1 = b$	$y_2 = -b$	$y_3 = -b$	$y_4 = b$	$y_5 = 0$

This gives us the following formula which has the same accuracy as (44) in the sense that only errors associated with the fourth and higher degree terms are present.

$$(45) \quad \int_{-a}^{a}\int_{-b}^{b} F(x,y)\,dx\,dy = \frac{ab}{3}\left[8F(0,0) + F(a,b) + F(a,-b) + F(-a,-b) + F(-a,b)\right]$$

The remainder error for (45) is $\frac{8ab}{45}(3A_{40}a^4 + 5A_{22}a^2b^2 + 3A_{04}b^4)$.

If we place several rectangles of equal dimensions so that some of them have common sides and apply the first five point formula to each rectangle, we can develop a very simple rule for approximating the value of the double integral of the function (which may be different for each rectangle) over the space represented by the sum of the areas of the rectangles. We can, for instance, think of Figure 6 as representing the surface of a lake, the volume of which we wish to determine.

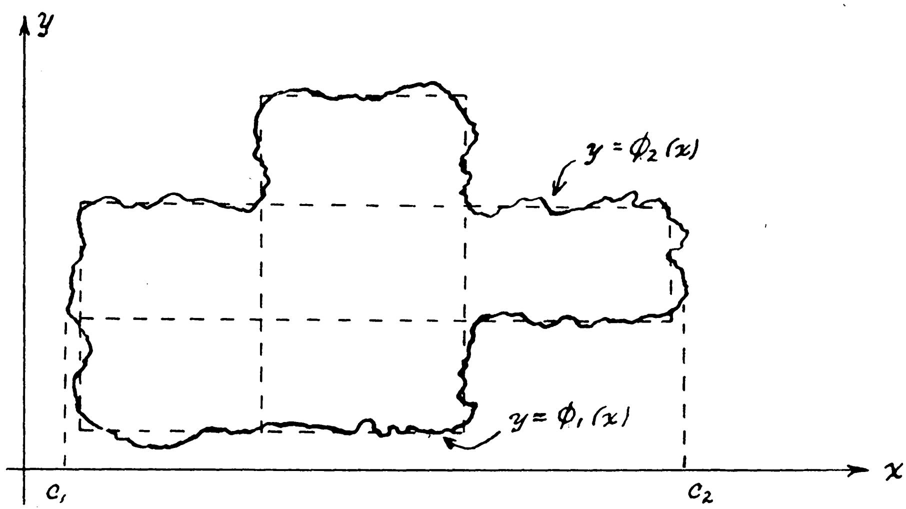


Figure 6.

If we had analytical expressions for the depth of the lake, say $\psi(x,y)$ and for the boundary as shown in Figure 6, the volume would be given precisely by,

$$(46) \qquad V = \int_{c_1}^{c_2}\int_{\phi_1(x)}^{\phi_2(x)} \psi\,(x,y)\;dx\;dy.$$

If we apply (44) to each rectangle, we observe that each point within the rectilinear figure will have a weight of two, either because it is at the center of a rectangle or because it occurs in two rectangles, while the points on the perimeter of the approximating figure will have unit weight. Thus if there are p perimeter points and q interior points, we would have V_1 as an approximation to V where,

(47) $$V_1 = \frac{\text{Total Area of Rectangles}}{p + 2q} \Big[\text{sum of depths at the p perimeter points plus twice the sum of the depths at the q interior points} \Big].$$

We can apply (45) to develop another rule similar to (47), but this will not be given in more detail. Rule (47) appears to be about the simplest such rule possible for third degree accuracy. This rule gives near equal weighting for the different determinations of the value of the function, and, as explained for Tchebichef's formula for evaluating the integral of a function of a single variable, this is of statistical advantage. This would be especially true in the lake volume problem where a portion of the function values at the perimeter points would be expected to be zero. In any application where we are striving for a high degree of accuracy, the equal weighting feature might be of considerable importance.

9. THE THIRTEEN POINT AND TWENTY-ONE POINT FORMULAE FOR APPROXIMATING DOUBLE INTEGRALS OVER A RECTANGLE.

If we seek greater accuracy than is given by (44) or (45) we find direct work with (41) somewhat tedious. For fourth and fifth degree accuracy, fifteen and twenty-one equations respectively of (41) must be satisfied.

For a fifth degree function (38) reduces to the following expression.

$$(48)\quad I = 4ab\left(A_{00} + \frac{A_{20}a^2}{3} + \frac{A_{02}b^2}{3} + \frac{A_{40}a^4}{5} + \frac{A_{22}a^2b^2}{9} + \frac{A_{04}b^4}{5}\right).$$

We can gain simplicity by writing (48) as,

$$(49)\quad I = \frac{4ab}{45}(45M + 15N + 9P + 5Q)$$

where $M = A_{00}$, $N = A_{20}a^2 + A_{02}b^2$, $P = A_{40}a^4 + A_{04}b^4$, $Q = A_{22}a^2b^2$.

We wish to find a small number of conveniently located points at which properly weighted values of $F(x,y)$ will give (49). If we choose the point at the center (0,0), the four points at the center of the sides $(0, \pm b)$ and $(\pm a, 0)$, and the four points at the corners $(\pm a, \pm b)$, we observe that $F(0,0) = M$; the sum of the values of the function at the four centers of the sides $= 4M + 2N + 2P$; and the sum of the values of the function at the four corners $= 4M + 4N + 4P + 4Q$. If these nine points are to give I as shown by (49), then the following identity in M, N, P, Q must be satisfied.

(50) $45M + 15N + 9P + 5Q = \alpha M + \beta(4M + 2N + 2P) + \gamma(4M + 4N + 4P + 4Q)$, where the Greek letters are unspecified constants representing the relative weights of the values of $F(x,y)$ at their respective groups of points as indicated above. But (50) is equivalent to (51).

$$(51)\quad \begin{cases} \alpha + 4\beta + 4\gamma = 45 \\ 2\beta + 4\gamma = 15 \\ 2\beta + 4\gamma = 9 \\ 4\gamma = 5 \end{cases}$$

We see that (51) is an inconsistent set of equations and this proves that it is impossible to obtain fifth degree accuracy in terms of these nine points as grouped here. It seems evident that no other grouping of these nine points would give a solution though this has not been proved.

If, in addition to the nine points considered in the last paragraph, we take the four points midway from the center to the midpoint of the sides we obtain a solution. The sum of $F(x,y)$ at these four points is $4M + \frac{N}{2} + \frac{P}{8}$, which, given a relative weighting of δ, would add $\delta\,(4M + N/2 + P/8)$, to the right hand side of (50). This leads us to the following set of linear equations.

$$(52)\qquad \begin{cases} \alpha + 4\beta + 4\gamma + 4\delta = 45 \\ 2\beta + 4\gamma + \delta/2 = 15 \\ 2\beta + 4\gamma + \delta/8 = 9 \\ 4\gamma = 5 \end{cases}$$

The solution of (52) is; $\alpha = -28$, $\beta = 1$

$\gamma = 5/4$, and $\delta = 16$.

This gives us (53) which we shall call the thirteen point rectangle formula.

$$(53)\qquad I_1 = \frac{4ab}{45}\left[-28\,F_1 + F_{2,5} + 5/4F_{6,9} + 16F_{10,13}\right]$$

where $F_1 = F(0,0)$

$F_{2,5} = F(0,b) + F(a,0) + F(0,-b) + F(-a,0)$

$F_{6,9} = F(a,b) + F(a,-b) + F(-a,-b) + F(-a,b)$

$F_{10,13} = F(0,b/2) + F(a/2,0) + F(0,-b/2) + F(-a/2,0)$

and $E = \frac{2ab}{21}(A_{60}a^6 + A_{06}b^6) + \frac{8ab}{45}(A_{42}a^4b^2 + A_{24}a^2b^4)$

The thirteen points are shown in, and the above notation suggested by, Figure 7.

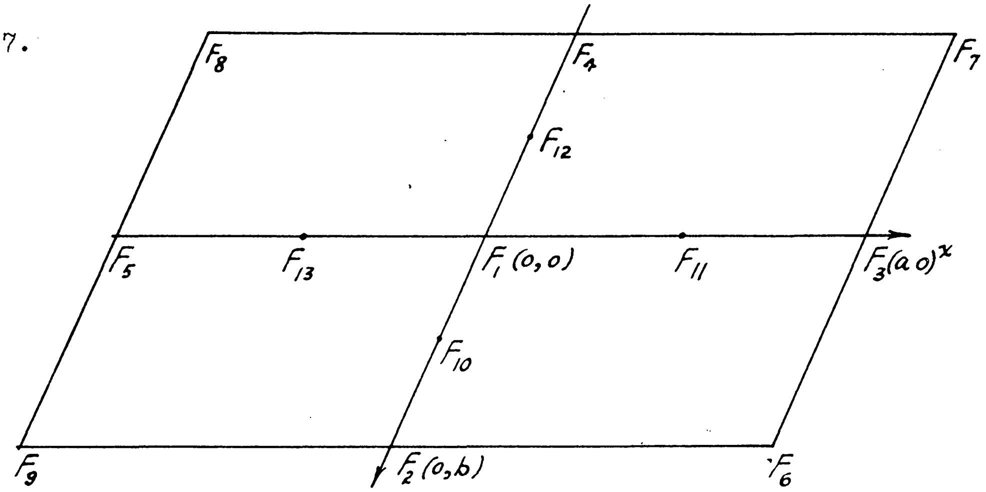


Figure 7.

The negative weighting for F(0,0) in (53) is not too surprising since certain of the Cotes numbers (weights) in the Newton-Cotes formula for a single integral are found to be negative in the cases for n = 8, and n = 10.

The solution of (52) implies that the values shown in the following table satisfy the first twenty-one equations of (41) for m = 13. This can be verified without difficulty.

α	1	2	3	4	5	6	7	8	9	10	11	12	13
$180R_\alpha$	-112	4	4	4	4	5	5	5	5	64	64	64	64
x_α	0	0	a	0	-a	a	a	-a	-a	0	a/2	0	-a/2
y_α	0	b	0	-b	0	b	-b	-b	b	b/2	0	-b/2	0

If F(x,y) is a seventh degree function (38) reduces to

$$(54)\qquad I = 4ab\left[N + \frac{N}{3} + \frac{P}{5} + \frac{Q}{9} + \frac{R}{7} + \frac{S}{15}\right]$$

where M, N, P, and Q have the same meaning as in (49), and

$$R = A_{60}a^6 + A_{06}b^6 ; \quad S = A_{42}a^4b^2 + A_{24}a^2b^4 .$$

It is evident that if we are to develop a formula by extending the method used in arriving at (53), and which will be accurate for seventh degree functions, we will most probably need a minimum of six groups of points. If we select, including the end points, five equally spaced points on each of the two diagonals and seven equally spaced points on each of the two segments joining the mid-points of opposite sides of the rectangle, and proceed as we did in arriving at the thirteen point formula, we obtain the following twenty-one point seventh degree accuracy formula.

$$(55) \quad \int_{-a}^{a}\int_{-b}^{b} F(x,y)dx\,dy = \frac{ab}{945}\Big[5388\,F_1 + 111\,\Sigma F_2 + 49\Sigma F_3 - 405\Sigma F_4 + 896\Sigma F_5 - 1863\Sigma F_6\Big]$$

where $F_1 = F(0,0)$

$$\Sigma F_2 = F(0,b) + F(a,0) + F(0,-b) + F(-a,0).$$

$$\Sigma F_3 = F(a,b) + F(a,-b) + F(-a,-b) + F(-a,b).$$

$$\Sigma F_4 = F(0,\tfrac{2b}{3}) + F(\tfrac{2a}{3},0) + F(0,\tfrac{-2b}{3}) + F(\tfrac{-2a}{3},0).$$

$$\Sigma F_5 = F(0,\tfrac{b}{2}) + F(\tfrac{a}{2},\tfrac{-b}{2}) + F(\tfrac{-a}{2},\tfrac{-b}{2}) + F(\tfrac{-a}{2},\tfrac{b}{2}).$$

$$\Sigma F_6 = F(0,\tfrac{b}{3}) + F(\tfrac{a}{3},0) + F(0,\tfrac{-b}{3}) + F(\tfrac{-a}{3},0).$$

and $E = \frac{1162}{25515}ab(A_{80}a^8 + A_{08}b^8) + \frac{2ab}{63}(A_{62}a^6b^2 + A_{26}a^2b^6) + \frac{14ab}{225}A_{44}a^4b^4$

The determination of the weights of the groups of points in (55) is the result of solving six linear equations. We note that there is a rather severe inequality of weighting between the center point and the points located one-third the distance from the center to the mid-points of the sides, and that this group of points is the only one with negative weighting.

To show the position of the points and the weights to be given to the function values at these points, we employ a presentation scheme which has been used by Bickley [17] in the closely related problem of the numerical solution of certain differential equations, and which he he has termed computation molecules. The computation molecule for (55) is shown in Figure 8.

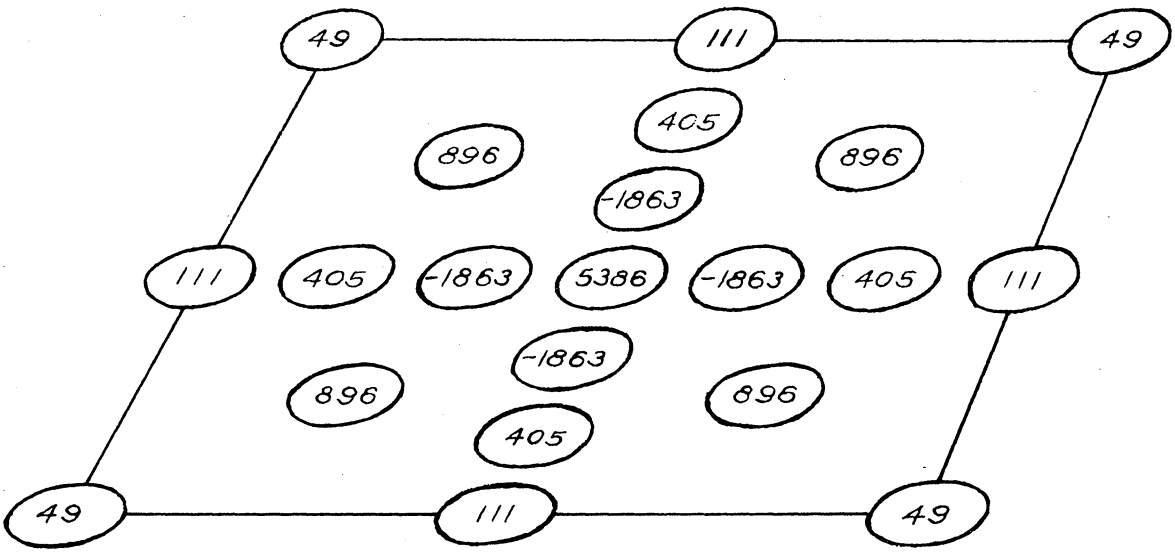


Figure 8.

The general binary polynomial of seventh degree contains 36 terms. The 21 point formula enables us to calculate without error the effects upon the double integral of the 36 terms, while in general it would be necessary to use a minimum of 36 arbitrarily located points to attain this accuracy. Formula (55) thus controls the polynomial error more efficiently than can be expected from random points but not nearly as well as the twelve point formula derived in the next section. However, an attractive feature of formula (55) is the simple location of the points.

It will be evident after reading the next section that the thirteen point and twenty one point formulae might have been obtained, after making the appropriate assumptions, in a manner analogous to the derivation of the eight point formula. It appears however, that the preceeding development, which was the way these formulae were first obtained, requires fewer judicious assumptions than would be required by the alternate method.

10. DERIVATION OF THE FOUR, EIGHT, AND TWELVE POINT FORMULAE.

In our search for greater accuracy per point than is given by the five point or thirteen point formulae, we can gain simplicity by considering the following function.

$$(56)\quad \phi(x,y) = \tfrac{1}{4}\left[F(x,y) + F(x,-y) + F(-x,y) + F(-x,-y)\right]$$

This grouping of points in sets of four which, in pairs, are symmetrically located with respect to the origin and both of the coordinate axes eliminates all terms except those in which both x and y occur in an even power. That is to say, (56) can be written as

(57) $\phi(x,y) = A_{00} + A_{20}x^2 + A_{02}y^2 + - - - + A_{2i,2j}x^{2i}y^{2j} + - - - +,$

which can be verified by writing a few terms in the expansion of each expression on the right hand side of (56). This leads to the system of equations,

(58) $$\left\{ \sum_{\alpha=1}^{m} R_\alpha x_\alpha^{2i} y_\alpha^{2j} = \frac{a^{2i}\, b^{2j}}{(2i+1)\,(2j+1)}, \right.$$ for all i, j for which $i+j \leq 2n$.

It will be observed that (58) is (41) after the equations whose right sides are zero have been deleted.

For $M = 1$ (58) becomes

(59) $$\left\{ \begin{aligned} R_1 &= 1 \\ x_1^2 &= a^2/3 \\ y_1^2 &= b^2/3 \end{aligned} \right.$$

The solution of (59) gives the following four point formula for third degree accuracy.

(60) $$I_1 = ab\left[F\left(\frac{a}{\sqrt{3}}, \frac{b}{\sqrt{3}}\right) + F\left(\frac{a}{\sqrt{3}}, \frac{-b}{\sqrt{3}}\right) + F\left(\frac{-a}{\sqrt{3}}, \frac{b}{\sqrt{3}}\right) + F\left(\frac{-a}{\sqrt{3}}, \frac{-b}{\sqrt{3}}\right)\right]$$

and $$E = \frac{16\,ab}{45}\left[A_{40}a^4 + A_{04}b^4\right]$$

This formula gives the same polynomial accuracy as the five point formulae developed earlier, and is therefore more efficient if we consider its application to a single integral. Applying this formula to a series of elemental rectangles a rule can be developed, corresponding to (47), for the approximation of an integral over an irregular area. The location of the points could not be described quite as simply as in (47), but all the function values to be determined would have equal weights. A rule developed in this way, perhaps in some applications, will be preferable to (47).

For $m = 2$ there is no solution for the first six equations of (58). This means that it is impossible to obtain fifth degree accuracy from two sets of four points each as grouped in (56).

If we consider (58) for $m = 3$ and make the following assumptions

1) $x_3 = 0$; 2) $y_2 = 0$ 3) $\frac{x_a}{a} = \frac{y_a}{b}$

we obtain the following equations.

$$(61)\quad \begin{cases} R_1 + R_2 + R_3 = 1 \\ R_1 x_1^2 + R_2 x_2^2 = a^2/3 \\ R_1 y_1^2 + R_3 y_3^2 = b^2/3 \\ R_1 x_1^4 + R_2 x_2^4 = a^4/5 \\ R_1 x_1^2 y_1^2 = a^2 b^2/9 \\ R_1 y_1^4 + R_3 y_3^4 = b^4/5 \end{cases}$$

The solution of (61) is found to be,

$$R_1 = \frac{9}{49} \qquad R_2 = R_3 = \frac{20}{49}$$

$$x_1^2 = \frac{7a^2}{9} \qquad x_2^2 = \frac{7}{15} a^2$$

$$y_1^2 = \frac{7b^2}{9} \qquad y_3^2 = \frac{7}{15} b^2$$

This gives us the eight point formula for fifth degree accuracy. The points and weights are shown in the following table.

α	1	2	3	4	5	6	7	8
196 R_α	9	9	9	9	40	40	40	40
x_α	$\frac{\sqrt{7}}{3}a$	$\frac{\sqrt{7}}{3}a$	$-\frac{\sqrt{7}}{3}a$	$-\frac{\sqrt{7}}{3}a$	$\sqrt{\frac{7}{15}}a$	$-\sqrt{\frac{7}{15}}a$	0	0
y_α	$\frac{\sqrt{7}}{3}b$	$-\frac{\sqrt{7}}{3}b$	$\frac{\sqrt{7}}{3}b$	$-\frac{\sqrt{7}}{3}b$	0	0	$\sqrt{\frac{7}{15}}b$	$-\sqrt{\frac{7}{15}}b$

Points 5 and 6 may be interpreted as four points which have become coincident in pairs, and hence the function at each of these points would have double the weight indicated by R_2 in the solution of (61). This interpretation also holds for points 7 and 8. The eight point formula, which is essentially Burnsides formula, can be written as follows,

$$(62) \qquad I_1 = \frac{ab}{49}\left[9\sum_{1}^{4} F\left(\pm\frac{\sqrt{7}}{3}a, \pm\frac{\sqrt{7}}{3}b\right) + 40\sum_{1}^{2} F\left(\pm\sqrt{\frac{7}{15}}\,a,\ 0\right) + 40\sum_{1}^{2} F\left(0, \pm\sqrt{\frac{7}{15}}\,b\right)\right]$$

where it is understood the summations extend over all distinct combinations of signs. For (62), $E = \frac{16ab}{14175}\left[-53\,(A_{60}a^6 + A_{06}b^6) + 70\,(A_{42}a^4b^2 + A_{24}a^2b^4)\right]$

The above formula, written for integrating over a square of side 2 units was given by W. Burnside [12] in 1908. He gave no details of its derivation but stated that it was constructed by a procedure closely similar to that which gives Gauss' two point third degree accuracy and three point fifth degree accuracy formulae for single integrals. Burnside illustrated the use of his formula by approximating the value of the two integrals.

$$\text{(i)} \quad \int_0^1\int_0^1 \frac{dx\,dy}{\sqrt{3-x^2-y^2}}$$

$$\text{(ii)} \quad \int_0^1\int_0^1 \frac{dx\,dy}{\sqrt{2-x^2-y^2}}$$

The formal values of these integrals are

$$\text{(i)}\ \tfrac{1}{2}\,\pi\left(1-\frac{1}{\sqrt{3}}\right), \qquad \text{(ii)}\ \pi\left(1-\frac{1}{\sqrt{2}}\right)$$

which reduced to 4 figure decimals are

(i) 0.6639 (ii) 0.9202.

Burnside gives the values of these integrals, as calculated from the formula (62) as,

(i) 0.6641, (ii) 0.9262.

He points out that in the second integral the conditions are unfavorable for applying the approximation formula since both first partial derivatives of the radical in this integral increase without limit as the point $x = 1$, $y = 1$, is approached.

The integrals (i) and (ii) were used also by Aitken and Frewin [14] to obtain a rough numerical check on some of the formulae for double integrals which they developed.

If we seek seventh degree accuracy and consider (58) for $m = 4$, and assume: 1) $\frac{x_\alpha}{a} = \frac{y_\alpha}{b}$ 2) $x_4 = y_3 = 0$.

We obtain the following system of equations.

$$
(63) \quad \begin{cases}
R_1 + R_2 + R_3 + R_4 = 1 \\
R_1 x_1^2 + R_2 x_2^2 + R_3 x_3^2 = a^2/3 \\
R_1 y_1^2 + R_2 y_2^2 + R_4 y_4^2 = b^2/3 \\
R_1 x_1^4 + R_2 x_2^4 + R_3 x_3^4 = a^4/5 \\
R_1 y_1^4 + R_2 y_2^4 + R_4 y_4^4 = b^4/5 \\
R_1 x_1^2 y_1^2 + R_2 x_2^2 y_2^2 = \frac{a^2 b^2}{9} \\
R_1 x_1^6 + R_2 x_2^6 + R_3 x_3^6 = a^6/7 \\
R_1 y_1^6 + R_2 y_2^6 + R_4 y_4^6 = b^6/7 \\
R_1 x_1^4 y_1^2 + R_2 x_2^4 y_2^2 = \frac{a^4 b^2}{15} \\
R_1 x_1^2 y_1^4 + R_2 x_2^2 y_2^4 = \frac{a^2 b^4}{15}
\end{cases}
$$

The solution for (63) is found to be

$$\frac{x_1}{a} = \frac{y_1}{b} = \sqrt{\frac{114 - 3\sqrt{583}}{287}} = 0.380555. \quad R_1 = \frac{178981 + 2769\sqrt{583}}{472230} = 0.520593$$

$$\frac{x_2}{a} = \frac{y_2}{b} = \sqrt{\frac{144 + 3\sqrt{583}}{287}} = 0.805980. \quad R_2 = \frac{178981 - 2769\sqrt{583}}{472230} = 0.237432$$

$$\frac{x_3}{a} = \frac{y_4}{b} = \sqrt{\frac{6}{7}} = 0.925820. \quad R_3 = R_4 = \frac{49}{405} = 0.120988$$

This leads to the following twelve point seventh degree accuracy formula.

$$(64)\quad \int_{-a}^{a}\int_{-b}^{b} F(x,y)dx\,dy = ab\left[R_1 \sum_1^4 F(\pm x_1, \pm y_1) + R_2 \sum_1^4 F(\pm x_2, \pm y_2) + 2R_3 \sum_1^2 F(\pm x_3, 0) + 2R_4 \sum_1^2 F(0, \pm y_4)\right],$$

where the values of the R_α, x_α and y_α are given by the solution of (63). The remainder error is:

$$ab\left[-0.013184(A_{80}a^8 + A_{08}b^8) + 0.020441\,(A_{62}a^6b^2 + A_{26}a^2b^6) - 0.010035\,A_{44}a^4b^4\right].$$

The position of the points and relative weights are shown in Figure 9.

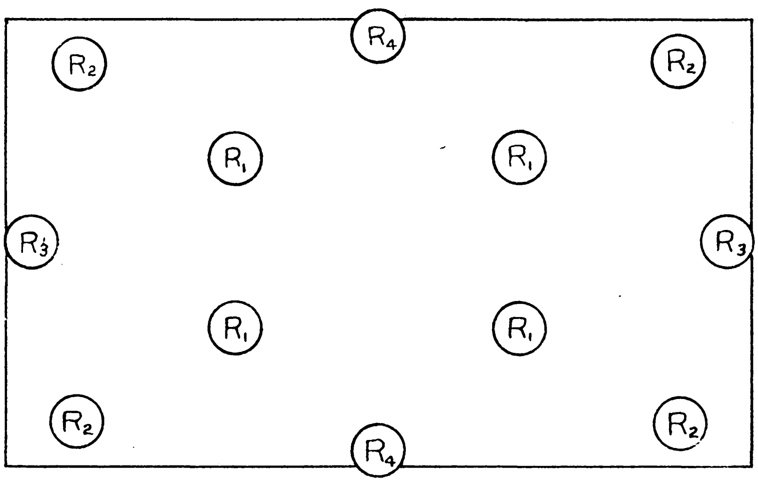


Figure 9.

For the value of the integrals which Burnside used as a rough check for his formula, the above twelve point formula gives (i) 0.6639 and (ii) 0.9161. The approximations are seen to be better than the approximations for these integrals from Burnside's formula, though the approximation for (ii) is still in error by 4 units in the third significant figure.

11. RELATIVE MERITS OF THE EIGHT AND THIRTEEN POINT FORMULAE AND THE TWELVE AND TWENTY-ONE POINT FORMULAE.

If F(x,y) is fifth degree, there are twenty-one coefficients ($A_{\alpha\beta}$). This means that there are twenty-one disposable constants, which can be used, except in special cases which we shall not discuss here, to make F(x,y) pass through twenty-one points of, or satisfy a variety of other conditions with respect to an experimentally obtained function. In statistical terms, this function has twenty-one degrees of freedom. It is evident then that the eight point or the thirteen point formulae, with their respective numbers of measurements, in so far as the integration is concerned will dispose of these twenty-one degrees of freedom without error. In the problem of estimating the value of the double integral of a function taken over a single rectangle, we can say that the eight point formula is 13/8 as efficient as the thirteen point formula. If however, we consider applying these formulae to a large number of equal sized elemental rectangles, we see that this advantage of the eight point formula is decreased, though apparently for all shapes of areas it will exist, at least to a small extent. The advantage of the eight point formula decreases, of course, because the points located on the perimeter of the elemental rectangles may be coincident for two, three, or four of these rectangles. A situation favorable to the thirteen point formula in this respect, occurs in the problem of estimating the integral over a rectangle, which, to increase the accuracy, has been sub-divided into n^2 smaller rectangles similar to the original. The eight point formula would require $8n^2$ function evaluations, compared with $8n^2 + 4n + 1$ evaluations for the thirteen point formula. It follows that for $n = 5$ an increase of about 10% in the number of function value determinations would be required to apply the thirteen point formula, but for $n > 50$, the corresponding increase would be less than 1%.

In addition to the matter discussed in the last paragraph, it is evident that application of the eight point formula would result in weights (R_α) for each point which would be more nearly equal than the weights that would result from applying the thirteen point formula. On the other hand, the location of the thirteen points could be described more simply and perhaps in some problems actually located with less error than will be the case with the eight points.

A very similar situation to that just discussed exists in regard to using the twelve or twenty-one point formulae. The twelve point formula which disposes of the effects of thirty-six coefficients ($A_{\alpha\beta}$) is highly efficient in controlling the polynomial error when applied to a single rectangle. One can readily envisage conditions under which it would seem advisable in the same problem to use a combination of different size rectangles and formulae of different degree accuracy.

No general recommendations will be attempted concerning what formulae should be used under the many possible conditions that these formulae might be applied. It is felt that such a decision should be made in the light of all available knowledge of the problem at hand.

12. USE OF ARBITRARILY LOCATED POINTS.

Under certain experimental conditions it may be impossible to evaluate $F(x,y)$ at predetermined positions. We can substitute the x and y coordinates of randomly, or arbitrarily located points into (41) and obtain linear equations in the R's. Except for the pathological cases in which the coefficients determinant of these linear equations vanishes, ν is both a necessary and sufficient number of R's to obtain a solution for ν equation. There are $\frac{1}{2}(K+1)(K+2)$ coefficients in the general two variable K^{th} degree series, and it follows that it will be necessary to evaluate $F(x,y)$ at $\frac{1}{2}(K+1)(K+2)$ arbitrarily chosen points to achieve K^{th} degree accuracy in approximating the integral (36). It is evident that in most applications, the degree accuracy employed may lose much of its meaning unless the points are well distributed over the integration area.

13. DOUBLE INTEGRALS OVER CURVILINEAR BOUNDED AREAS.

In problems requiring integration over an irregularly bounded area, one can see possibilities of obtaining a more accurate and efficient approximation by the use of formulae which involve variable limits for the integrals. This general problem of developing numerical integration formulae for double integrals with variable limits has not been studied exhaustively, but it is evident that complexities increase rapidly as we allow the bounding surface and cylinder greater freedom.

Let

$$(65) \quad I = \int_{-a}^{a} \int_{0}^{b(1 - \frac{x^2}{a^2})} F(x,y)\, dx\, dy.$$

Geometrically, I represents the volume under the surface F (x,y) and bounded by the parabolic cylinder, $y = b(1 - \frac{x^2}{a^2})$, and the x y and x z planes. If we select the five points shown in Figure 10 and proceed in a manner analogous to the procedure for developing the thirteen point rectangle formula of Section 9, we find that we can achieve second degree accuracy for $F(x,y)$ in terms of these points. The resulting formula is shown as (66).

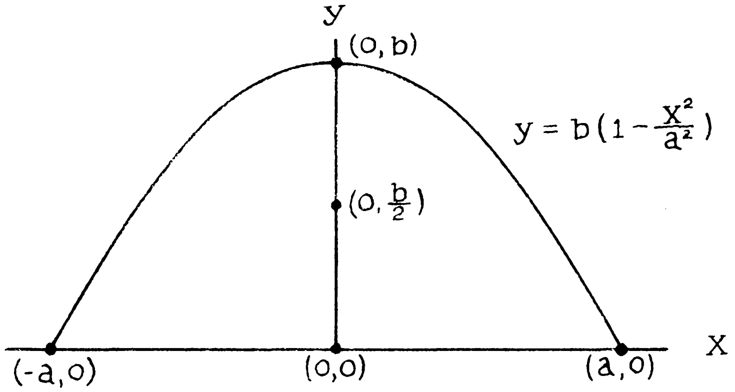


FIGURE 10.

$$(66) \quad \int_{-a}^{a} \int_{0}^{b(1 - \frac{x^2}{a^2})} F(x,y)\, dx\, dy = \frac{4\,ab}{3}\left[\frac{2}{35}\{F(0,0) + F(0,b)\} + \frac{1}{10}\{F(-a,o) + F(a,o)\} + \frac{24}{35} F(o, b/2)\right]$$

$$= \frac{\text{Area of parobolic segment}}{70}\left[4\,F(o,o) + 4\,F(o,b) + 7F(-a,o) + 7F(a,o) + 48F(o,\tfrac{b}{2})\right]$$

$$E = -\frac{4\,ab}{315}\,(6A_{21}\,a^2 b + 17A_{03}b^3)$$

Seeking greater freedom for F(x,y), we can gain simplicity by considering the doubly symmetrical integral,

$$(67) \qquad \int_{-a}^{a}\int_{-b(1-x^2/a^2)}^{b(1-x^2/a^2)} F(x,y)\,dx\,dy.$$

Making use of the symmetrical location of the points we can group the thirteen points shown in Figure 11 into six groups and derive, as the result of solving a set of six (we now see it could have been done with five) linear equations for the weights, the following thirteen point parabolic formula which has fifth degree accuracy.

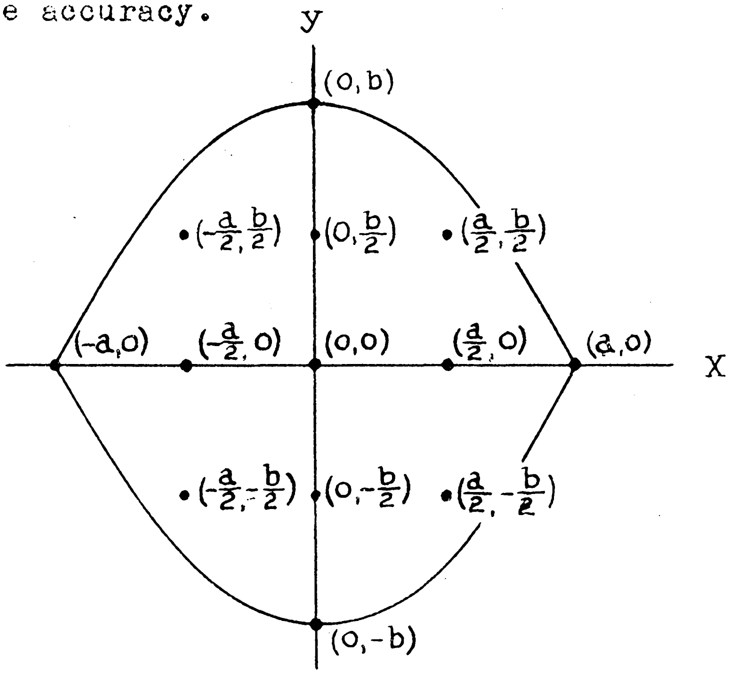


FIGURE 11.

$$\int_{-a}^{a}\int_{-b(1-\frac{x^2}{a^2})}^{b(1-\frac{x^2}{a^2})} F(x,y)\,dx\,dy = \frac{8\,ab}{3}\;\frac{1}{6930}\Big[344\,F(o,o) + 248\{F(o,b)+F(o,-b)\}$$

$$+ 768\,\{F(o,\tfrac{b}{2})+F(o,\tfrac{-b}{2})\} + 165\{F(a,o)+F(-a,o)\}$$

$$(68) \qquad + 704\{F(\tfrac{a}{2},o)+F(\tfrac{-a}{2},o)\}$$

$$+ 704\{F(\tfrac{a}{2},\tfrac{b}{2})+F(\tfrac{a}{2},\tfrac{-b}{2})+F(\tfrac{-a}{2},\tfrac{b}{2})+F(\tfrac{-a}{2},\tfrac{-b}{2})\Big]$$

$$E = \frac{-8ab}{3}\left[\frac{11\,A_{60}a^6}{70} + \frac{59}{630}\left(A_{42}a^4b^2 + A_{24}a^2b^4\right) + \frac{307}{2310}A_{06}b^6\right]$$

Attractive features of formula (68) are the simple position of the points and the near equality of weighting for all these points.

14. THE APPROXIMATION OF TRIPLE INTEGRALS OVER A RECTANGULAR PARALLELEPIPED.

The methods employed in the preceeding sections can immediately be extended to triple integrals. Changing the notation slightly, we let

(69) $$I = \int_{-a_1}^{a_1}\int_{-a_2}^{a_2}\int_{-a_3}^{a_3} F(x_1,x_2,x_3)\, dx_1\, dx_2\, dx_3$$

and assume $F(x_1,x_2,x_3)$ can be expressed in series form as follows:

(70) $$F(x_1,x_2,x_3) = A_{000} + A_{100}x_1 + - - - + A_{\alpha\beta\gamma}x_1^{\alpha}x_2^{\beta}x_3^{\gamma} + - - - + A_{00,2n}x_3^{2n}$$

$$= \sum_{\alpha=0}^{2n}\sum_{\beta=0}^{2n}\sum_{\gamma=0}^{2n} A_{\alpha\beta\gamma}x_1^{\alpha}x_2^{\beta}x_3^{\gamma} \text{ for all } \alpha,\beta,\gamma \text{ for which } \alpha+\beta+\gamma \leq 2n.$$

If we express the value of the function at the i^{th} point by

(71) $$x_{4i} = F(x_{1i}, x_{2i}, x_{3i})$$

and let the approximating function be

(72) $$I_1 = \sum_{i=1}^{m} R_i x_{4i}$$

then by setting the error, $I-I_1$, equal to zero we obtain the following system of equations.

(73) $$\sum_{a=1}^{m} R_a x_{1a}^{i}x_{2a}^{j}x_{3a}^{k} = C_{ijk} \text{ for all } i,j,k \text{ for which } i+j+k \leq 2n$$

Where
$$C_{ijk} = \frac{a_1^i\, a_2^j\, a_3^k}{(i+1)(j+1)(k+1)} \text{ for } i,\ j \text{ and } k \text{ all even}$$

$$= 0 \text{ for } i, \text{ or } j, \text{ or } k \text{ odd.}$$

Grouping the points in sets of 8, one in each octant, and proceeding in a manner analogous to the development of (58), we would obtain a simpler system than (73). One, or both of these systems can, perhaps, be employed advantageously in deriving formulae of higher degree accuracy, though they are not needed for the results of this section.

If we integrate (69) directly, and omit all fourth and higher degree terms we obtain,

$$\text{(74)} \qquad I = 2^3 a_1 a_2 a_3 \left[A_{000} + \frac{1}{3} \left(A_{200} a_1^2 + A_{020} a_2^2 + A_{002} a_3^2 \right) \right]$$

By considering the values of $F(x_1, x_2, x_3)$ at the center of each of the six faces of the parallelepiped, it will be observed that the volume of the integration space multiplied by the mean of these six values is identical with (74). Hence we have the following six point formula for third degree accuracy.

$$\text{(75)} \quad I_1 = \frac{2^3 a_1 a_2 a_3}{6} \left[F(a_1,0,0) + F(-a_1,0,0) + F(0,a_2,0) + F(0,-a_2,0) + F(0,0,a_3) + F(0,0,-a_3) \right]$$

$$E = \frac{8\, a_1 a_2 a_3}{45} \left[6(A_{400} a_1^4 + A_{040} a_2^4 + A_{004} a_3^4) - 5(A_{220} a_1^2 a_2^2 + A_{202} a_1^2 a_3^2 + A_{022} a_2^2 a_3^2) \right]$$

In a similar manner, by considering the value of the function at the corners and center of the parallelepiped, we obtain the following five point formula for near third degree accuracy.

$$\text{(76)} \quad I = \frac{2^3 a_1 a_2 a_3}{12} \left[8F(0,0,0) + F(a_1,a_2,a_3) + F(-a_1,a_2-a_3) + F(a_1,-a_2-a_3) + F(-a_1,-a_2,a_3) \right]$$

Using appropriate R's and coordinates as indicated by (75), it is found that the first twenty equations of (67) for $m = 6$

are satisfied. Using the R's and coordinates as shown by (76), we find that nineteen of the first twenty equations of (73) for $m = 5$ are satisfied. The only term less than fourth degree which contributes an error when using (76) is the $x_1\, x_2\, x_3$ term. If the coefficient, A_{111}, of this term is available, then subtracting $\frac{8}{3}\ A_{111}\ a_1^2\ a_2^2\ a_3^2$ from (76) will eliminate this error and enable us to make a full third degree precision estimate from these five points.

We can obtain a nine point, third degree accuracy formula by considering the center and all eight vertices of the parallelepiped. This formula is written as (77) and while it does not control the polynomial error as efficiently as either of the two preceding formulae, it gives better coverage of the integration space and in certain problems it might be employed advantageously.

$$(77)\quad I_1 = \frac{a_1 a_2 a_3}{3}\left[16F(o,o,o) + \sum_{1}^{8} F(\pm a_1,\ \pm a_2,\ \pm a_3)\right];\ \text{with error,}$$

$$E = \frac{16\ a_1 a_2 a_3}{45}\left[3(A_{400} a_1^4 + A_{040} a_2^4 + A_{004} a_3^4) + 5(A_{220} a_1^2 a_2^2 + A_{202} a_1^2 a_3^2 + A_{022} a_2^2 a_3^2)\right]$$

If we seek greater accuracy and consider the twenty-one points which are located at

(1) the center of the parallelepiped,

(2) the eight vertices,

(3) the six centers of the faces, and

(4) the six mid-points of the segments joining the center of the parallelepiped to the center of each face,

we obtain formula (78) which has fifth degree accuracy. The details of the derivation will be omitted but proceeding as we did in developing the thirteen point formula of section 9, we can derive (78) as the result of solving only four linear equations.

$$(78) \qquad I_1 = \frac{a_1\, a_2\, a_3}{45} \left[-496F_1 + 5\, F_8 + 8\, F_{6,1} + 128F_{6,1/2}\right],$$

where: $F_1 = F(o,o,o)$

F_8 = Sum of values of the function at the 8 vertices.

$F_{6,1}$ = Sum of values of the function at the 6 centers of the faces.

$F_{6,1/2}$ = Sum of values of the function at the 6 points located mid-way from the center of the parallelepiped to the six faces.

A poor feature of this formula is the large negative weight of F_1. This can be improved somewhat by choosing the last six points nearer the center, but the negative weighting cannot be eliminated in this way and it is doubtful if a more useful formula will result from such an adjustment. The general ternary quintic has 56 terms each of which might contribute an error in estimating the value of the triple integral and thus formula (78), which utilizes only twenty-one points, has high efficiency for controlling the polynomial error.

It is clear that rules can be developed, based on any one of the last four formulae, for estimating the triple integral of a function over a domain which has been sub-divided into elemental parallelepipeds. In view of the equal weighting for the points and the general simplicity of (75), it appears that such a rule based on this formula would possess the greatest practical merits.

Sadowsky [16] developed the following 42 point formula:

$$\int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1}\mu(x,y,z) = \frac{4}{225}\left[91\sum\mu_6 - 40\sum\mu_{12} + 16\sum\mu_{24}\right]$$

where: $\Sigma\mu_6$ denotes the sum of the six values of μ (x,y,z) determined at the centers of the six faces of the cube, $\Sigma\mu_{12}$ denotes the sum of the values of $\mu(x,y,z)$ at mid points of the twelve edges of the cube, and $\Sigma\mu_{24}$ denotes the sum of the twenty-four values of $\mu(x,y,z)$ at the four points on the diagonals of each face and at a distance of $\frac{1}{2}\sqrt{5}$ from the center of the face. This formula has fifth degree accuracy and the points are all located on the surface of the cube. Sadowsky concludes that 42 points is the smallest number that can be used to achieve this accuracy under the restraint that the points must lie on the surface. He also points out that the sixth degree function $\mu(x,y,z) = (x^2 - 1)(y^2 - 1)(z^2 - 1)$ vanishes at all points on the surface of the cube and hence it is impossible in general to attain as high as sixth degree accuracy under the above restraint.

15. GENERALIZATION FOR FIRST DEGREE AND THIRD DEGREE ACCURACY.

The possibilities of writing formulae with a given degree accuracy for any number of variables have not been explored extensively, but it is evident that some of the formulae of the preceding sections are special cases of more general formulae that can be written.

Let us consider,

$$I = \int_{-a_1}^{a_1} - - - \int_{-a_n}^{a_n} F(X_1, - - -, X_n)\, dx_1 - - - dx_n \tag{79}$$

where $F(X_1, - - -, X_n)$ can be expressed in series form by

$$F(x_1, - - -, x_n) = \sum_{\alpha_1=0}^{N} - - - \sum_{\alpha_n=0}^{N} A_{\alpha_1 - - - \alpha_n} X_1^{\alpha_1} - - - X_n^{\alpha_n} \tag{80}$$

for all α_i for which $\sum_{i=1}^{n} \alpha_i \leq N$

If $F(X_1, - - -, X_n)$ is linear, all the coefficients, except the first (the constant), will be neutralized in the successive integrations and we have immediately the following formula for first degree accuracy.

$$I = 2^n \prod_{i=1}^{n} a_i\, F(0,0, - - -,0) \tag{81}$$

In terms of n dimensional geometry, the almost trivial result (81) simply asserts that the integral of any linear function taken over a rectangular domain is the product of the volume of the integration space and the value of the function at the center of this domain.

If we assume $F(X_1, - - -, X_n)$ is a third degree polynomial then by direct intergration of (79) we obtain

$$(82) \qquad I = 2^n \prod_{i=1}^{n} a_i \left[A_{00--0} + \frac{1}{3} \left(A_{20--0} a_1^2 + A_{02--0} a_2^2 + --- + A_{00--2} a_n^2 \right) \right]$$

It is evident that the expression in brackets in (80) is a weighted average of the value of the function at the center of the integration space and at the "center of the faces" of this space. Equation (83) gives the weighting which for all positive values of n yields (82) and is therefore a 2n + 1 point formula with third degree accuracy for integrating over a rectangular n space.

$$(83) \qquad I = \frac{2^n \prod_{i=1}^{n} a_i}{6} \Big[(6-2n)F(0,0,--,0) + F(a_1,0,--,0) + F(-a_1,0,--,0) + F(0,a_2,--,0) + --- + F(0,0,---,-a_n) \Big]$$

Equations (75) and (44) are special cases of (83)

16. ESTIMATES FROM EQUALLY-SPACED DATA.

From the practical point of view there are strong arguments in certain type problems for insisting that estimates be based upon equally-spaced data. This restriction enables one who is conducting an experiment to simplify instruction to his helpers and this should tend to decrease the number of human errors and, in general, expedite the progress that can be made on a given experimental program. It is evident also that function value measurements or calculations, at equally-spaced intervals, gives one the most general coverage of the integration region and the best chance to detect discontinuities or other radical changes in the function to be integrated. In the case of the function of a single variable, the use of equally-spaced abscissa points permits us to employ orthogonal polynomial theory in a much simplified manner from that which is possible otherwise. Under the equal spacing restriction, we can make simple use of tabulated coefficients to sub-divide the variance of the ordinates into components attributable to particular terms of best fitting polynomials. This device has been very useful in studying the trends in economic data and it has been employed in numerous other fields.

In the problem of estimating the value of a single integral, it has been shown in previous sections that the Newton-Cotes formulae, which can be directly applied to ordinate measurements made at equally-spaced abscissa points, do not control the polynomial nor observational error as efficiently as do Gauss' and Tchebichef's formulae. This relative inefficiency, resulting from using equally-spaced data, not only persists, but appears to become more pronounced as the number of variables upon which the function depends increases.

In the case of double integrals, if we consider the four points at the corners of a rectangle as the basis for estimating the value of the integral over a region which has been sub-divided into elemental rectangles, we can attain only first degree precision in each rectangle for the polynomial error. This would give equal weighting for the points and maximum control for the observational error. If the rectangular net can be made sufficiently fine or the polynomial error is of little relative importance, this may be the best procedure. The measurement scheme based on the first five point third-degree accuracy formula and illustrated in the lake volume problem of section 7, can be applied to equally-spaced points along the diagonals and should give better control of the polynomial error than the slightly simpler scheme based on the four point first degree accuracy formula.

In attempting to improve error control by basing our estimate upon a larger number of lattice points, we are essentially searching for more effective ways of averaging the function values for these points. Considering a nine point rectangular lattice, it is evident from equation (51) section 8, that we cannot achieve fifth degree accuracy with these nine points. Even if fifth degree precision for the whole rectangle were possible from these nine points, it appears that, in many applications, we would prefer third degree accuracy for each of the four sub-rectangles.

Proceeding to the sixteen point lattice, it is somewhat disturbing to find the situation is virtually the same as with the nine points.

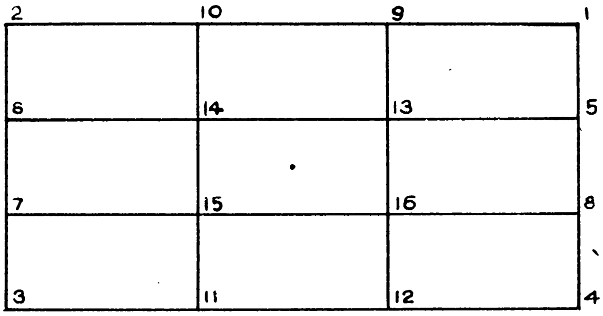


FIGURE 12

Considering the sixteen points shown in Figure 12, and grouping them as

(1) $\sum_{i=1}^{4} i$ with weight α,

(2) $\sum_{i=5}^{12} i$ with weight β,

(3) $\sum_{i=13}^{16} i$ with weight γ,

and employing the notation and method of section 9, we are led to the following system of equation which must be satisfied for fifth degree accuracy.

$$(84)\qquad \begin{cases} 4\alpha + 8\beta + 4\gamma = 1 \\ 4\alpha + \frac{40}{9}\beta + \frac{4}{9}\gamma = 1/3 \\ 4\alpha + \frac{328}{81}\beta + \frac{4}{81}\gamma = 1/5 \\ 4\alpha + \frac{8}{9}\beta + \frac{4}{81}\gamma = 1/9 \end{cases}$$

The first three equations of (84) are inconsistent and hence we cannot attain fifth degree accuracy nor eliminate simultaneously what may be construed as the three most important sources of error, using this grouping of points. If we break the second group of points into two groups of four each with different weights and in such a manner that the grouping still produces simplification, it is readily shown that the associated system of equations reduces to (84). The first, second and fourth equation of (84) have a solution for α, β, and γ which are in the ratios 1 : 3 : 9 and a rule for approximating the integral over a network of elemental rectangles based on these weights would give, except for the perimeter points, three sets of weights having the ratios 4 : 6 : 9. Such a rule would give near fifth degree accuracy for each rectangle and would control the observational error reasonably well, but even so, it would appear to be rarely more appropriate to use than the simple procedure of (47) section 8.

It seems evident that the above grouping of the 16 points gives maximum simplification and that we could not gain full fifth degree accuracy employing these points even if we allow separated weights for individual points. This statement is a conjective, however, and no simpler proof seems available than that which would involve examining the rank of a 17 x 21 matrix. This has not been done.

We can attain fifth degree accuracy using the 25 points in a 5 x 5 rectangular lattice; but this is no better than the accuracy expected from twenty-one random points, and further, a simple weighting procedure does not seem possible even though two of the six weights for symmetrically located groups of points can be chosen arbitrarily. If we try to attain seventh degree accuracy from either 25 or 36 lattice points, we find that the equations for determining the weights for natural groupings of the points constitute inconsistent sets. In the case of 36 points, this was somewhat surprising since 36 random points would, except in special cases, yield this accuracy. The points were grouped in three groups of four points and three groups of eight points each. Three of the groupings are identical with those described for the 16 points of Figure 12 and the twenty perimeter points are similarly grouped into two groups of eight and one group of four points. The 36 x 37 Matrix has not been examined to prove the following statement but again it seems a safe conjecture that the 36 equations, which would result from providing individual weights for these points, also would constitute an inconsistent set.

The possibility of forming rules for the approximation of triple integrals over an irregularly bounded domain, and based on the formulae of section 14 has been discussed briefly in that section.

It was pointed out that the six point third degree accuracy formula for estimating triple integrals (75) appeared to be the most feasible basis for such a rule. This estimating scheme would be applied to a network of points equally-spaced in each of the three directions though not necessarily the same spacing for any two of these directions.

If we attempt to achieve greater polynomial accuracy with equally-spaced points and consider the 27 points determined by three sets of three or the 64 points determined by three sets of four mutually perpendicular planes, we find that this cannot be accomplished in terms of the function values at these points. In both these cases, the points fall quite naturally into four symmetrically located groups, and the equations to determine the weights for these groups so that the estimate of the value of the integral will have fifth degree accuracy again constitute inconsistent sets. This condition is not too surprising in view of the similar situation found to exist in the two dimensional case.

The inconsistent sets of equation encountered in these cases of equally-spaced points stem from a relatively high degree of dependence among the polynomial function values at these points. These inconsistent equations simply assert that it is impossible to develop a simple rule to utilize equally-spaced data in two and three dimensions to effectively control the polynomial error. These discoveries are disappointing insofar as possible useful results are concerned but at the same time they place emphasis upon the efficient polynomial error control that is given by many of the formulae developed in the preceding sections.

17. ESTIMATING THE ERROR OF OBSERVATION.

The estimation of the error of observation, when no direct estimate is available, might be approached in the following manner.

Just as the analysis of variance may be thought of in terms of an orthogonal transformation, certain components of which are assigned to whatever systematic effects may be present, leaving the remaining components free from these effects and therefore dependent on error alone, so, in this case, in which the mean varies in some continuous manner, we may think of an orthogonal transformation so chosen that the whole of this variation is isolated in a subset of the components.

Let $y_1, y_2, \ldots, y_n$ be a set of observed ordinates, corresponding to abscissae $x_1, x_2, \ldots, x_n$, where $x_{i+1} - x_i$ is constant. Let

$$(85)\quad Z_i = \sum_{j=1}^{n} a_{ij} y_j, \quad i = 1, 2, \ldots, \quad \sum_{j=1}^{n} a_{ij} a_{jk} = \delta_{ik} \qquad \delta_{ik} = 1 \text{ for } i = j$$
$$= 0 \text{ for } i \neq j$$

be an orthogonal transformation, so chosen that Z_1 is expressible entirely in terms of the first differences of the y's, but not expressible in terms of the second differences, Z_2 is expressible in terms of second but not third differences, etc. Then Z_{n-1} would be expressible in terms of the single $(n-1)^{th}$ difference and Z_n, if we wish to speak of it at all, would have to be a multiple of $\bar{y}$ if the orthogonality conditions are to be satisfied.

It can be seen that the conditions prescribed for this transformation can be satisfied in an essentially unique manner. Consider

$$Z_i = a_{i1}y_1 + a_{i2}y_2 + \ldots + a_{in}y_n \quad ,$$

which is to be expressible as a linear function of the i^{th} differences of the y's, $\Delta_1^{(i)}, \Delta_2^{(i)}, \ldots, \Delta_{n-i}^{(i)}$. That is, we must have identically

$$(86)\quad a_{i1}y_1 + a_{i2}y_2 + a_{in}\, y_n \equiv A_1\Delta_1^{(i)} + A_2\Delta_2^{(i)} + \ldots + A_{n-1}\Delta_{n-i}^{(i)} ,$$

where the A's are unspecified constants. Substituting the expressions for the Δ's in terms of the y's in the right side of this identity, it becomes

$$\begin{aligned} & A_1(y_{i+1} - {}_iC_1 y_i + {}_iC_2 y_{i-1} + \ldots + (-1)^i\,{}_iC_i y_1) \\ + \; & A_2(y_{i+2} - {}_iC_1 y_{i+1} + {}_iC_2 y_i + \ldots + (-1)^i\,{}_iC_i y_2) \\ \vdots \; & \\ + \; & A_{n-i}(y_n - {}_nC_1 y_{n-1} + {}_iC_2 y_{n-2} + \ldots + (-1)^i\,{}_iC_i y_{n-i}). \end{aligned}$$

Regrouping this expression in terms of the y's, the coefficient of y_k is found to be

$$A_{k-i} - {}_iC_1 A_{k-i+1} + {}_iC_2 A_{k-i+2} + \ldots + (-1)^i\,{}_iC_i A_k ,$$

Equating coefficients of corresponding y's in this identity leads to the relations

(87)

$$\begin{aligned} a_{i1} &= (-1)^i\,{}_iC_i A_1 \\ a_{i2} &= (-1)^{i-1}\,{}_iC_{i-1} A_1 + (-1)^i\,{}_iC_i A_2 \\ a_{i3} &= (-1)^{i-2}\,{}_iC_{i-2} A_1 + (-1)^{i-1}\,{}_iC_{i-1} A_2 + (-1)^i\,{}_iC_i A_3 \\ &\vdots \\ a_{i,i+1} &= {}_iC_o A_1 - {}_iC_1 A_2 + \ldots + (-1)^{i-2}\,{}_iC_{i-2} A_{i-1} + (1)^{i-1}\,{}_iC_{i-1} A_i + (-1)^i\,{}_iC_i A_{i+1} \\ &\vdots \\ a_{i,n-i} &= {}_iC_o A_{n-2i} - {}_iC_1 A_{n-2i+1} + \ldots + (-1)^{i-1}\,{}_iC_{i-1} A_{n-i-1} + (-1)^i\,{}_iC_i A_{n-i} \\ &\vdots \\ a_{i,n} &= {}_iC_o A_{n-1} \end{aligned}$$

In these equations, A_λ is to be replaced by zero whenever λ is less than 1 or greater than n-i.

In order that the n quantities a_{i1}, a_{i2} ... a_{in} may be expressible in terms of n-i parameters $A_1, A_2, \ldots, A_{n-i}$, it is clearly necessary that the a's satisfy i independent linear restraints. One of these is $\sum_{k=1}^{n} a_{ik} = 0$, since, in this sum, the coefficient of each of the A's is

$$\sum_{\lambda=0}^{i} (-1)^{i-\lambda} {}_iC_{i-\lambda} = (-1)^i \sum_{\lambda=0}^{i} (-1)^\lambda {}_iC_\lambda = (-1)^i (1-1)^i = 0.$$

The remaining i-1 restraints may be derived by considering sums of the form

$$(88) \quad \sum_{k=1}^{n} {}_{r+k-1}C_r a_{ik} \equiv S_r,$$

It can be shown that $S_r = 0$, $r = 0, 1, 2, \ldots, (i-1)$.

The coefficient of any A, A_q say, in S_r may be written in the form

$$\sum_{\lambda=0}^{i} (-1)^\lambda {}_iC_\lambda ({}_{r+i+q-1-\lambda}C_r)$$

or, equivalently,

$$(-1)^i \sum_{\mu=0}^{i} (-1)^\mu {}_iC_\mu ({}_{r+q-1+\mu}C_r)$$

Consider now the identity in <u>t,</u>

$$t^{r+q-1}(1-t)^i \equiv \sum_{\mu=0}^{i} (-1)^\mu {}_iC_\mu \, t^{r+q-1+\mu}$$

Differentiation with respect to t yields

$$(r+q-1)t^{r+q-2}(1-t)^i \quad i t^{r+q-1}(1-t)^{i-1} \equiv \sum_{\mu=0}^{i} (-1)^\mu {}_iC_\mu . (r+q-1+\mu) t^{r+q-2+\mu}$$

the right side of this identity, when $r = 1$ and $t = 1$, reduces to the coefficient of A_q in S_1, $q = 1, 2, \ldots$, whereas the left side reduces to zero. Hence, $S_1 = 0$.

Another differentiation yields an identity which, when $r = 2$ and $t = 1$, reduces on the right to twice the coefficient of A_q in S_2 and on the left to zero, proving that $S_2 = 0$. Clearly this process can be continued through i-1 differentiations, proving that S_3, $S_4, \ldots$ S_{i-1} are zero, but thereafter breaks down, since the left side no longer reduces to zero when t is given the value 1.

Thus it is seen that the coefficients a_{ik} of Z_i must satisfy the i linear relations $S_r = 0$, $r = 0, 1, 2, \ldots, i-1$. These relations express the condition that Z_i shall be expressible in terms of the i^{th} differences of the y's. To these must be added the orthogonality conditions. It may not be obvious that all these conditions can be satisfied, but the following considerations show that they can be, and in essentially only one way.

The coefficients of Z_{n-1} must satisfy the n-1 conditions $S_o = S_1 = \ldots = S_{n-2} = 0$. Thus, except for a multiplied constant, these coefficients are completely determined, since the equations are homogeneous. While these numbers may be found by solving the equations, it is evident that they must be proportional to the coefficients in the expression giving the $(n-1)^{th}$ difference of the y's in terms of the y's, that is, the coefficients of t in the expansion of $(1-t)^{n-1}$. Therefore, $a_{n-1,k} = (-1)^{k-1}\,{}_{n-1}C_{k-1}$, except, perhaps, for a multiplied constant which can be chosen for convenience.

The coefficients of Z_{n-2} must satisfy the n-2 conditions $S_0 = S_1 = \ldots = S_{n-3} = 0$ and, in addition, the condition that Z_{n-2} be orthogonal to Z_{n-1}, which is

$$\sum_{k=1}^{n} (-1)^{k-1} {}_{n-1}C_{k-1}\, a_{n-2,k} = 0.$$

Thus the n coefficients $a_{n-2,k}$ must satisfy n-1 linear homogeneous equations and therefore are uniquely determined, apart from a multiplied constant.

A continuation of this argument shows that the orthogonal transformation of the y's into the Z's is uniquely determined by the condition that Z_i, i = 1, 2,...,n-1, be expressible in terms of the i^{th} differences of the y's and also supplies a method of determining the coefficients of the transformation.

For tabulation and use, the solutions of the equations which yield the coefficents of these transformations would be extracted in integral form, with the integers reduced to their lowest terms. The linear functions of the y's, computed with these coefficients, would be converted to the corresponding Z-values by dividing by the square root of the sum of the squares of the integer coefficients.

When, in practice, all the Z's have been computed, it is to be expected that, somewhere in the sequence of Z^2 values, a point will be reached at which the systematic variation in the y's is entirely eliminated or virtually so and therefore, from this point on, the Z^2 values will reflect only errors of observation. The sum of these Z^2 values then furnishes a valid basis for an estimate of the error variance.

While this approach may possess some intrinsic interest, it should be clear that the procedure to which it leads is wholly equivalent to fitting, by orthogonal polynomials, a curve of sufficiently high degree to remove the trend from the observed series and taking the residual mean square to estimate the error variance. It follows that the coefficients of the orthogonal transformation must be, except perhaps for a multiplied constant, identical with the values of the orthogonal polynomials which are tabulated up to the fifth degree. It is a matter of experience, however, that a fifth degree curve frequently does not completely remove the trend from an empirical series. Before methods of this sort can be employed, tables are required which give the values of the orthogonal polynomials to degrees much higher than 5 up to sample sizes of 40 or 50.

Such tables could be prepared by the methods already used in constructing the tables up to degree 5, but conceivably the equations developed above for the coefficients of the orthogonal transformation contain the possibility of improved methods. Certainly they furnish a better means of checking the tables than any heretofore employed.

The equations expressing the a's in terms of the A's can be used to lighten the arithmetical work somewhat.

Note 1. The coefficients of the equations $S_0 = S_1 = \ldots\ldots = S_{i-1} = 0$ are given below

	a_{i1}	a_{i2}	a_{i3}	$\cdots\cdots\cdots$	$a_{i(n-1)}$	a_{in}
S_0:	1	1	1		1	1
S_1:	1	2	3		$n-1$	n
S_2:	1	3	6		${}_{n}C_2$	${}_{n+1}C_2$
S_{i-1}:	${}_{i-1}C_{i-1}$	${}_{i}C_{i-1}$	${}_{i+1}C_{i-1}$			${}_{n+1-2}C_{i-1}$

These numbers may be written down by a rule similar to that used in building up the Pascal triangle of binomial coefficients. Each entry is the sum of that immediately above it and that immediately to the left.

A direct application of Chio's rule reduces this set of coefficients to an equivalent and perhaps simpler set.

$$\begin{array}{ccccc}
1 & 1 & 1 & 1 & 1 \\
 & 1 & 2 & n-2 & n-1 \\
 & & 1 & {}_{n-2}C_2 & {}_{n-1}C_2 \\
 & & {}_{i-1}C_{i-1} & {}_{n-2}C_{i-1} & {}_{n-1}C_{i-1}
\end{array}$$

Note 2. The determination of the value of the integral.

If we proceed, as was advocated above, to fit a polynomial of degree sufficiently high to account for all the trend, it is possible that the area under this fitted curve is a better estimate of the value of the integral than is furnished by any of the ordinary numerical integration formulae, which depend on polynomial curves which pass through all the observed points. Indeed, it seems likely that this is the case. If it is, it may be possible to formulate some simple rules for combining the first few items in the sequence of Z's to obtain this area.

18. ORTHOGONAL POLYNOMIAL METHODS IN EVALUATING MULTIPLE INTEGRALS.

The methods of orthogonal polynomials may be useful in estimating both the observational error and the integral of functions of two or more variables. Let

$$z = f(x,y) = \sum_{\alpha=0}^{N} \sum_{\beta=0}^{N} b_{\alpha\beta} x^{\alpha} y^{\beta} \quad ; \qquad \alpha + \beta \leq N$$

be a polynomial of degree N in x and y. This equation can be rearranged and written as

$$(89) \quad z = \sum_{\alpha=0}^{N} \sum_{\beta=0}^{N} B_{\alpha\beta}\, \xi'_{\alpha}(x)\, \xi'_{\beta}(y); \qquad \alpha + \beta \leq N$$

where $\xi'_{\alpha}(x)$ and $\xi'_{\beta}(y)$ are orthogonal polynomials of degree α and β in x and y respectively. In the case of functions of a single variable the properties of orthogonal polynomials are most effective in simplifying calculations incident to the statistical analysis of a set of measurements represented as ordinates when the abscissa points are equally spaced. It is not surprising to find that the equal-spacing restriction is necessary for the effective extension of these methods to functions of several variables. We therefore assume that,

$$z_{ij}, \quad i = 1, 2, \ldots\ldots, n; \; j = 1, 2, \ldots\ldots, m.$$

represents a set of mn equally-spaced observations to which we wish to fit the function (89). To determine the B's for the best fit we would minimize the sum of squares of the deviations given by the function,

$$(90) \quad G(B_{\alpha\beta}) = \sum_{i=1}^{n} \sum_{j=1}^{m} \left[Z_{ij} - \sum_{\alpha=0}^{N} \sum_{\beta=0}^{N} B_{\alpha\beta}\, \xi'_{\alpha}(x)\, \xi'_{\beta}(y) \right]^2$$

The partial derivative of G with respect to any B, say B_{pq}, is given by;

$$\frac{\partial G}{\partial B_{pq}} = -2 \sum_{i=1}^{n} \sum_{j=1}^{m} \left[Z_{ij} - \sum_{\alpha=0}^{N} \sum_{\beta=0}^{N} B_{\alpha\beta} \xi'_{\alpha}(x_i)\, \xi'_{\beta}(y_j) \right] \xi'_{p}(x_i)\, \xi'_{q}(y_j).$$

The normal equation resulting from equating this quantity to zero can be written

$$\sum_{\alpha=0}^{N} \sum_{\beta=0}^{N} B_{\alpha\beta} \sum_{i=1}^{n} \xi'_{p}(x_i) \xi'_{\alpha}(x_i) \sum_{j=1}^{m} \xi'_{q}(y_j) \xi'_{\beta}(y_j) = \sum_{i=1}^{n} \sum_{j=1}^{m} Z_{ij}\, \xi'_{p}(x_i)\, \xi'_{q}(y_j).$$

Since: $\displaystyle\sum_{i=1}^{n} \xi'_{p}(x_i)\, \xi'_{\alpha}(x_i) = 0$ for $\alpha \neq p$,

$\displaystyle\sum_{j=1}^{m} \xi'_{q}(y_j)\, \xi'_{\beta}(y_j) = 0$ for $\beta \neq q$;

this equation reduces to,

$$B_{pq} \sum_{i=1}^{n} [\xi'_{p}(x_i)]^2 \sum_{j=1}^{m} [\xi'_{q}(y_j)]^2 = \sum_{i=1}^{n} \sum_{j=1}^{m} Z_{ij}\, \xi'_{p}(x_i)\, \xi'_{q}(y_j)$$

Therefore the regression coefficient B_{pq} is given by

$$(91.) \quad B_{pq} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{m} Z_{ij}\, \xi'_{p}(x_i)\, \xi'_{q}(y_j)}{\sum_{i=1}^{n} [\xi'_{p}(x_i)]^2 \sum_{j=1}^{m} [\xi'_{q}(y_j)]^2}$$

The orthogonal polynomials (ξ'_{α}) have been tabulated in reference [18] for $\alpha \leq 5$ and sample sizes up to 104 and these tables are being extended to considerably higher values of α in a project, under the direction of Dr. D. B. DeLury at the Ontario Research Foundation. With the orthogonal polynomials available the calculation of the regression coefficients as given by (91) for the best fitting surface to a set of observations at rectangular lattice points is a simple matter.

One of the most valuable benefits that one gains from employing orthogonal polynomials is the ease with which the reduction in the residual sum of squares attributable to a particular term of the best fitting polynomial can be calculated. This is a well known and frequently used property of this system when one is studying the variation in a quantity as the function of a single independent variable. That the reduction in residual sum of squares due to a particular term in the equation of a regression surface can be calculated very simply can be seen from the following considerations. For a regression surface, z, the residual sum of squares is given by

$$(92) \qquad \sum_{i=1}^{n} \sum_{j=1}^{m} (Z_{ij} - z_{ij})^2$$

Expressing z_{ij} in terms of x and y as given by (89) and taking expected values, we see that (92) reduces to

$$(93) \qquad \sum_{i=1}^{n} \sum_{j=1}^{m} Z_{ij}^2 - \sum_{\alpha=0}^{N} \sum_{\beta=0}^{N} B_{\alpha\beta}^2 \sum_{i=1}^{n} [\xi_\alpha'(x_i)]^2 \sum_{j=1}^{m} [\xi_\beta'(y_j)]^2.$$

A comparison of (93) and (91) will show that the reduction in sum of squares that is attributable to the term in $x^p y^q$ is simply the coefficient of this term, B_{pq}, multiplied by the numerator of this quantity as given by (91). To each $B_{\alpha\beta}$ in the regression equation will be associated a single degree of freedom and hence, if an estimate of error is available, we can use the "F" or "t" functions to test for the statistical significance of the variance found to be due to any of these coefficients. If no other estimate of observational error is available for a given set of measurements we should in many cases, simply by inspection of the reduction in residual sum of squares for the several $B_{\alpha\beta}$ calculated, be able to determine when the trends have been removed from the data. The residual sum of squares after the trends have been removed could be used for making an estimate of the observational error.

When the observational error is large and we wish to estimate the integral of the experimental function it seems even more desirable in the case of multiple integrals than in the single integral case that we make the estimate from observed values after they have been adjusted statistically. The adjustment should be made by fitting the observations to a function which is sufficiently flexible to provide for the trend or trends which may be physically meaningful.

The origin and x and y scales can be so chosen that the double integral of a function, z, of degree N, taken over any rectangle, can be written in the form

$$\int_1^n \int_1^m z \, dy \, dx = \int_1^n \int_1^m \sum_{\alpha=0}^{N} \sum_{\beta=0}^{N} B_{\alpha\beta} \, \xi'_\alpha(x) \, \xi'_\beta(y) \, dy \, dx$$

(94)

$$= \sum_{\alpha=0}^{N} \sum_{\beta=0}^{N} B_{\alpha\beta} \int_1^n \xi'_\alpha(x) \, dx \int_1^m \xi'_\beta(y) \, dy$$

The $B_{\alpha\beta}$ in (94) can be calculated as shown by (91) and $\int_1^n \xi'_\alpha(x)dx$, which can be shown to be zero for odd α, is always easy to calculate though its exact form depends upon both n and α. If the values of the above integral were calculated for the useful range of n and α and made available in tables, then the calculations in (94) would reduce to a series of simple multiplications and divisions. Moreover it appears that these calculations are of such nature that punched cards and mechanical computing methods could be applied to them, without difficulty. This would seem to greatly increase the feasibility of applying these estimation methods on an extensive scale.

The restriction, $\alpha+\beta \leq N$, was placed on (89) merely to hold this function to degree N and exclude terms which it appears we usually would not wish to consider. None of the developments of this section depend upon this restriction, and we can remove it if we wish and deal with a function each term of which will have the sum of its x and y exponents not greater than 2N.

It seems evident that the methods of this section can be extended to functions of three or more variables. Some of the details of the calculations necessary to apply the orthogonal polynomial methods to fitting a function of two variables are shown in the second illustrative example.

ILLUSTRATIVE EXAMPLE I.

The following calculations were made in connection with a radio directional antenna design problem. We wish to evaluate the integral.

$$J = \frac{240\ I^2 \sin^2 \alpha}{\pi} \iint_R f(\phi,\theta)\, d\phi\, d\theta,$$

$$= \frac{240\ I^2 \sin^2 \alpha}{\pi} \iint_R \frac{\sin^2\{\frac{kl}{2}[1-\sin\theta\cos(\phi+\alpha)]\}\sin^2\{\frac{kl}{2}[1-\sin\theta\cos(\phi-\alpha)]\}}{[1-\sin\theta\cos(\phi+\alpha)][1-\sin\theta\cos(\phi-\alpha)]}\sin\theta\, d\phi\, d\theta$$

where $\alpha = \frac{\pi}{10}$, $\frac{kl}{2} = \frac{73}{2}$ (case I), $\frac{kl}{2} = \frac{73}{4}$ (case II) and R is the region of the ϕ,θ - plane containing the principal lobe of $f(\phi,\theta)$.

The function $f(\phi,\theta)$ is symmetric about the planes $\phi = \theta$, $\theta = \frac{\pi}{2}$. The principal lobe of $f(\phi,\theta)$ is centered at $(0,\frac{\pi}{2})$, the intersection of the planes of symmetry. The locus of zeros of $f(\phi,\theta)$ is given by

$$\frac{kl}{2}[1-\sin\theta\cos(\phi\pm\alpha)] = n\pi, \quad n=0, 1, 2, ---,$$

or $\sin\theta\cos(\phi\pm\alpha) = \frac{2n\pi}{kl}$, $n = 0, 1, 2, ---.$

For the locus of zeros of the principal lobe, $n = 1$.

Case I: $\frac{kl}{2} = \frac{73}{2}$.

In Figure 1 is shown the locus of zeros of $f(\phi,\theta)$. Figure 2 and 3 show the intersections of $f(\phi,\theta)$ with the planes $\theta = \frac{\pi}{2}$ and $\phi = 0$ respectively.

From formula (68), section 13, we know that for a plane region W, bounded by the parabolas

$$y = \pm b(1-\frac{x^2}{a^2}),$$

$$\iint_{W} f(x,y)\,dx\,dy = \frac{8ab}{3}\ \frac{1}{6930}\ \{344f(0,0)+248[f(0,b)+f(0,-b)] + 768[f(0,\tfrac{b}{2})+f(0,-\tfrac{b}{2})]+165[f(a,0)+f(-a,0)] + 704[f(\tfrac{a}{2},0)+f(-\tfrac{a}{2},0)] + 704[f(\tfrac{a}{2},\tfrac{b}{2})+f(\tfrac{a}{2},-\tfrac{b}{2})+f(-\tfrac{a}{2},\tfrac{b}{2})+f(-\tfrac{a}{2},-\tfrac{b}{2})]\},$$

providing $f(x,y)$ is a polynomial of degree ≤ 5 in x and y. (See Figure 4).

In Figure 1 is shown the pair of parabolas with vertices at $(\frac{6\pi}{180},\frac{\pi}{2})$ $(-\frac{6\pi}{180},\frac{\pi}{2})$, respectively, and passing through the points $(0, \frac{74\pi}{180})$, $(0, \frac{106\pi}{180})$. These parabolas provide a very close approximation for the locus of zeros of the principal lobe. Furthermore, it seems reasonable to assume that $f(\phi,\theta)$ can be approximated by a polynomial of degree ≤ 5 in ϕ and θ, at least in the region of the principal lobe. Consequently we will use the above integration formula to evaluate $\iint_R f(\phi,\theta)d\phi d\theta$

TABLE I.

Values of $f(\phi,\theta)$ for certain values of ϕ and θ (ϕ,θ expressed in degrees for convenience).

ϕ	θ	$f(\phi,\theta)$	ϕ	θ	$f(\phi,\theta)$
0^0	90^0	380.10	4^0	90^0	78.16
0^0	88^0	363.88	5^0	90^0	17.97
0^0	86^0	316.15	6^0	90^0	0.01
0^0	84^0	241.67	3^0	82^0	42.68
0^0	82^0	153.12	0^0	$73^0 56' 16''$	0.00
0^0	80^0	72.08	1^0	$75^0 08' 52''$	0.00
0^0	78^0	20.05	2^0	$76^0 3' 3' 09''$	0.00
0^0	76^0	1.67	3^0	$78^0 13' 24''$	0.00
0^0	74^0	0.00	4^0	$80^0 18' 01''$	0.00
1^0	90^0	350.71	5^0	$83^0 08' 50''$	0.00
2^0	90^0	272.49	$5^0 56' 45''$	90^0	0.00
3^0	90^0	171.21			

Using the necessary values from table I, and making use of symmetry, we have

$$\iint_R f(\phi,\theta)\,d\phi\,d\theta = \frac{8}{3}\ \frac{16\pi}{180}\ \frac{6\pi}{180}\ \frac{1}{6930}\ [344 \times 380.10 + 2 \times 768 \times 171.21 + 2 \times 704 \times 153.12 + 4 \times 704 \times 42.68]$$

$$= 8.208$$

As a rough check on this result, assume that the surface in question is a cone of height 380.104 having the parabolic base described above. Then,

$$\iint_R f(\phi,\theta)\,d\phi\,d\theta = \frac{1}{3}\ \frac{8}{3}\ \frac{16\pi}{180}\ \frac{6\pi}{180}\ 380.10,$$

$$= 9.880,$$

which is not an unreasonable result if we consider that this cone lies almost entirely outside the surface $f(\phi,\theta)$.

Finally,

$$J = \frac{240\ I^2 \sin^2 \frac{\pi}{10}}{\pi}\ 8.208$$

$$= 59.87\ I^2$$

Case II $\frac{kl}{2} = \frac{73}{4}$

In Figure 5 is shown the locus of zeros of $f(\phi,\theta)$ and the approximating parabolas. These parabolas have vertices at $(\frac{16\pi}{180}, \frac{\pi}{2})$ and $(-\frac{16\pi}{180}, \frac{\pi}{2})$ respectively and pass through the points $(0,\frac{\pi}{3})$ and $(0, \frac{2\pi}{3})$. Figures 6, 7 and 8 show the intersections of $f(\phi,\theta)$ with the planes $\theta = \frac{\pi}{2}$, $\phi = 0$ and $\theta = \frac{5\pi}{12}$ respectively.

TABLE II.

Values of $f(\phi,\theta)$ for certain values of ϕ and θ.

ϕ	θ	$f(\phi,\theta)$	ϕ	θ	$f(\phi,\theta)$
0^{0}	90^{0}	153.02	$7^{0}30'$	$86^{0}15'$	57.12
0^{0}	85^{0}	162.05	$3^{0}45'$	$82^{0}30'$	135.00
0^{0}	$82^{0}30$	169.00	$7^{0}30'$	$82^{0}30'$	65.06
0^{0}	80^{0}	172.20	$11^{0}15$	$82^{0}30'$	14.48
0^{0}	75^{0}	143.80	15^{0}	$82^{0}30'$	0.08
0^{0}	70^{0}	62.92	$7^{0}30'$	$78^{0}45'$	65.95
0^{0}	65^{0}	5.49	$0^{0}53'08''$	$89^{0}06'52''$	152.02
2^{0}	90^{0}	143.52	$14^{0}06'52''$	$89^{0}06'52''$	0.58
4^{0}	90^{0}	116.01	$7^{0}30'$	$87^{0}37'25''$	54.83
6^{0}	90^{0}	79.81	$2^{0}22'35''$	$82^{0}30'$	154.57
$7^{0}30'$	90^{0}	53.15	$12^{0}37'25''$	$82^{0}30'$	5.67
8^{0}	90^{0}	45.11	$7^{0}30'$	$77^{0}22'35''$	62.03
10^{0}	90^{0}	19.54	$0^{0}53'08''$	$75^{0}53'08''$	151.45
12^{0}	90^{0}	5.59	$14^{0}06'52''$	$75^{0}53'08''$	0.72
14^{0}	90^{0}	0.69	$16^{0}07'15''$	90^{0}	0.00
15^{0}	90^{0}	0.11	$15^{0}47'47''$	85^{0}	0.00
2^{0}	75^{0}	134.45	$14^{0}47'36''$	80^{0}	0.00
4^{0}	75^{0}	109.11	$13^{0}00'43''$	75^{0}	0.00
6^{0}	75^{0}	74.90	$10^{0}14'18''$	70^{0}	0.00
$7^{0}30'$	75^{0}	48.87	$6^{0}00'53''$	65^{0}	0.00
10^{0}	75^{0}	14.90	0^{0}	$60^{0}30'44''$	0.00
15^{0}	75^{0}	5.27	8^{0}	75^{0}	40.85

Using the necessary values from Table II, and making use of symmetry, we have

$$\iint_R f(\phi,\theta)\ d\phi d\theta = \frac{8}{3}\ \ \frac{\pi}{6}\ \ \frac{16\pi}{180}\ \ \frac{1}{6930}\ [344 \times 153.02 + 2 \times 768 \times 45.11$$

$$+ 2 \times 704 \times 143.80 + 4 \times 704 \times 40.85]$$

$$= 24.724$$

As an approximate check on this result, for each of the four quadrants formed by the planes of symmetry, we can integrate over the square with sides of length $\frac{\pi}{12}$ and over the remaining triangular portion of the region (see Figure 5). For a rectangular region having sides 2a, 2b and centered at the origin, we have the integration formula (53) which can be written more explicitly as,

$$\int_{-a}^{a}\int_{-b}^{b} f(x,y)dy\,dx = \frac{ab}{45}\{-112f(0,0)+ 4[f(a,o)+ f(-a,o)+ f(o,b)+ f(o,-b)]$$
$$+ 64[f(-\tfrac{a}{2},o)+ f(-\tfrac{a}{2},o)+ f(o,\tfrac{b}{2})+ f(o,-\tfrac{b}{2})]$$
$$+ 5[f(a,b)+ f(a,-b)+ f(-a,b)+ f(-a,-b)]\}.$$

providing $f(x,y)$ is a polynomial of degree ≤ 5 in x and y. Using this formula, and assuming that the surface over the triangular portion of the region is approximately a pyramid, we have

$$\iint_R f(\phi,\theta)d\phi d\theta = 4\left(\frac{15\pi}{180}\right)^2 \frac{1}{180}[-112 \times 65.06 + 4(53.15 + 169.00 + 0.08 + 48.87)$$
$$+ 64(57.12 + 135.00 + 14.48 + 65.95)$$
$$+ 5(153.02 + 0.11 + 143.80 + 5.27)]$$
$$+ 4\ \frac{1}{3}\ \frac{1}{2}\ \frac{13\pi}{180}\ \frac{15\pi}{180}\ 143.80$$
$$= 19.423 + 5.694$$
$$= 25.116$$

which checks quite well.

As a further check, for the same rectangular region described above, and for a polynomial of degree ≤ 5 in x and y, we have the integration formula (62)

$$\int_{-a}^{a}\int_{-b}^{b} f(x,y)dy\,dx = \frac{ab}{49}\{9[f(\sqrt{\tfrac{7}{3}}\,a, \sqrt{\tfrac{7}{3}}\,b)+f(\sqrt{\tfrac{7}{3}}\,a,-\sqrt{\tfrac{7}{3}}\,b)+f(-\sqrt{\tfrac{7}{3}}\,a,\sqrt{\tfrac{7}{3}}b)+f(-\sqrt{\tfrac{7}{3}}a,-\sqrt{\tfrac{7}{3}}b)]$$

$$+ 40[f(\sqrt{\tfrac{7}{15}}\,a,o)+ f(-\sqrt{\tfrac{7}{15}}\,a,o)+f(o,\sqrt{\tfrac{7}{15}}\,b)+f(o,-\sqrt{\tfrac{7}{15}}\,b)]\}$$

Using this formula

$$\iint_R f(\phi,\theta)d\phi d\theta = 4(\tfrac{15\pi}{180})^2 \tfrac{1}{196}\ [9(152.02 + 0.58 + 151.45 + 0.72)$$

$$+ 40(54.83 + 154.57 + 5.67 + 62.03)]$$

$$+ 5.694$$

$$= 19.340 + 5.694$$

$$= 25.034$$

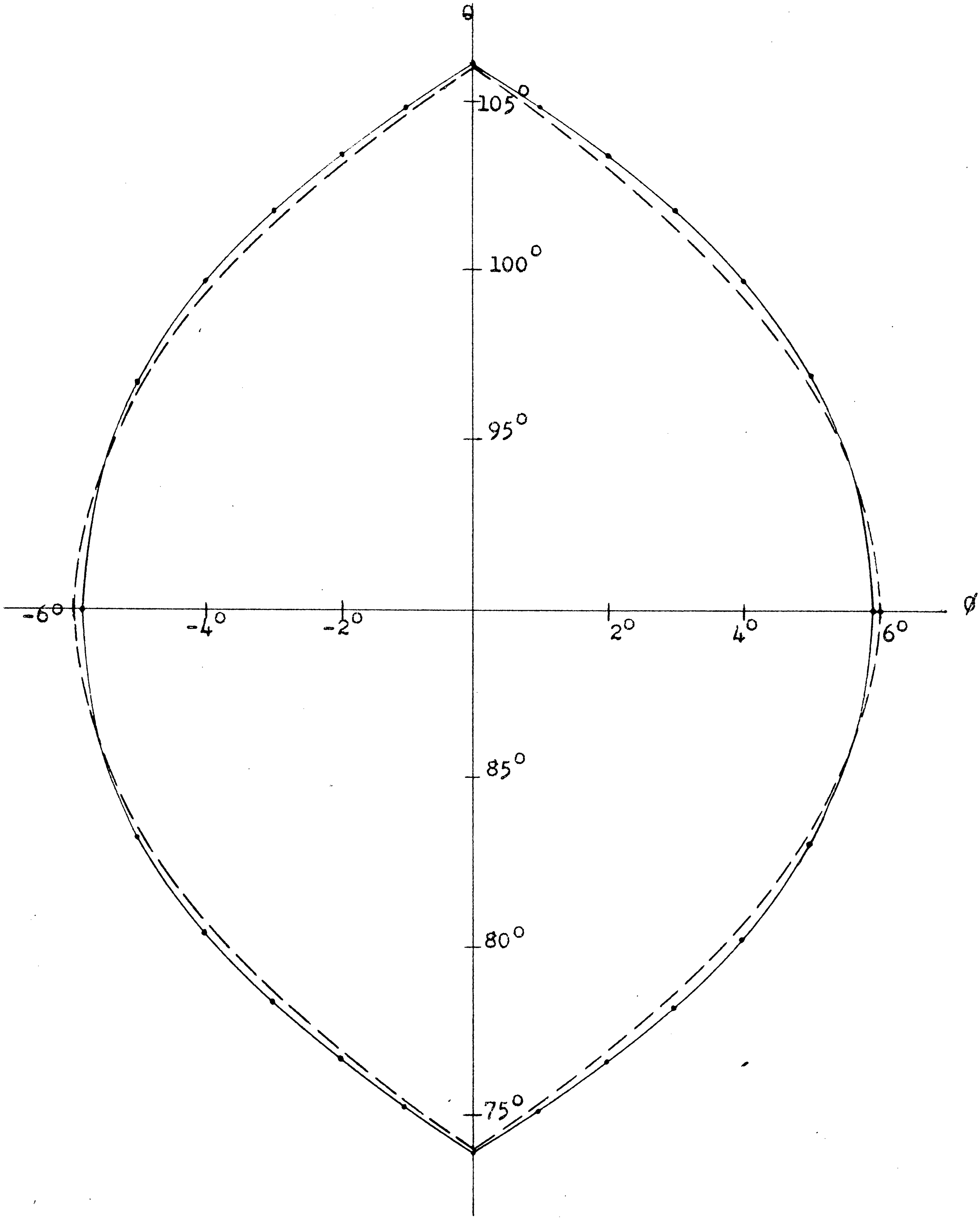
which also checks quite well.

Finally,

$$J = \frac{240\ I^2\ \sin^2\ \frac{\pi}{10}}{\pi} 24.724$$

$$= 180.36\ I^2$$

Figure 1. Locus of zeros and approximating parabolas for principal lobe of $f(\phi,\theta)$. Case I.



——— Locus of zeros.

– – – Approximating parabolas.

Figure 2. Intersection of $f(\phi,\theta)$ with $\theta = \frac{\pi}{2}$ plane. Case I.

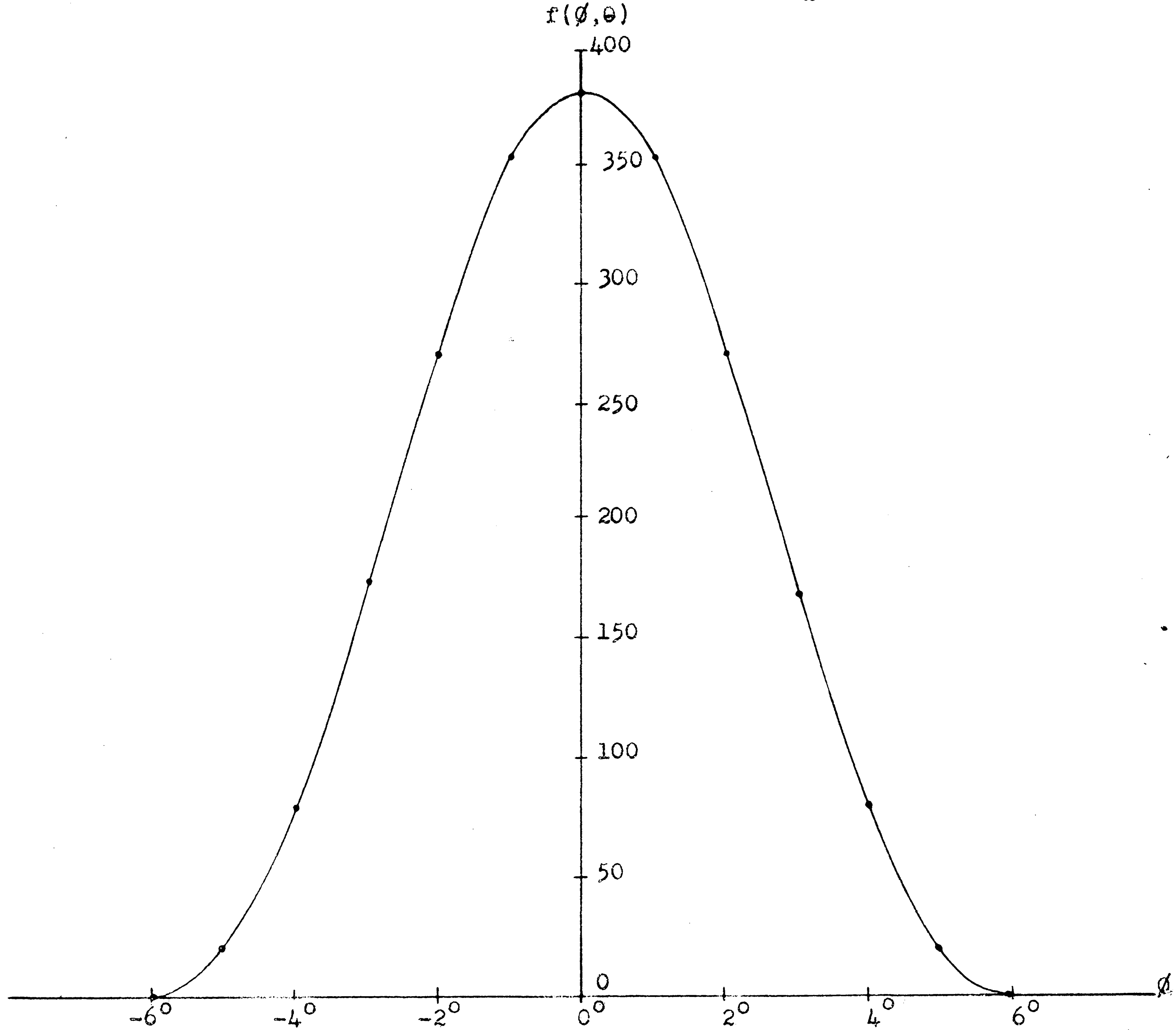


Figure 3. Intersection of $f(\emptyset,\theta)$ with $\theta = 0$ plane. Case I.

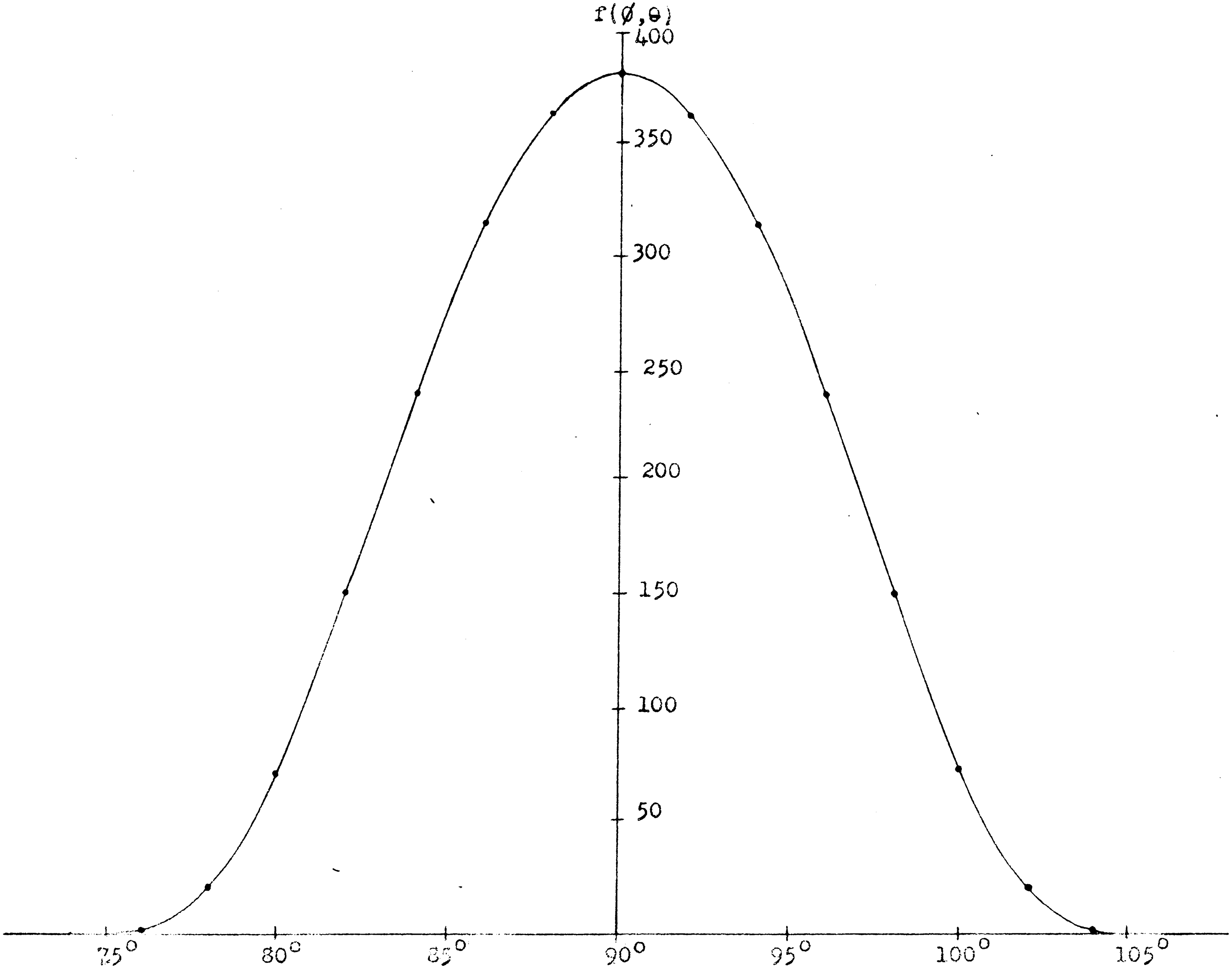


Figure 4. Parabolic region of integration showing 13 points at which function is evaluated.

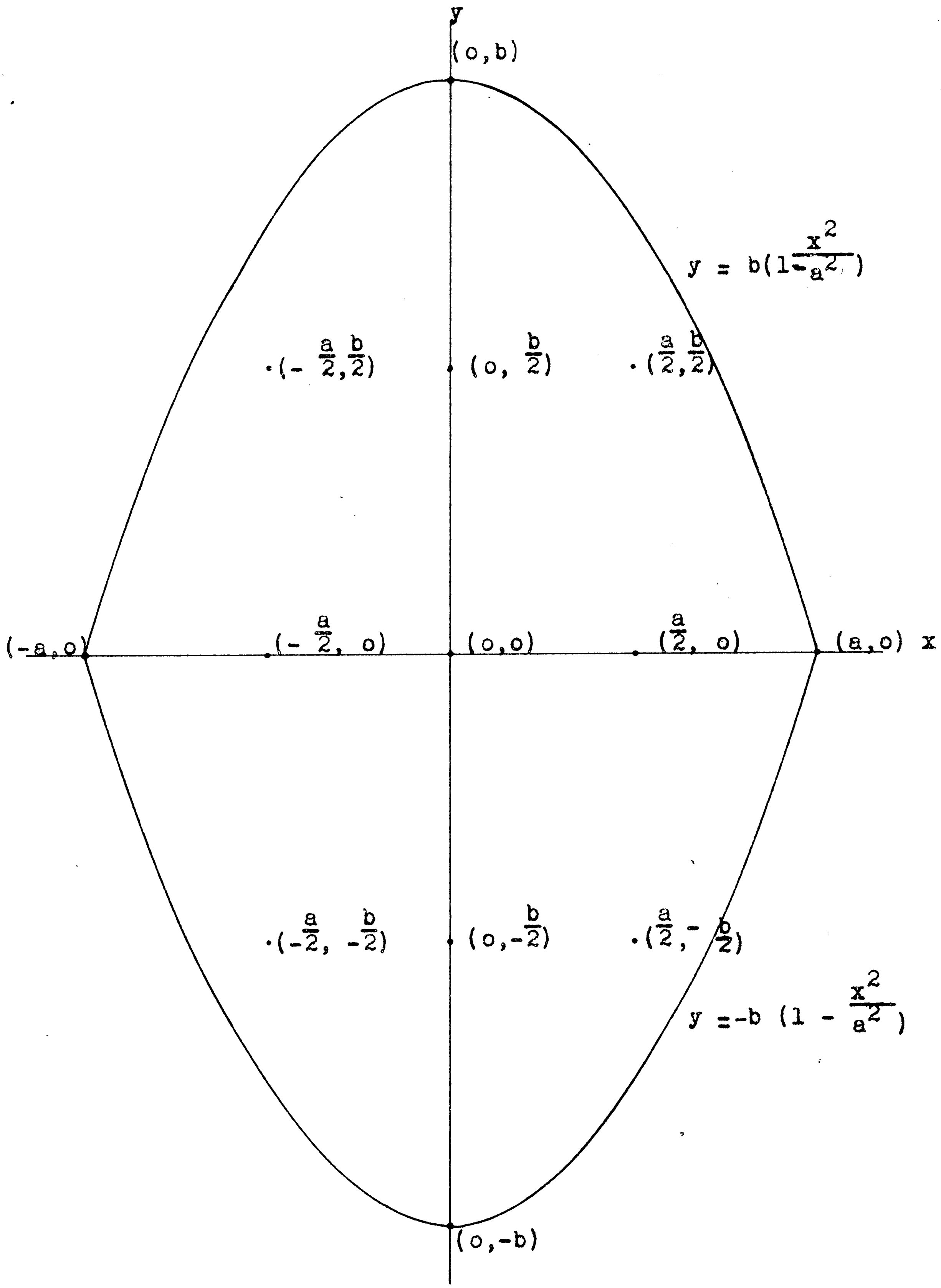
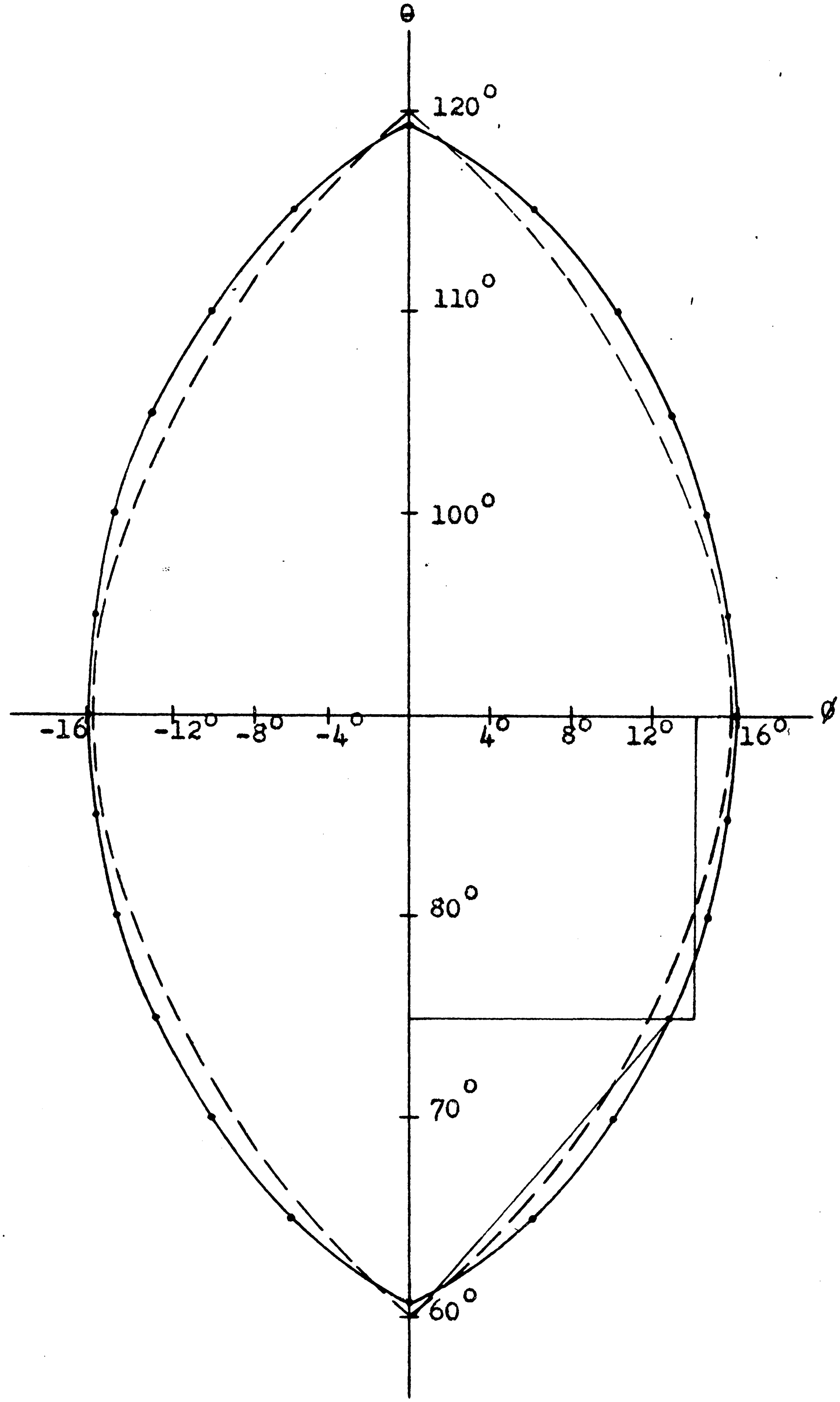


Figure 5. Locus of zeros and approximating parabolas for principal lobe of $f(\phi,\theta)$. Case II.



——— Locus of zeros

– – – Approximating parabolas

Figure 6. Intersection of $f(\emptyset,\theta)$ with $\theta = \frac{\pi}{2}$ plane. Case II.

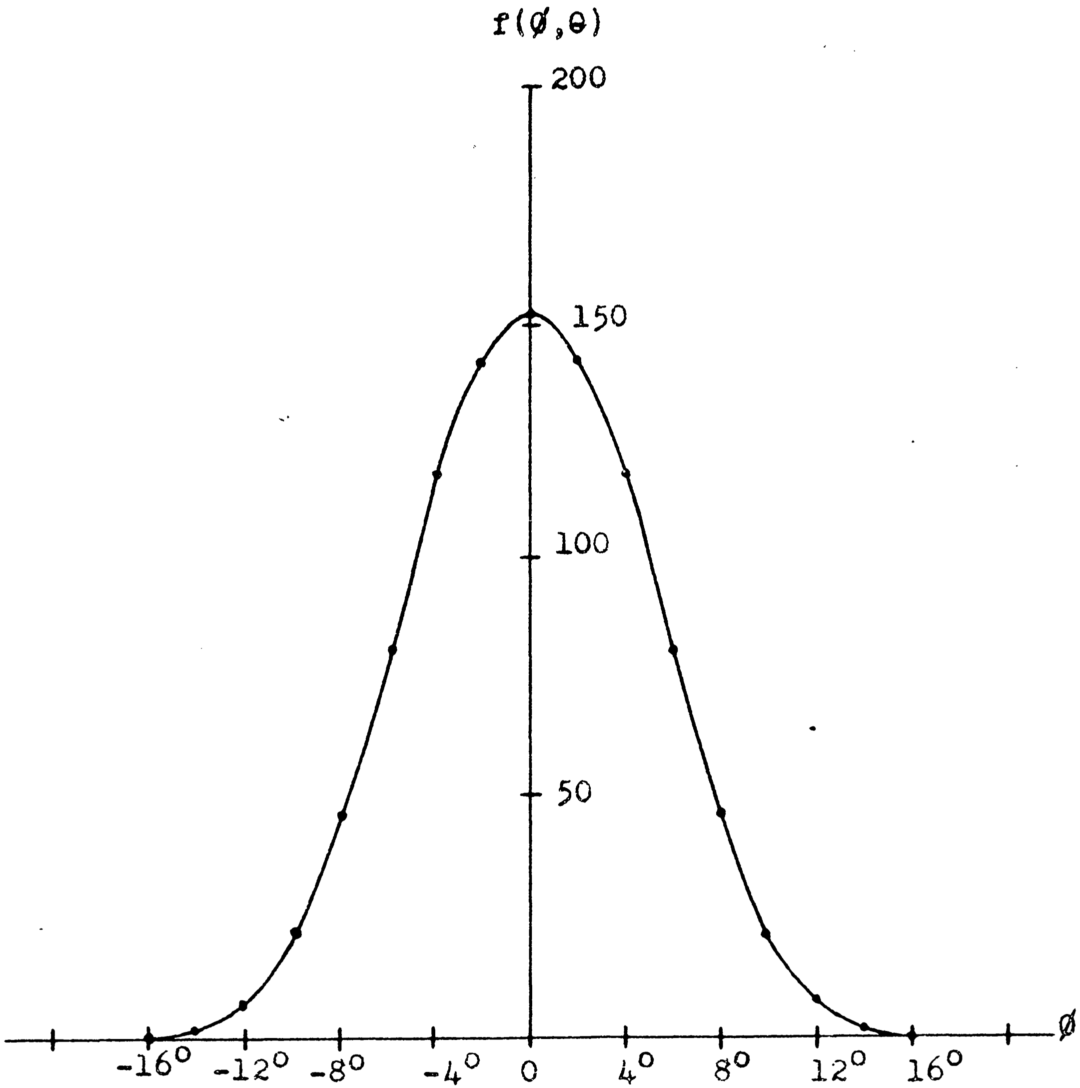


Figure 7. Intersection of $f(\phi,\theta)$ with ϕ = o plane. Case II.

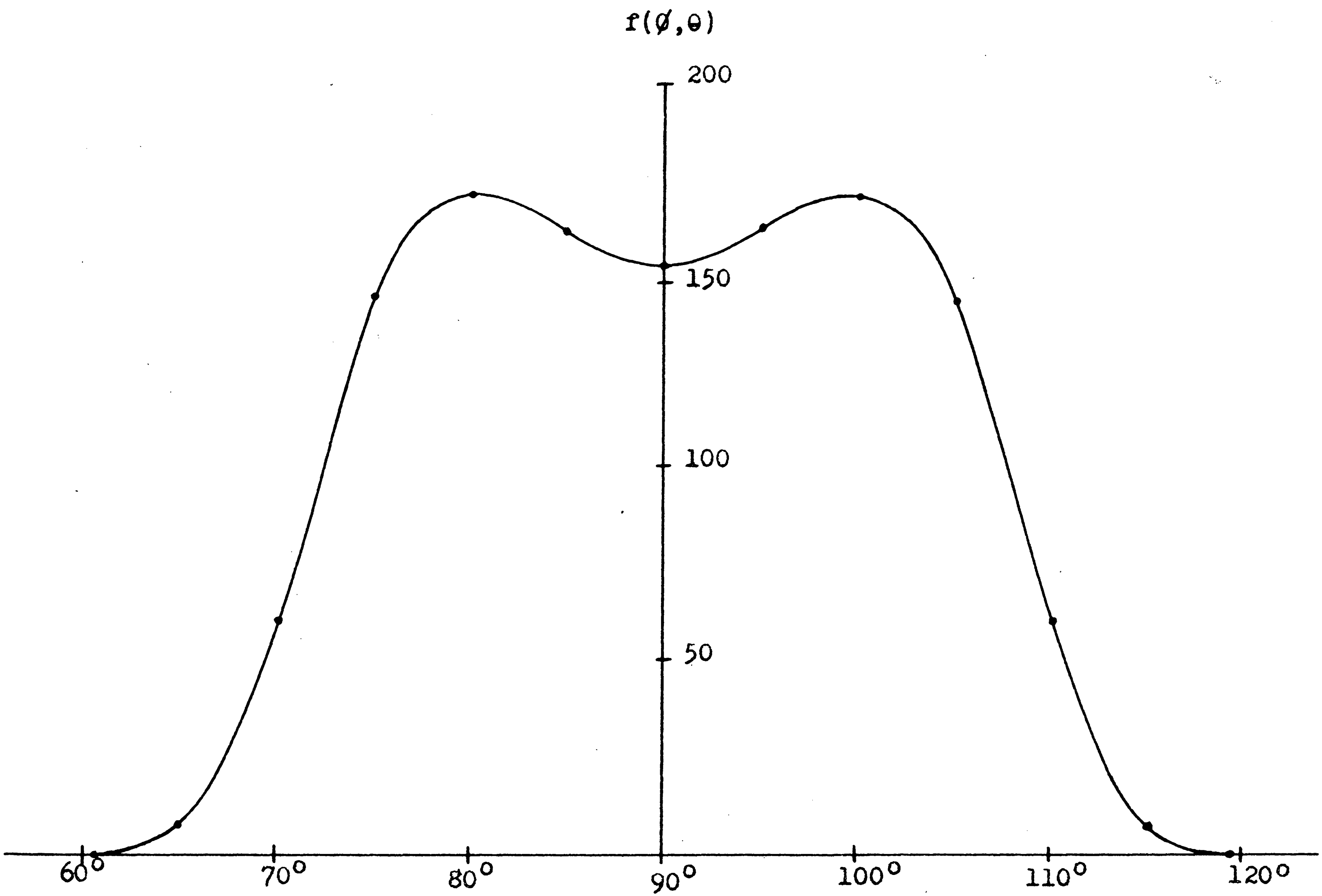
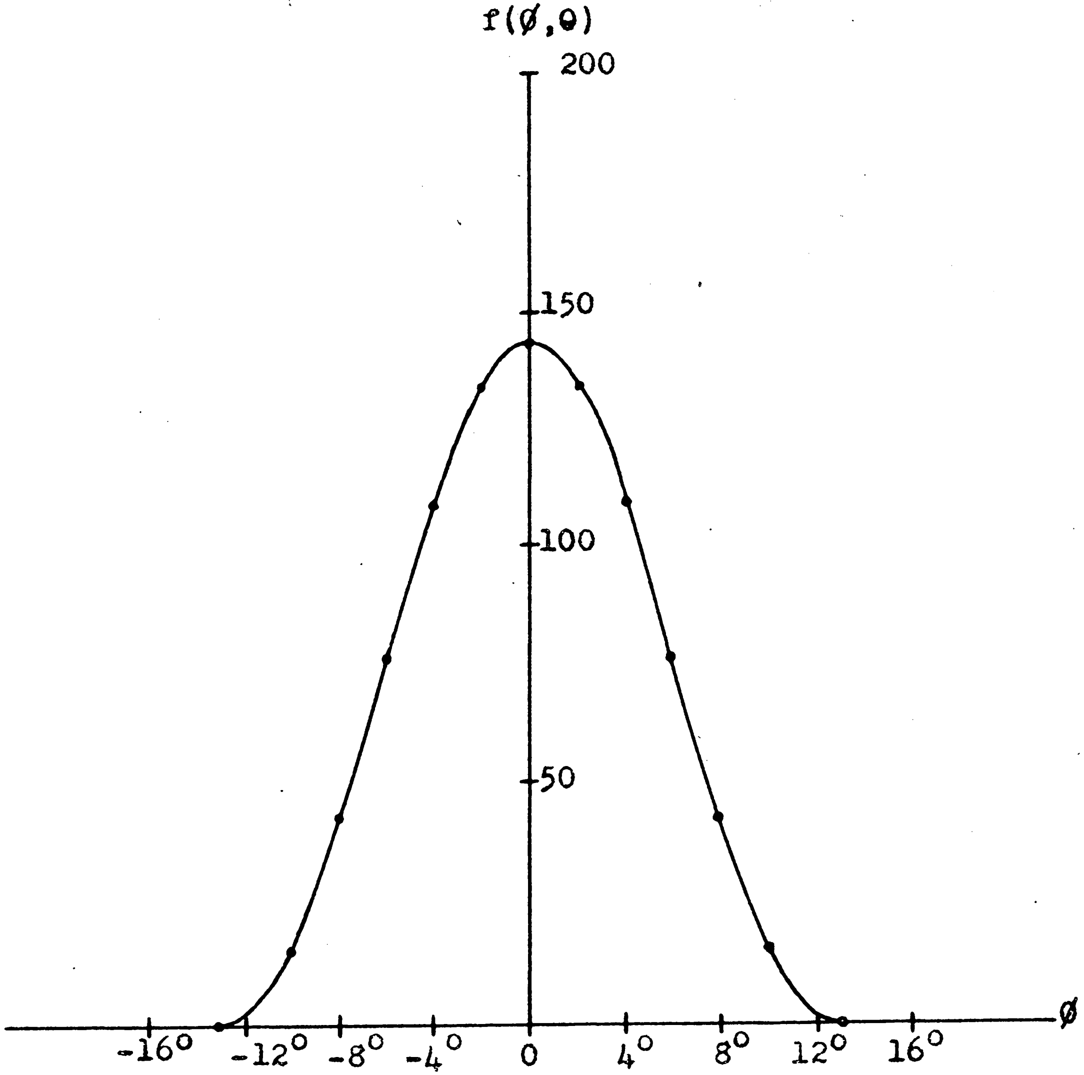


Figure 8. Intersection of $f(\phi,\theta)$ with the plane $\theta = \frac{5\pi}{12}$ (or $\theta = \frac{7\pi}{12}$).
Case II.



ILLUSTRATIVE EXAMPLE II

To illustrate the application of orthogonal polynomial methods to the study of trends, the estimation of observational error, and the evaluation of the integral we construct the following numerical example. The values shown in Table I are based on the function $f(x,y) = 65 + 4x - y + 2x^2 - xy^2$. In column 3 are shown the values calculated from the function. A random error drawn from a normal distribution with zero mean and variance 9 has been added to each of the function values in column 3 and the result entered as Z_{ij} in column 4. The values in column 4 are used in the analysis.

We define the first of the orthogonal polynomials, $\xi'_0 \equiv 1$, and the next four are given by;

$$\xi'_1 = \lambda_1 (X - \overline{X})$$

$$\xi'_2 = \lambda_2 \left[(X-\overline{X})^2 - \frac{n^2 - 1}{12} \right]$$

$$\xi'_3 = \lambda_3 \left[(X - \overline{X})^3 - (X - \overline{X}) \frac{3n^2-7}{20} \right]$$

$$\xi'_4 = \lambda_4 \left[(X-\overline{X})^4 - (X-\overline{X})^2 \frac{3n^2 - 13}{14} + \frac{3(n^2-1)(n^2-9)}{560} \right]$$

Where the λ_i are constants which can be chosen so that ξ'_i will be an integer for any value of n. For $n = 5$, the above set reduces to,

$$\xi'_1 = X - \overline{X}$$

$$\xi'_2 = (X - \overline{X})^2 - 2$$

$$\xi'_3 = \frac{5}{6} (X - \overline{X})^3 - \frac{17}{6} (X - \overline{X})$$

$$\xi'_4 = \frac{35}{12} (X - \overline{X})^4 - \frac{155}{12} (X - \overline{X})^2 + 6$$

$\bar{X} = 3$

The integers to which these polynomials reduce for given values of x or y are shown in the appropriate column of Table 1.

TABLE I

X	Y	f(x,y)	$\cong$ f(x,y) + error	$\xi'_1(x)$	$\xi'_2(x)$	$\xi'_3(x)$	$\xi'_4(x)$	$\xi'_1(y)$	$\xi'_2(y)$	$\xi'_3(y)$	$\xi'_4(y)$
1	1	69	66	-2	2	-1	1	-2	2	-1	1
2	1	78	81	-1	-1	2	-4	-2	2	-1	1
3	1	91	92	0	-2	0	6	-2	2	-1	1
4	1	108	110	1	-1	-2	-4	-2	2	-1	1
5	1	129	125	2	2	1	1	-2	2	-1	1
1	2	65	68	-2	2	-1	1	-1	-1	2	-4
2	2	71	74	-1	-1	2	-4	-1	-1	2	-4
3	2	81	79	0	-2	0	6	-1	-1	2	-4
4	2	95	99	1	-1	-2	-4	-1	-1	2	-4
5	2	113	115	2	2	1	1	-1	-1	2	-4
1	3	59	58	-2	2	-1	1	0	-2	0	6
2	3	60	61	-1	-1	2	-4	0	-2	0	6
3	3	65	63	0	-2	0	6	0	-2	0	6
4	3	74	73	1	-1	-2	-4	0	-2	0	6
5	3	87	93	2	2	1	1	0	-2	0	6
1	4	51	54	-2	2	-1	1	1	-1	-2	-4
2	4	45	44	-1	-1	2	-4	1	-1	-2	-4
3	4	43	39	0	-2	0	6	1	-1	-2	-4
4	4	45	50	1	-1	-2	-4	1	-1	-2	-4
5	4	51	49	2	2	1	1	1	-1	-2	-4
1	5	41	43	-2	2	-1	1	2	2	1	1
2	5	26	23	-1	-1	2	-4	2	2	1	1
3	5	15	16	0	-2	0	6	2	2	1	1
4	5	8	3	1	-1	-2	-4	2	2	1	1
5	5	5	5	2	2	1	1	2	2	1	1

$$f(x,y) = 65 + 4x - y + 2x^2 - xy^2$$

The $B_{\alpha\beta}$ as given by (91) can now be calculated conveniently from Table 1.

Since $\xi'_o \equiv 1$, B_{oo} is the total sum of the z's divided by the product of the sum of squares of the ξ'_o (x) and ξ'_o (y) which is $\frac{1582}{(5)(5)} = 63.28$. This value is simply the mean of the z's. The reduction in sum of squares due to B_{oo} is the product of 63.28 and 1582 which is 100108.96. It will be observed that this term is the correction term for the mean in the usual method of calculating the variance of a set of numbers.

B_{21} would be calculated as $(66)(2)(-2) + (78)(-1)(-2) + \ldots\ldots + (3)(-1)(2) + 5\,(2)(2)$ divided by $(2)^2 + (-1)^2 + (-2)^2 + (-1)^2 + (2)^2$ times $(-2)^2 + (-1)^2 + (o)^2 + (1)^2 + (2)^2$, which reduces to $\frac{61}{(14)(10)} = 0.4357$. The reduction in residual sum of squares attributable to B_{21} is the product of 0.4357 and 61 which is 26.58.

The value of 15 $B_{\alpha\beta}$ and the reduction in residual sum of squares which is associated with each of these is shown in Table 2.

TABLE 2

Coefficient	Value	Reduction in R.S.S.	Coefficient	Value	Reduction in R.S.S.
B_{oo}	63.28	100108.96	B_{21}	0.4357	26.58
B_{1o}	4.96	1230.08	B_{12}	-1.2643	223.78
B_{o1}	-19.34	18701.78	B_{03}	0.28	3.92
B_{2o}	2.2571	356.62	B_{40}	-0.1943	13.21
B_{11}	-6.09	3708.81	B_{31}	-0.12	1.44
B_{o2}	-3.3857	802.41	B_{22}	-0.3418	22.90
B_{3o}	-0.12	.72	B_{13}	-0.02	.04
			B_{o4}	-0.1086	4.13

The total sum of squares of the 25 z values is 125, 322 and the total reduction in this quantity explained by the 15 coefficients shown in Table 2 is 125,205. The difference of these numbers is 117 and is associated with the remaining 10 degrees of freedom, each of the B's accounting for a single degree of freedom. If we assume that there is no further trend in this data, our estimate of error variance is 11.7. It is evident that a number of the B's which have been calculated reflect only error and it would seem reasonable in this example to assume that only the 7 coefficients which account for the largest reduction in sum of squares represent real effects and under this assumption, the residual sum of squares would be 190 which divided by 18 degrees of freedom would yield an error variance estimate of 10.5. Consulting an F table, we see that any of the 7 coefficients which account for the largest portion of reduction in variance tested against either of the above estimates of error are significant at the 1% level.

It is not intended to be implied that in an analysis such as is illustrated by this example, one will always be able to say with great confidence exactly when the trends have been eliminated. Some coefficients may account for a reduction in residual variance which leaves the question of their representing a real effect quite doubtful. This method, nevertheless, appears to possess merit for assisting one in making an estimate of the observational error and usually we should have little trouble in deciding when the trends have been eliminated.

In order to calculate the double integral as given by (94) of the fourth degree function which we have fitted to the z values, we need the value of the integral of each of the orthogonal polynomials written at the bottom of page 1. These are given by

$$\int_1^n \xi_0' \, dx = \int_1^5 dx = 4 \qquad \int_1^5 \xi_3' \, dx = 0$$

$$\int_1^n \xi_1' \, dx = \int_1^5 (x-3) \, dx = 0 \qquad \int_1^5 \xi_4' \, dx = -\frac{68}{9}$$

$$\int_1^5 \xi_2' \, dx = -\frac{8}{3}$$

Since the value of the integrals of the orthogonal polynomials of odd degree are zero, all terms in (94) except those in which both subscripts are even vanish. For this example (94) reduces to

$$\begin{aligned} & (63.28)(4)(4) + (2.2571)(-8/3)(4) + (-3.3857)(4)(-8/3) \\ + \; & (-0.1943)\left(-\frac{68}{9}\right)(4) + (-0.3418)(-8/3)^2 + (-0.1086)(4)\left(-\frac{68}{9}\right) \\ = \; & 1031.24 \end{aligned}$$

For comparison with this value, the double integral of the function $f(x,y)$ taken over this rectangle yields 1018.67 and the thirteen point fifth degree accuracy formula given by (53) of section 9, yields 1014.67. The calculation from the orthogonal polynomials seems somewhat higher than we would expect from the other two calculations and presumably reflects the failure of the effects of the added errors to neutralize themselves as well as would be expected usually. The average of the errors added to obtain the values in column 4 of Table I is +.28 and this explains a part of the increased value shown in our first estimate of the integral.

ACKNOWLEDGEMENT

During the preparation of this thesis, I have had the encouragement and assistance of each member of my Doctoral Committee composed of Drs. D. B. DeLury, Boyd Harshbarger, T. W. Hatcher and H. W. Jackson, and of Dr. Louis O'Shaughnessy, Director of Graduate Studies. I am especially indebted to Dr. DeLury, now Director of the Department of Mathematical Statistics at the Ontario Research Foundation, under whose guidance this study was started during the 1946-47 college session and who kindly consented to continue in active supervision though he knew this would necessitate the interchange of numerous letters.

Since January 1948, I have served as a member of the Consulting Staff at the U. S. Navy Electronics Laboratory, San Diego, California and while, for the most part, the theoretical work on this problem was continued independently of my duties with the Navy, some applications of the theory was made at the Laboratory. During this period, a keen interest in certain phases of this study has been shown by the following colleagues of mine: Drs. A. B. Focke and G. H. Curl, Physicists; and Dr. F. H. Fisher and Mr. C. J. Van Vliet, Mathematicians; discussion with whom have helped clarify a number of troublesome points.

It is a pleasure for me to express my thanks for the stimulation and help received from each of those named above.

GEORGE W. TYLER.

San Diego, California

April 1949

B I B L I O G R A P H Y

1. Whittaker, E. T. and Robinson, G.
The Calculus of Observations - Blackie & Sons, London 1937.

2. Gauss, Carl Frederick
Werke - Dritter Band, Gottingen 1876.

3. Whittaker, E. T. and Watson, G. N.
Modern Analysis - Cambridge, 1946.

4. Jordan, Charles
Calculus of Finite Differences - New York, 1947.

5. Milne - Thomson, L. M.
The Calculus of Finite Differences - London, 1933.

6. Walsh, J. L.
Interpolation and Approximation of Rational Functions in the Complex Domain - American Mathematical Society Colloquium Publications - New York, 1935.

7. Ritt, J. F.
Integration in Finite Terms - New York, 1948.

8. Saks, S.
Theory of the Integral - Warsaw, 1937. (G.E.Stechert Co., New York).

9. Tchebichef, M. P.
Sur Les Quadratures - Journal De Mathematique, Pures et Appliques, Tome XIX, 1874.

10. Sheppard, W. F.
On Central - Difference Formulae - Proceedings of the London Mathematical Society, Vol. XXXI, 1899.

11. Moors, B. P.
Valuers Approximative d'une Integrale Definie - Gauthier Villars, Paris, 1905.

12. Burnside, W.
An approximate Quadrature Formula - Messenger of Mathematics, Vol. 37, 1908.

13. Moors, B. P.
Etude sur les Formules (Specialement de Gauss) Servant a Calculer des Valuers Approximative d'une Integrale Definie.
Verhandelingen der Koninklijke Akademie van Wetenschappen te Amsterdam deel XI No. 6. - Amsterdam, 1913.

14. Aitken, A.C. and Frewin, G. L.
The Numerical Evaluation of Double Integrals - Proceedings of the Edinburgh Mathematical Society, Vol. 42, 1923.

B I B L I O G R A P H Y (Cont'd.)

15. O'Toole, A. L.
On the Degree of Approximation of Certain Quadrature Formulas. The Annals of Mathematical Statistics, Vol. 4, 1933.

16. Sadowsky, Michael
A Formula for Approximate Computation of a Triple Integral - American Mathematical Monthly. Vol. 47, 1940.

17. Bickley, W. G.
Finite Difference Formulae for the Square Lattice - The Quarterly Journal of Mechanics and Applied Mathematics, Vol 1, 1948.

18. Anderson, R. L. and Houseman, E. E.
Tables of Orthogoral Polynomial Values Extended to N = 104. Iowa State College Research Bulletin 297, April 1942.

19. Cramer, Harald
Mathematical Methods in Statistics - Princeton - 1946.

20. Fisher, R. A.
Statistical Methods for Research Workers - London - 1948.

21. Snedecor, George W.
Statistical Methods - Ames - 1940.

22. Tintner, Gerhard
The Variate Difference Method - Bloomington - 1940.