III. *Contributions to the Mathematical Theory of Evolution.*

By KARL PEARSON, *University College, London.*

Communicated by Professor HENRICI, *F.R.S.*

Received October 18,—Read November 16, 1893.

[PLATES 1—5.]

CONTENTS.

I.—*On the Dissection of Asymmetrical Frequency-Curves.*

(1.) IF measurements be made of the same part or organ in several hundred or thousand specimens of the same type or family, and a curve be constructed of which the abscissa x represents the size of the organ and the ordinate y the number of specimens falling within a definite small range δx of organ, this curve may be termed a *frequency-curve.* The centre or origin for measurement of the organ may, if we please, be taken at the *mean* of all the specimens measured. In this case the frequency-curve may be looked upon as one in which the frequency—per thousand or per ten thousand, as the case may be—of a given small range of deviations from the mean, is plotted up to the mean of that range. Such frequency-curves play a large part in the mathematical theory of evolution, and have been dealt with by Mr. F. GALTON, Professor WELDON, and others. In most cases, as in the case of errors of observation, they have a fairly definite symmetrical shape* and one that

* Symmetrical shapes may of course occur which are not of the normal or error-curve form. See Part II., § 11 of this paper.

9.5.94

approaches with a close degree of approximation to the well-known error or probability-curve. A frequency-curve, which, for practical purposes, can be represented by the error curve, will for the remainder of this paper be termed a *normal curve.* When a series of measurements gives rise to a normal curve, we may probably assume something approaching a stable condition; there is production and destruction impartially round the mean. In the case of certain biological, sociological, and economic measurements there is, however, a well-marked deviation from this normal shape, and it becomes important to determine the direction and amount of such deviation. The asymmetry may arise from the fact that the units grouped together in the measured material are not really homogeneous. It may happen that we have a mixture of 2, 3, . . . n homogeneous groups, each of which deviates about its own mean symmetrically and in a manner represented with sufficient accuracy by the normal curve. Thus an abnormal frequency-curve may be really built up of normal curves having parallel but not necessarily coincident axes and different parameters. Even where the material is really homogeneous, but gives an abnormal frequency-curve the amount and direction of the abnormality will be indicated if this frequency-curve can be split up into normal curves. The object of the present paper is to discuss the dissection of abnormal frequency-curves into normal curves. The equations for the dissection of a frequency-curve into n normal curves can be written down in the same manner as for the special case of $n = 2$ treated in this paper; they require us only to calculate higher moments. But the analytical difficulties, even for the case of $n = 2$, are so considerable, that it may be questioned whether the general theory could ever be applied in practice to any numerical case.

There are reasons, indeed, why the resolution into two is of special importance. A family probably breaks up first into two species, rather than three or more, owing to the pressure at a given time of some particular form of natural selection; in attempting to procure an absolutely homogeneous material, we are less likely to have got a mixture of three or more heterogeneous groups than of two only. Lastly, even where the heterogeneity may be threefold or more, the dissection into two is likely to give us, at any rate, an approximation to the two chief groups. In the case of homogeneous material, with an abnormal frequency-curve, dissection into two normal curves will generally give us the amount and direction of the chief abnormality. So much, then, may be said of the value of the special case dealt with here.

A distinction must be made between the two cases which may *theoretically* occur. If we have a real mixture of two normal groups represented by our abnormal frequency-curve, then, theoretically, it is possible to find the two components, and these two components must be unique. If they were not unique, a relation of the following kind must hold for every value of x:—

$$\frac{c_1}{\sigma_1\sqrt{(2\pi)}} e^{-\frac{(x-b_1)^2}{2\sigma_1^2}} + \frac{c_2}{\sigma_2\sqrt{(2\pi)}} e^{-\frac{(x-b_2)^2}{2\sigma_2^2}} = \frac{c_3}{\sigma_3\sqrt{(2\pi)}} e^{-\frac{(x-b_3)^2}{2\sigma_3^2}} + \frac{c_4}{\sigma_4\sqrt{(2\pi)}} e^{-\frac{(x-b_4)^2}{2\sigma_4^2}}.$$

Between the six constants on either side of this equation an infinite variety of relations can be reached by giving x an infinite variety of values, and it seems impossible to satisfy this series by the same set of values of the constants. For example, let x be very great, and suppose σ_1 to be the largest of all the quantities σ_1, σ_2, σ_3, and σ_4. Dividing by $\frac{1}{\sqrt{(2\pi)}} e^{-\frac{(x-b_1)^2}{2\sigma_1^2}}$ and putting x very great we have

$$\frac{c_1}{\sigma_1} + \frac{c_2}{\sigma_2} e^{-\frac{x^2}{2}\left(\frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2}\right)} = \frac{c_3}{\sigma_3} e^{-\frac{x^2}{2}\left(\frac{1}{\sigma_3^2} - \frac{1}{\sigma_1^2}\right)} + \frac{c_4}{\sigma_4} e^{-\frac{x^2}{2}\left(\frac{1}{\sigma_4^2} - \frac{1}{\sigma_1^2}\right)},$$

whence, proceeding to the limit,

$$c_1/\sigma_1 = 0,$$

unless $\sigma_1 = \sigma_3$ or σ_4.

The first is impossible by hypothesis, therefore the latter must be true, say $\sigma_1 = \sigma_3$. This gives us at once $c_1 = c_3$.

Returning to the original equation, and making x large in it, we see that the first two terms become equal on either side. Hence, the second two terms must become equal as x approaches infinity, or

$$\frac{c_2}{\sigma_2} e^{-\frac{x^2}{2\sigma_2^2}} = \frac{c_4}{\sigma_4} e^{-\frac{x^2}{2\sigma_4^2}}.$$

Dividing again by $e^{-\frac{x^2}{2\sigma_2^2}}$, this leads in the same manner as before to $\sigma_2 = \sigma_4$, and, ultimately, to $c_2 = c_4$.

Our original equation may now be written

$$\frac{c_1}{\sigma_1\sqrt{(2\pi)}}\left\{e^{-\frac{(x-b_1)^2}{2\sigma_1^2}} - e^{-\frac{(x-b_3)^2}{2\sigma_1^2}}\right\} = \frac{c_2}{\sigma_2\sqrt{(2\pi)}}\left\{e^{-\frac{(x-b_4)^2}{2\sigma_2^2}} - e^{-\frac{(x-b_2)^2}{2\sigma_2^2}}\right\} \quad . \quad . \quad (\eta).$$

Put $x = \frac{1}{2}(b_1 + b_3)$, then the left-hand side vanishes and, accordingly, the right must vanish, but this involves either

$$b_2 = b_4,$$

or

$$b_1 + b_3 = b_2 + b_4.$$

Similarly, putting $x = \frac{1}{2}(b_2 + b_4)$, we find that either

$$b_1 = b_3,$$

or

$$b_1 + b_3 = b_2 + b_4 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\alpha).$$

Thus, either the two sets of components are identical, or (α) is true, Multiply equation (η) above by x, x^3 and x^5 in succession, and integrate the results respectively between the limits α and $-\alpha$.* We find

$$(b_1 - b_3)\, c_1 = (b_4 - b_2)\, c_2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\beta),$$

$$\{3\sigma_1^2 (b_1 - b_3) + b_1^3 - b_3^3\}\, c_1 = \{3\sigma_2^2 (b_4 - b_2) + b_4^3 - b_2^3\}\, c_2,$$

reducing by aid of (α) and (β) to

$$3\sigma_1^2 - b_1 b_3 = 3\sigma_2^2 - b_2 b_4 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\gamma),$$

and

$$\begin{aligned}&\{15\sigma_1^4 (b_1 - b_3) + 10\sigma_1^2 (b_1^3 - b_3^3) + b_1^5 - b_3^5\}\, c_1 \\ &\quad = \{15\sigma_2^4 (b_4 - b_2) + 10\sigma_2^2 (b_4^3 - b_2^3) + b_4^5 - b_2^5\}\, c_2,\end{aligned}$$

reducing by aid of (α), (β), and (γ) to the two forms,

$$2\sigma_2^2 + 8\sigma_1^2 + 3b_1^2 + 3b_3^2 + 4b_1 b_3 = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\delta),$$

$$2\sigma_1^2 + 8\sigma_2^2 + 3b_2^2 + 3b_4^2 + 4b_2 b_4 = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\epsilon).$$

Equations (α), (β), (δ), and (ϵ) are four independent equations, which suffice to determine b_1, b_2, b_3, b_4, as definite functions of σ_1, σ_2, c_1, and c_2. But b_1, b_2 are in general independent of σ_1, σ_2, c_1, and c_2; hence it follows that (α) cannot in general be true, or we must have $b_1 = b_3$ and $b_2 = b_4$. That is, a curve which breaks up into two normal components can break up in one way, and one way only.

Now it is clear that in actual statistical practice our abnormal frequency-curve will never be the absolutely true sum of two normal-curves; indeed, if it be not a mixture, but an asymmetrical frequency-curve, it is not necessarily a very close approach to the sum of *two* frequency-curves of normal type,—it may be the limit to an asymmetrical binomial.† We must not, therefore, be surprised if more than one solution be given by any method of dissection. A mathematical criterion for discriminating the "true" solution might easily be given. For example, in the method of the present paper, we might define that as the "true," or at any rate the "best," solution which gave for the compound-curve a sixth moment nearest in value to that of the observation-curve. Such a theoretical criterion, however, may not have much

* The values of the successive moments of the normal-curve are given in § 5 of this paper, and permit of these integrations being performed at once.

† The general form of the limit to asymmetrical binomials is

$$y = \mathrm{C}\left(\beta + \frac{x}{c}\right)^{2\beta - 1} e^{-x/2c}$$

where C, c, and β are constants, and x is to have positive values only. β is always positive. [A slightly fuller form is given in the abstract of this paper, 'Roy. Soc. Proc.,' vol. 54, p. 331.]

practical value. For after we have made the areas and first five moments of two curves identical, their sixth moments will in general be (like their contours) much closer together than either are to that of the curve of observations. Added to this the great labour involved in the calculation of the sixth moment is sufficient to deter the practical statistician, if any other convenient mode—*e.g.*, results of measurement on other organs—suffices in the particular case to discriminate between the solutions found. Thus, while the mathematical solution should be unique, yet from the utilitarian standpoint we have to be content with a compound curve which fits the observations closely, and more than one such compound curve may arise. All we can do is to adopt a method which minimizes the divergences of the actual statistics from a mathematically true compound. The utilitarian problem is to find the *most likely* components of a curve which is not the true curve, and would only be the true curve had we an infinite number of absolutely accurate measurements. As there are different methods of fitting a normal curve to a series of observations, depending on whether we start from the mean or the median, and proceed by "quartiles," mean error or error of mean square, and as these methods lead in some cases to slightly different normal-curves, so various methods for breaking up an abnormal frequency-curve may lead to different results. As from the utilitarian standpoint good results for a simple normal curve are obtained by finding the mean from the first moment, and the error of mean square from the second moment, so it seems likely that the present investigation, based on the first five or six moments of the frequency-curve, may also lead to good results. While a method of equating chosen ordinates of the given curve and those of the components leaves each equation based only on the measurements of organs of one size, the method of moments uses *all* the given data in the case of each equation for the unknowns, and errors in measurement will, thus, individually have less influence. At the same time it would be of great interest to discover whether other methods of dissection lead to results identical or nearly identical with the method of moments adopted by the present writer. Any other method analytically possible has not yet, however, occurred to him; nor any criterion for distinguishing practically between two solutions so close as those of figs. 1 and 2, other than that adopted by Professor WELDON when he appeals to the measurements of a correlated organ.

(2.) In the case of a frequency-curve whose components are two normal curves, the complete solution depends in the method adopted in finding the roots of a numerical equation of the *ninth* order. It is possible that a simpler solution may be found, but the method adopted has only been chosen after many trials and failures. Clearly each component normal curve has three variables: (i.) the position of its axis, (ii.) its "standard-deviation" (GAUSS's "Mean Error," AIRY's "Error of Mean Square"), and (iii.) its area. Six relations between the given frequency-curve and its component curves would therefore suffice to determine the six unknowns. Innumerable relations of this kind can be written down, but, unfortunately, the majority of them lead to exponential equations, the solution of which seems more beyond the wit of man than that of a numerical equation even of the ninth order.

(3.) In any given example the conditions will be sufficient to reduce the suitable roots of this equation very largely, possibly to two or even one. These limiting conditions will be considered later. A suitable root of this equation leads to a quadratic for the areas of the two component normal curves. This quadratic is fundamental, and appears to be highly suggestive for the problem of evolution. We have two cases:

(i.) *Both its roots are positive.*

In this case the given frequency-curve is the *sum* of two normal curves. The units of the frequency-curve may be considered as composed of definite proportions of two species, each of which is stable about its mean. The process of differentiation here appears complete.

(ii.) *One root is positive and the other negative.*

The given frequency-curve is now the *difference* of two probability-curves. The probability-curve, with positive area, may possibly now be looked upon as the birth-population (unselectively diminished by death). The negative probability-curve is a selective diminution of units about a certain mean; that mean may, perhaps, be the average of the less "fit."

It is possible that in some numerical cases solutions of both the types (i.) and (ii.) will be found to exist, but I imagine that in most cases of a well-marked and characteristic asymmetrical frequency-curve, either only one type of solution will exist, or, if two types do exist, then one will give a much better agreement with the actual shape of the curve than the other. That the two types of solutions should exist side by side *occasionally* is, perhaps, to be expected. In such cases we have examples of groups, which are, perhaps, in process of differentiation into separate species by the elimination of members round a selected mean.

(iii.) From the nature of the problem, the case of both roots negative does not occur.

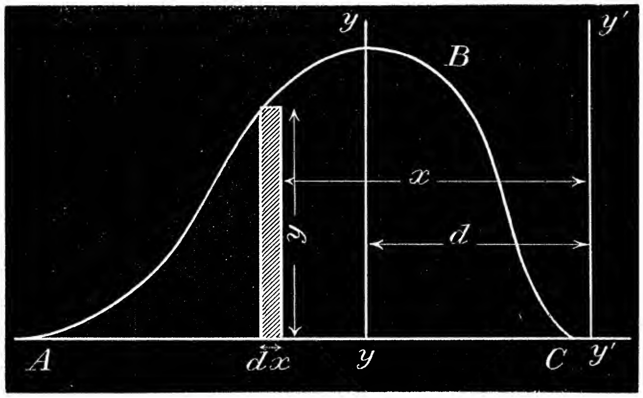
We now pass to the solution of the problem:

Given an asymmetrical frequency-curve to break it up, if possible, into two component probability-curves, or into two normal curves.

(4.) *Preliminary Definitions and Problems.*

(i.) Given any curve ABC, and the line $y'y'$, if we take the sum of the products of every element of area by the nth power of the distance of the element from the line $y'y'$, we form the nth moment of the area about the line yy'.

Clearly, if y be the length of a strip parallel to $y'y'$ and x its distance from $y'y'$, then the nth moment $= \int x^n y \, dx$, the integration extending all over ABC, or from A to C in our case, where the curve is always bounded by a straight line, AC, perpendicular to $y'y'$.



If h be any standard length, say 10 or 100 units, then the nth moment is of the order $h^n\alpha$, if α be the area of ABC. It therefore equals $\mu'_n h^n \alpha$, where μ'_n is a purely numerical factor. We shall invariably represent it as the product of these three factors.

(ii.) Given the first n moments about $y'y'$, or the coefficients $\mu'_1, \mu'_2, \mu'_3, \mu'_4 \ldots \mu'_n$, to find the nth moment about yy or the coefficient μ_n.

Let the distance between yy and $y'y'$ be $d = qh$, then

$$\mu_n h^n \alpha = \int (x - d)^n y \, dx,$$

or

$$\mu_n = \mu'_n - nq\mu'_{n-1} + \frac{n(n-1)}{1 \cdot 2} q^2 \mu'_{n-2} - \frac{n(n-1)(n-2)}{\lfloor 3} q^3 \mu'_{n-3} +, \&c. \ldots$$

In particular, since $\mu'_0 = 1$,

$$\left.\begin{aligned}
\mu_1 &= \mu'_1 - q \\
\mu_2 &= \mu'_2 - 2q\mu'_1 + q^2 \\
\mu_3 &= \mu'_3 - 3q\mu'_2 + 3q^2\mu'_1 - q^3 \\
\mu_4 &= \mu'_4 - 4q\mu'_3 + 6q^2\mu'_2 - 4q^3\mu'_1 + q^4 \\
\mu_5 &= \mu'_5 - 5q\mu'_4 + 10q^2\mu'_3 - 10q^3\mu'_2 + 5q^4\mu'_1 - q^5
\end{aligned}\right\} \quad \ldots \quad (1).$$

When the line $y'y'$ passes through the centroid of the curve, and the curve is symmetrical about $y'y'$ μ'_1, μ'_3, μ'_5 are all zero. Hence if in this case we take yy to the right of $y'y'$, or d negative,

$$\left.\begin{aligned}\mu_1 &= q\\ \mu_2 &= \mu'_2 + q^2\\ \mu_3 &= 3q\mu'_2 + q^3\\ \mu_4 &= \mu'_4 + 6q^2\mu'_2 + q^4\\ \mu_5 &= 5q\mu'_4 + 10q^3\mu'_2 + q^5\end{aligned}\right\} \quad \ldots \ldots \ldots \quad (2).$$

(iii.) The distance of the centroid of ABC from $y'y'$ is the ratio of its first moment $\mu'_1 h\alpha$ to its area α, and $= \mu'_1 h$.

(iv.) To find the successive moments of a given curve about a given line.

For the purposes of the present problem we require only the first five moments of a curve like ABC about a line yy passing through its centroid. The solution may be obtained either analytically or graphically according to the accuracy or rapidity with which we wish to work.

(a.) *Analytically*.—Suppose the frequency-curve to be obtained by plotting up the results of 1000 measurements, each unit of length along AC corresponding to an equal change in the deviation. Starting from the point C, beyond which no individual occurs, we may have in practice, perhaps, 20 to 30 equal ranges of deviations before we reach the point A, which terminates the deviations on the left. The equal range being taken as the unit of length, let the numbers in the groups at 1, 2, 3, 4, 5 . . . units of distance from C be $y_1, y_2, y_3, y_4, y_5 \ldots$

Then the n^{th} moment clearly equals very approximately

$$1^n \times y_1 + 2^n \times y_2 + 3^n \times y_3 + 4^n \times y_4 + \ldots,$$

or since $\alpha = 1000$, and h may be conveniently taken $= 100$,

$$\mu'_n = \frac{1^n \times y_1 + 2^n \times y_2 + 3^n \times y_3 + 4^n \times y_4 + \ldots}{100^n \times 1000} \quad \ldots \ldots \quad (3).$$

Sufficiently accurate values can then be found for $\mu'_1, \mu'_2, \mu'_3, \mu'_4, \mu'_5$, provided we know the 2nd, 3rd, 4th, and 5th powers of the natural numbers up to about 20 to 30. The values of these powers up to 30 are given later in this paper.

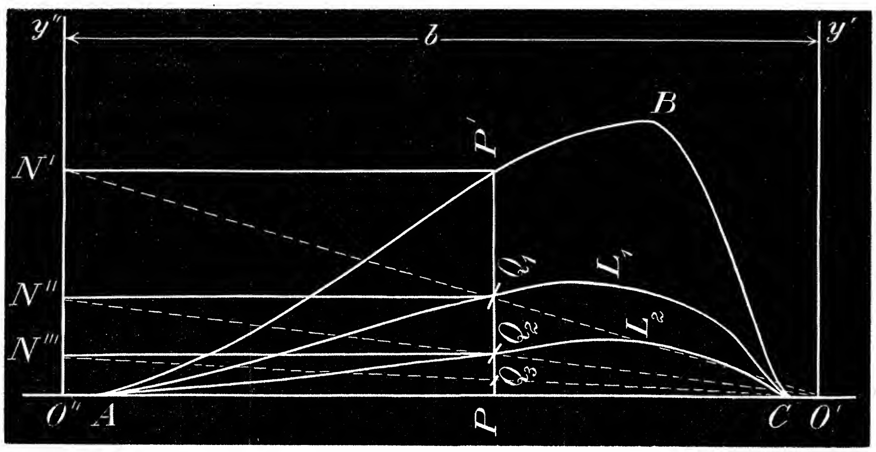
Knowing the first five moments about the vertical through C, we can find the centroid by aid of (iii.) above, and then the moments about the vertical through the centroid by aid of equations (1).

Since $\mu_1 = 0$ for the centroid $\mu'_1 = q$, and therefore we have the following to determine the other moments :—

$$\left.\begin{aligned}\mu_2 &= \mu'_2 - q^2\\ \mu_3 &= \mu'_3 - 3q\mu'_2 + 2q^3\\ \mu_4 &= \mu'_4 - 4q\mu'_3 + 6q^2\mu'_2 - 3q^4\\ \mu_5 &= \mu'_5 - 5q\mu'_4 + 10q^2\mu'_3 - 10q^3\mu'_2 + 4q^5\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad (4).$$

The centroid having been found, it may be asked : Why we should not calculate μ_2, μ_3, μ_4, μ_5 directly ? The answer lies in the fact that the centroid will not generally coincide with a unit division on the deviation axis, and the powers to be calculated, instead of being those of two place figures, become in general powers of numbers containing three or four figures. Thus the labour of the arithmetic is much increased.

(*b*.) *Graphically.*—If the figure be drawn on a large scale, the moments may be found with a fair degree of accuracy by aid of the following process, which has long been of use in graphical statics for finding the first, second, and third moments of plane areas.*



It is required to find the moments about $O'y'$ of the curve ABC, bounded by the straight line $O'CA$. Take $O''y''$ parallel to $O'y'$ and at distance h. Take any line PP', first to $O'y'$ from AC to ABC; let the perpendicular from P' on $O''y''$ meet it in N', and let $O'N'$ meet PP' in Q_1 ; let the perpendicular from Q' on $O''y''$ meet it in N'', and let $O'N''$ meet PP' in Q_2; let the perpendicular from Q_2 on $O''y''$ meet it in N''', and let $O'N'''$ meet PP' in Q_3. In this manner a series of points Q_1, Q_2, Q_3, Q_4, Q_5, are determined. Let these points be determined for a series of positions of PP' taken at short intervals from C to A, then all the corresponding Q being joined, we obtain curves termed respectively the first, second, third, fourth, and fifth moment-

* The third moment of a plane area is used in determining graphically the moment of inertia of a spindle about its axis. The method described is sometimes attributed to COLLIGNON, but seems to have been long in use to find "equivalent figures" in the case of beam sections.

curves. Let the areas AQ_1L_1C, AQ_2L_2C, &c., be read off with a planimeter, and be $\alpha_1, \alpha_2, \alpha_3 \ldots$ Then

$$\left.\begin{aligned}\mu_1' &= \alpha_1/\alpha \\ \mu_2' &= \alpha_2/\alpha \\ \mu_3' &= \alpha_3/\alpha \\ \mu_4' &= \alpha_4/\alpha \\ \mu_5' &= \alpha_5/\alpha\end{aligned}\right\} \quad \ldots \ldots \quad (5).$$

A good draughtsman will construct these curves with great readiness, and if on a sufficiently large scale, the results may be read to within the one per cent. error.*

Equations (4) then enable us to complete the problem of finding the moments about a line through the centroid. Or, the first moment being found about $O'y'$, and so the centroid determined; we may shift $O'y'$ till it passes through the centroid, and then proceed to find $\mu_2 \ldots \mu_5$ directly in the above manner. In this case care will have to be taken in reading the areas of the moment-curves, which have now pieces of their areas *negative*, to carry the planimeter point, in the proper sense, round their contours.

(5.) *Properties of the probability-curve.*

Let the equation to the probability-curve be—

$$y' = \frac{c}{\sigma\sqrt{(2\pi)}}\, e^{-x^2/(2\sigma^2)} \quad \ldots \ldots \quad (6).$$

Then σ will be termed its *standard-deviation* (error of mean square). c is the total number of units measured, or the area of the probability curve.

(i.) To find the second and fourth moments of the probability-curve about the axis of y'.

Let them be M_2' and M_4'.

Then

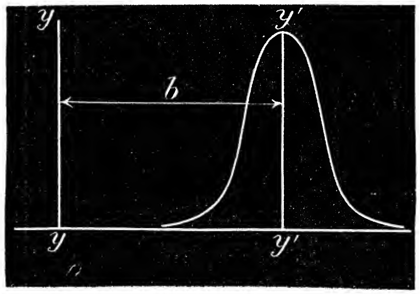
$$M_2' = 2\int_0^\alpha y'x^2\,dx = c \times \sigma^2.$$

$$M_4' = 2\int_0^\alpha y'x^4\,dx = c \times 3\sigma^4.$$

Clearly M_3' and M_5' are zero.

[* My demonstrator, Mr. G. U. YULE, has graphically calculated the first four moments of a number of statistical frequency-curves, with the object of fitting them to the generalized probability-curve (see footnote, p. 74). The method is sufficiently accurate in practice, and I hope soon to have an instrument to construct these curves mechanically, designed by him.—February 9, 1894.]

(ii.) Now let α be a standard area and h a standard length. Let us use



Equations (2) of Art. 4 (ii.), taking $y'y'$ as the axis of symmetry of the probability-curve, and yy at a distance b to the left, then—

$$\mu_1 hc = bc.$$
$$\mu_2 h^2 c = (\sigma^2 + b^2)\,c.$$
$$\mu_3 h^3 c = (3b\sigma^2 + b^3)\,c.$$
$$\mu_4 h^4 c = (3\sigma^4 + 6b^2\sigma^2 + b^4)\,c.$$
$$\mu_5 h^5 c = (15\sigma^4 b + 10b^3\sigma^2 + b^5)\,c.$$

Now let $c/\alpha = z$, $\sigma/b = u$, and $b/h = \gamma$.

Then z, u, and γ are purely numerical quantities, and we have for the first five moments round yy—

$$\left.\begin{aligned} M_1 &= \gamma z \alpha h, \\ M_2 &= \gamma^2 z\,(1 + u^2)\,\alpha h^2, \\ M_3 &= \gamma^3 z\,(1 + 3u^2)\,\alpha h^3, \\ M_4 &= \gamma^4 z\,(1 + 6u^2 + 3u^4)\,\alpha h^4, \\ M_5 &= \gamma^5 z\,(1 + 10u^2 + 15u^4)\,\alpha h^5, \end{aligned}\right\} \quad \ldots\ldots\ldots \quad (7).$$

(6.) We are now in a position to write down the equations which give the general solution of our problem. Let the deviation-axis of the asymmetrical frequency-curve be taken as axis of x, and let the axis of y be a perpendicular on this axis through the centroid of the frequency-curve. Let this centroid and the first five moment-coefficients about the axis of y of the frequency-curve, *i.e.*, 0, μ_2, μ_3, μ_4, μ_5, be found either analytically or graphically by the methods suggested in Art. 4 (iv.).

Then, if the position and magnitude of the component normal curves be given by the quantities b_1, c_1, σ_1, and b_2, c_2, σ_2, or the corresponding numerics

$$\gamma_1,\ z_1,\ u_1, \text{ and } \gamma_2,\ z_2,\ u_2,$$

we have, since moments round the vertical axis are clearly additive—

$$c_1 + c_2 = \alpha,$$

$$(\gamma_1 z_1 + \gamma_2 z_2)\, \alpha h = 0,$$

$$\{\gamma_1^2 z_1 (1 + u_1^2) + \gamma_2^2 z_2 (1 + u_2^2)\}\, \alpha h^2 = \mu_2 \alpha h^2,$$

$$\{\gamma_1^3 z_1 (1 + 3u_1^2) + \gamma_2^3 z_2 (1 + 3u_2^2)\}\, \alpha h^3 = \mu_3 \alpha h^3,$$

$$\{\gamma_1^4 z_1 (1 + 6u_1^2 + 3u_1^4) + \gamma_2^4 z_2 (1 + 6u_2^2 + 3u_2^4)\}\, \alpha h^4 = \mu_4 \alpha h^4,$$

$$\{\gamma_1^5 z_1 (1 + 10u_1^2 + 15u_1^4) + \gamma_2^5 z_2 (1 + 10u_2^2 + 15u_1^4)\}\, \alpha h^3 = \mu_5 \alpha h^5.$$

The first equation here represents the equality of the areas of the resultant curve and its components. Reducing to the simplest terms, we have the following six equations to find the six unknowns, z_1, z_2, γ_1, γ_2, u_1, u_2 :—

$$z_1 + z_2 = 1 \quad \text{. . . . .} \quad (8).$$

$$\gamma_1 z_1 + \gamma_2 z_2 = 0 \quad \text{. . . . .} \quad (9).$$

$$\gamma_1^2 z_1 (1 + u_1^2) + \gamma_2^2 z_2 (1 + u_2^2) = \mu_2 \quad \text{. . . . . . .} \quad (10).$$

$$\gamma_1^3 z_1 (1 + 3u_1^2) + \gamma_2^3 z_2 (1 + 3u_2^2) = \mu_3 \quad \text{. . . . . . .} \quad (11).$$

$$\gamma_1^4 z_1 (1 + 6u_1^2 + 3u_1^4) + \gamma_2^4 z_2 (1 + 6u_2^2 + 3u_2^4) = \mu_4 \quad \text{. . .} \quad (12).$$

$$\gamma_1^5 z_1 (1 + 10u_1^2 + 15u_1^4) + \gamma_2^5 z_2 (1 + 10u_2^2 + 15u_2^4) = \mu_5 \quad \text{. .} \quad (13).$$

Equations (8)–(13) give the complete solution of the problem.* After several trials, I find that the elimination of z_1, z_2, u_1, u_2 from these equations, and the determination of equations giving $\gamma_1 \gamma_2$ and $\gamma_1 + \gamma_2$ appear to lead to a resulting equation of the lowest possible order.

(7.) Eliminating z_2 between (8) and (9), we have

$$z_1 = -\frac{\gamma_2}{\gamma_1 - \gamma_2} \quad \text{. . . . .} \quad (14).$$

Similarly,

$$z_2 = \frac{\gamma_1}{\gamma_1 - \gamma_2} \quad \text{. . . . .} \quad (15).$$

* All my attempts to obtain a simpler set have failed. Equating of selected ordinates, or of selected portions of area, or of moments round the axis of x, all appear to lead to exponential equations defying solution. It is possible, however, that some other six equations of a less complex kind may ultimately be found.

Equations (14) and (15) clearly give the numbers in the component groups so soon as γ_1 and γ_2 are found.

Substituting these values of z_1 and z_2 in (10) and (11), we have two equations to determine u_1^2 and u_2^2 in terms of γ_1, γ_2. Solving them we find

$$\gamma_1 u_1^2 = \frac{\mu_2}{\gamma_1} - \frac{1}{3}\frac{\mu_3}{\gamma_1\gamma_2} - \tfrac{1}{3}(\gamma_1 + \gamma_2) + \gamma_2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (16).$$

$$\gamma_2 u_2^2 = \frac{\mu_2}{\gamma_2} - \frac{1}{3}\frac{\mu_3}{\gamma_1\gamma_2} - \tfrac{1}{3}(\gamma_1 + \gamma_2) + \gamma_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (17).$$

These equations clearly give u_1^2 and u_2^2, and, therefore, the standard-deviations of the component groups when γ_1 and γ_2 are known.

For brevity, put

$$v_1 = (\gamma_1 u_1)^2, \qquad v_2 = (\gamma_2 u_2)^2,$$

$$p_1 = \gamma_1 + \gamma_2, \qquad p_2 = \gamma_1\gamma_2.$$

Then

$$v_1 = \mu_2 - \tfrac{1}{3}\mu_3/\gamma_2 - \tfrac{1}{3}p_1\gamma_1 + p_2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (18),$$

$$v_2 = \mu_2 - \tfrac{1}{3}\mu_3/\gamma_1 - \tfrac{1}{3}p_1\gamma_2 + p_2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (19),$$

while from (12) and (13) we have

$$2(\gamma_1 v_1 - \gamma_2 v_2) + \frac{v_1^2}{\gamma_1} - \frac{v_2^2}{\gamma_2} = (\gamma_1 - \gamma_2)\{\tfrac{1}{3}p_2 - \tfrac{1}{3}p_1^2 - \tfrac{1}{3}\mu_4/p_2\} \quad . \quad . \quad (20),$$

$$2(\gamma_1^2 v_1 - \gamma_2^2 v_2) + 3(v_1^2 - v_2^2) = (\gamma_1 - \gamma_2)\{\tfrac{2}{5}p_1p_2 - \tfrac{1}{5}p_1^3 - \tfrac{1}{5}\mu_5/p_2\} \quad . \quad (21).$$

We must now substitute (18) and (19) in (20) and (21). We find

$$\gamma_1 v_1 - \gamma_2 v_2 = (\gamma_1 - \gamma_2)\left\{\mu_2 - \tfrac{1}{3}\mu_3\frac{p_1}{p_2} - \tfrac{1}{3}p_1^2 + p_2\right\},$$

$$\gamma_1^2 v_1 - \gamma_2^2 v_2 = (\gamma_1 - \gamma_2)\left\{\mu_2 p_1 - \tfrac{1}{3}\mu_3\frac{p_1^2}{p_2} + \tfrac{1}{3}\mu_3 - \tfrac{1}{3}p_1^3 + \tfrac{4}{3}p_1p_2\right\},$$

$$\frac{v_1^2}{\gamma_1} - \frac{v_2^2}{\gamma_2} = (\gamma_1 - \gamma_2)\left\{-\frac{\mu_2^2}{p_2} + \tfrac{1}{9}\frac{\mu_3^2}{p_2^2} + \tfrac{1}{9}p_1^2 - p_2 - 2\mu_2 + \tfrac{2}{9}\mu_3\frac{p_1}{p_2}\right\},$$

$$v_1^2 - v_2^2 = (\gamma_1 - \gamma_2)\left\{\tfrac{1}{9}\frac{\mu_3^2 p_1}{p_2^2} + \tfrac{1}{9}p_1^3 - \tfrac{2}{3}\frac{\mu_2\mu_3}{p_2} - \tfrac{2}{3}\mu_2 p_1 + \tfrac{2}{9}\mu_3\frac{p_1^2}{p_2} - \tfrac{2}{3}\mu_3 - \tfrac{2}{3}p_1p_2\right\},$$

whence,

$$\frac{\mu_3^2}{p_2^2} - \frac{4\mu_3 p_1}{p_2} - 2p_1^2 + 6p_2 - \frac{9(\mu_2^2 - \frac{1}{3}\mu_4)}{p_2} = 0.$$

$$\frac{5\mu_3^2 p_1}{p_2^2} - 20\mu_3 - 2p_1^3 + 4p_1p_2 - \frac{15(2\mu_2\mu_3 - \frac{1}{5}\mu_5)}{p_2} = 0.$$

Write

$$\lambda_4 = 9\mu_2^2 - 3\mu_4, \qquad \lambda_5 = 30\mu_2\mu_3 - 3\mu_5 \quad \ldots \ldots \quad (22),$$

and put

$$p_3 = p_1 p_2 \quad \ldots \ldots \quad (23),$$

then, multiplying up, the above equations become

$$\mu_3^2 - 4\mu_3 p_3 - 2p_3^2 - \lambda_4 p_2 + 6p_2^3 = 0 \quad \ldots \ldots \quad (24),$$

$$5\mu_3^2 p_3 - 2p_3^3 + 4p_3 p_2^3 - 20\mu_3 p_2^3 - \lambda_5 p_2^2 = 0 \quad \ldots \ldots \quad (25).$$

From these equations let us first find p_3 in terms of p_2. Multiply the first by p_3 and subtract from the second

$$4\mu_3 p_3^2 + p_3(4\mu_3^2 + \lambda_4 p_2 - 2p_2^3) - 20\mu_3 p_2^3 - \lambda_5 p_2^2 = 0 \quad \ldots \quad (26).$$

Multiply (24) by $2\mu_3$ and add to (26) we find

$$2\mu_3^3 + p_3(-4\mu_3^2 + \lambda_4 p_2 - 2p_2^3) - 2\mu_3\lambda_4 p_2 - \lambda_5 p_2^2 - 8\mu_3 p_2^3 = 0,$$

or

$$p_3 = \frac{2\mu_3^3 - 2\mu_3\lambda_4 p_2 - \lambda_5 p_2^2 - 8\mu_3 p_2^3}{4\mu_3^2 - \lambda_4 p_2 + 2p_2^3} \quad \ldots \ldots \quad (27).$$

Hence, so soon as p_2 is known, $p_1 = p_3/p_2$ can be found, and then γ_1 and γ_2 will be the two roots of the quadratic

$$\gamma^2 - p_1\gamma + p_2 = 0 \quad \ldots \ldots \quad (28).$$

Returning to (27), substitute this value of p_3 in (24), and we have an equation containing p_2 only, on which the whole solution of the problem now turns.

This equation is the following one :—

$$24p_2^9 - 28\lambda_4 p_2^7 + 36\mu_3^2 p_2^6 - (24\mu_3\lambda_5 - 10\lambda_4^2)p_2^5 - (148\mu_3^2\lambda_4 + 2\lambda_5^2)p_2^4$$
$$+ (288\mu_3^4 - 12\lambda_4\lambda_5\mu_3 - \lambda_4^3)p_2^3 + (24\mu_3^3\lambda_5 - 7\mu_3^2\lambda_4^2)p_2^2 + 32\mu_3^4\lambda_4 p_2 - 24\mu_3^6 = 0 \quad . \quad (29).$$

(8.) Some remarks may be made on this equation. Since this equation is of an odd order, one real root may always be found. Further, remembering that $\lambda_4 = 9\mu_2^2 - 3\mu_4$ and $\lambda_5 = 30\mu_2\mu_3 - 3\mu_5$, we see that in the case of a normal curve, for which $\mu_4 = 3\mu_2^2$, while μ_3 and $\mu_5 = 0$, all the coefficients of the above equation of the ninth order vanish except the first.

Thus p_2, as we should naturally expect, will be zero. Accordingly, since, with increasing symmetry, the coefficients become small, it will be needful to work their values out to a greater degree of exactness the slighter the degree of asymmetry.

Given that a frequency-curve is compounded of two normal curves, equations (29), (28), (27), (14), (15), (16), and (17) form the complete solution of the problem.

We may throw the whole solution into the following form :—

Stage I.—Find the centroid of the frequency-curve and calculate μ_2, μ_3, μ_4, μ_5, λ_4, and λ_5.

Stage II.—Solve (29) for p_2 and find the corresponding values of p_1 from (27).

Stage III.—Find the positions of the axes of the component normal curves from (28).

Stage IV.—The fractions z_1 and z_2 that the areas of the normal curves are of the area of the frequency-curve are the roots of the quadratic :

$$z^2 - z - \frac{p_2}{p_1^2 - 4p_2} = 0 \qquad (30).$$

Stage V.—Since $\sigma_1/h = \sqrt{v_1}$ and $\sigma_2/h = \sqrt{v_2}$, the standard-deviations are given at once on substituting in (18) and (19).

(9.) The whole method may be illustrated by the following numerical example :—

Breadth of "Forehead" of Crabs.—Professor W. F. R. WELDON has very kindly given me the following statistics from among his measurements on crabs. They are for 1000 individuals from Naples. The abscissæ of the curve are the ratio of "forehead" to body-length, and one unit of abscissa = ·004 of body-length. No. 1 of the abscissæ corresponds to ·580 — ·583 of body-length. The ordinates represent the number of individual crabs corresponding to each set of ratios of forehead to body-length. Thus there was one crab fell into the range ·580 — ·583, three fell into the range ·584 — ·587, five into the range ·588 — ·591, and so on. The average length of animals measured 35 millims., and measurements were recorded to ·1 millim.

Abscissæ.	Ordinates.	Abscissæ.	Ordinates.
1	1	16	74
2	3	17	84
3	5	18	86
4	2	19	96
5	7	20	85
6	10	21	75
7	13	22	47
8	19	23	43
9	20	24	24
10	25	25	19
11	40	26	9
12	31	27	5
13	60	28	0
14	62	29	1
15	54		

This curve is plotted out as the dark continuous line in Plate 1, fig. 1, and is clearly asymmetrical. I proceeded to calculate its first five moments in the analytical method suggested on p. 78 (a), each calculation being made twice independently. I took $h = 1$, and clearly $\alpha = 1000$. The moments were taken about the vertical through the point 0, and were calculated by the aid of Table I. of the powers of the first 30 natural numbers given at the end of this memoir. The following results were obtained :—

$$\begin{aligned}
\mu_1' &= 16\cdot799 \\
\mu_2' &= 304\cdot923 \\
\mu_3' &= 5{,}831\cdot759 \\
\mu_4' &= 116{,}061\cdot435 \\
\mu_5' &= 2{,}385{,}609\cdot719
\end{aligned}$$

μ_1', since $h = 1$, is clearly the distance of the centroid vertical of the frequency-curve from the origin O, *i.e.* $= q$ of p. 77 (ii.).

The moments about this centroid vertical were now calculated by aid of (1), p. 77. There resulted :—

$$\begin{aligned}
\mu_1 &= 0 \\
\mu_2 &= 22\cdot716{,}599 \\
\mu_3 &= -\ 53\cdot874{,}770 \\
\mu_4 &= 1576\cdot533{,}413 \\
\mu_5 &= -\ 9598\cdot313{,}922 \\
\lambda_4 &= -\ 85\cdot205{,}407 \\
\lambda_5 &= -\ 7920\cdot604{,}761
\end{aligned}$$

where λ_4, λ_5 are given in terms of the μ's by (22) of p. 84.

Turning now to the fundamental nonic (29), let it be divided by 24, and written in the form

$$p_2^9 + a_2 p_2^7 + a_3 p_2^6 + a_4 p_2^5 + a_5 p_2^4 + a_6 p_2^3 + a_7 p_2^2 + a_8 p_2 + a_9 = 0.$$

Then the coefficients a_2, a_3 . . . were calculated, and the following values found :—

$$\begin{aligned}
a_2 &= 99\cdot406 \\
a_3 &= 4{,}353\cdot742 \\
a_4 &= -\ 423{,}696 \\
a_5 &= -\ 3{,}702{,}933 \\
a_6 &= 119{,}298{,}911 \\
a_7 &= 1{,}232{,}409{,}400 \\
a_8 &= -\ 957{,}080{,}900 \\
a_9 &= -\ 24{,}451{,}990{,}000
\end{aligned}$$

Put $p_2 = 10\chi$ and divide by 10^9 we then have for the fundamental nonic the following equation, where only three decimal places are retained :—

$$\chi^9 + \cdot994\chi^7 + 4\cdot354\chi^6 - 42\cdot370\chi^5 - 37\cdot029\chi^4 + 119\cdot299\chi^3 + 123\cdot241\chi^2 - 9\cdot571\chi - 24\cdot452 = 0.$$

After a somewhat laborious calculation, the values of STURM's functions $f(\chi)$, $f_1(\chi)$, $f_2(\chi)$, $f_3(\chi)$, $f_4(\chi)$, $f_5(\chi)$, $f_6(\chi)$, $f_7(\chi)$, $f_8(\chi)$, $f_9(\chi)$ were ascertained and gave the following results :—

$f(\infty) = +$	$f(-\infty) = -$
$f_1(\infty) = +$	$f_1(\infty) = +$
$f_2(\infty) = -$	$f_2(\infty) = +$
$f_3(\infty) = -$	$f_3(\infty) = -$
$f_4(\infty) = -$	$f_4(\infty) = +$
$f_5(\infty) = +$	$f_5(\infty) = +$
$f_6(\infty) = +$	$f_6(\infty) = -$
$f_7(\infty) = +$	$f_7(\infty) = +$
$f_8(\infty) = -$	$f_8(\infty) = +$
$f_9(\infty) = -$	$f_9(\infty) = -$
3 changes.	6 changes.

Thus there are $6 - 3 = 3$ real roots.

These three real roots were then localized as follows :—

Two roots between 0 and -1, χ_1 and χ_2.

One root between 0 and 1, χ_3.

As successive approximations, I found :—

To χ_1 :	-1,	$-\cdot89$,	$-\cdot870$,	$-\cdot8757$,
,, χ_2 :	$-\cdot5$,	$-\cdot65$,	$-\cdot670$,	$-\cdot6724$,
,, χ_3 :	$\cdot5$,	$\cdot40$,	$\cdot422$,	$\cdot4170$.

With sufficient accuracy we may then take for the values of p_2 :—

1st solution, $p_2 = -8\cdot757$.

2nd ,, $p_2 = -6\cdot724$.

3rd ,, $p_2 = 4\cdot170$.

Discussion of first solution. $p_2 = -\cdot 8\cdot 757$. p_3 was first calculated from (27) on p. 84, and then $p_1 = p_3/p_2$ found. There resulted : $p_1 = -1\cdot 027$.

The quadratic for γ_1, γ_2, which are here identical with b_1, b_2 (the distances of the centroids of the component probability-curves from the centroid vertical of the frequency-curve), is :—

$$\gamma^2 + 1\cdot 027\gamma - 8\cdot 757 = 0,$$

whence

$$\gamma_1 = -3\cdot 517, \qquad \gamma_2 = 2\cdot 490.$$

The values of z_1 and z_2 were now found from (14) and (15) of p. 82.

$$z_1 = \cdot 4145, \qquad z_2 = \cdot 5855,$$

thus the numbers of individuals in either group are respectively

$$c_1 = 414\cdot 5, \qquad c_2 = 585\cdot 5.$$

The values of the standard-deviations, σ_1 and σ_2, were now determined from (18) and (19), where, since $h = 1$, $v_1 = \sigma_1^2$, and $v_2 = \sigma_2^2$. At the same time the maximum ordinates of the component probability-curves, y_1 and y_2, were found from

$$y_1 = \frac{c_1}{\sqrt{(2\pi)}\,\sigma_1}, \qquad y_2 = \frac{c_2}{\sqrt{(2\pi)}\,\sigma_2}.$$

There resulted

$$\sigma_1 = 4\cdot 4685, \qquad \sigma_2 = 3\cdot 1154.$$
$$y_1 = 37\cdot 008, \qquad y_2 = 74\cdot 976.$$

Thus the 1st solution may be summed up as follows :—

1st Component.	2nd Component.
$c_1 = 414\cdot 5$,	$c_2 = 585\cdot 5$.
$b_1 = -3\cdot 517$,	$b_2 = 2\cdot 490$.
$\sigma_1 = 4\cdot 4685$,	$\sigma_2 = 3\cdot 1154$.
$y_1 = 37\cdot 008$,	$y_2 = 74\cdot 976$.

These two normal curves were now drawn by aid of the Table II., which was calculated afresh for this purpose from the exponential.* These curves are plotted out in fig. 1, and their ordinates added together give the resultant curve. It will be seen that this curve is in remarkably close agreement with the original asymmetrical frequency-curve, an agreement quite as close as we could reasonably expect from the com-

* I have always found it more convenient to work with the standard-deviation than with the probable error or the modulus, in terms of which the error-function is usually tabulated.

parative smallness of the number of individuals dealt with, and the resulting fact that the observation-curve can at best only be an approximation to the true resultant.

2nd Solution.—Precisely similar calculations were undertaken for the value $p_2 = -6{\cdot}724$, and it will, accordingly, be sufficient to cite the final conclusions here.

Quadratic for γ : $\gamma^2 - {\cdot}3412\gamma - 6{\cdot}724 = 0$.

1st Component.	2nd Component.
$c_1 = 467{\cdot}2$,	$c_2 = 532{\cdot}8$.
$b_1 = 2{\cdot}769$,	$b_2 = -2{\cdot}428$.
$\sigma_1 = 2{\cdot}878$,	$\sigma_2 = 4{\cdot}7702$.
$y_1 = 64{\cdot}764$,	$y_2 = 44{\cdot}559$.

These component-curves are drawn in fig. 2, and their ordinates added together. We see that we have again broken up our asymmetrical frequency-curve into two probability-curves, whose sum is a very close approximation to the original curve.

3rd Solution : $p_2 = 4{\cdot}170$.

While the first two solutions have been additive, this solution makes γ_1 and γ_2 ($p_2 = \gamma_1\gamma_2$) of the same sign, or the centroids of the component curves fall both on the same side of the centroid vertical of the frequency-curve. Accordingly the area of one of them must be negative, and the solution promised to be a subtractive one, *i.e.*, to represent the frequency-curve as the difference of two normal curves.

Determining p_3 and then p_1 from (27), we find $p_1 = -3{\cdot}605$; hence

$$\gamma^2 + 3{\cdot}605\gamma + 4{\cdot}170 = 0.$$

The roots of this equation are, however, imaginary. In the case of crabs' foreheads, therefore, we cannot represent the frequency-curve for their forehead lengths as the *difference* of two normal curves.

(10.) So far as the nonic is concerned, our work is now accomplished. Taking the biologist's measurements and assuming them to be the chance distribution of two unequal groups about two different means, then one or other of our solutions is the correct answer. Applying the test of the sixth moment, we find for the observations $\mu_6 = 177{,}004$, while for the first solution it is 188,099 and for the second solution 192,446. According to this test, the first solution is the required one,* but, as we have noticed, the two solutions are themselves much closer together than either to

* The theory of correlation will here, perhaps, confirm this result. Professor WELDON tells me that the first and not the second solution is in good accordance with his other measurements.

the observations (see p. 75). In fact, the contours of the compound-curve for both solutions are very close together, and neither differs more from the observations than most normal curves differ from symmetrical frequency-curves in statistical measurements of this kind.

The contours are so close that, notwithstanding we have demonstrated a *theoretical* uniqueness for the solution of the problem (see p. 72, *et seq.*), we see that, from the standpoint of practical statistics, it is possible for the given material to be broken up into more than one pair of normal curves. Thus the problem indeed becomes somewhat arbitrary—at any rate till the asymmetry of the frequency-curve becomes much more marked than is the case with that of the foreheads of Naples crabs. Indeed, although the method adopted leads to only two solutions, it is quite possible that pairs of component normal curves might be tentatively found lying in the neighbourhood of those determined by the above solutions, which would give resultant-curves fairly close to the frequency-curve. Professor WELDON had, indeed, found by repeated trials one such solution, but this solution differs widely in the third and higher moments from the observations; it cannot, therefore, be considered to have the same justification as those given by the present theory. Granted that the original observations represent a mixture of two species varying about their mean according to exact normal curves, our method gives *two solutions, and two only.* Without correlated measurements, it might be difficult to discriminate between these solutions—at any rate from the standpoint of practical statistics. The perhaps over-fine theoretical test of the sixth moment decides for the first solution.

II.—*The Dissection of Symmetrical Frequency-Curves.*

(11.) Another important case of the dissection of a frequency-curve can arise, when the frequency-curve, without being asymmetrical, still consists of the sum or difference of two components, *i.e.*, when the means about which the component groups are distributed are identical. This case is all the more interesting and important, as it is not unlikely to occur in statistical investigations, and the symmetry of the frequency-curve is then in itself likely to lead the statistician to believe that he is dealing with an example of the normal frequency-curve. It seems to me that without very strong grounds for belief in the homogeneity of any statistical material, we ought not to be satisfied by its representation by the ordinary normal curve, simply because our results are symmetrical and fit the normal curve fairly well. We ought first to ascertain whether or not they would fit still better the sum or difference of two normal curves. This, at any rate, is a first stage to demonstrating the homogeneity of our material, although possibly our test for two may fail, not because our material is homogeneous, but because its heterogeneity is multiple rather than double.*

* Symmetry might arise in the case of compound frequency-curves, even without identity of the means of the components. In this case, for two components we should have for different means,

We will now modify the results of our previous investigation to suit the case of an asymmetrical frequency-curve which has arisen from the superposition of two normal-curves having the same axis. In this case if we unite, $b_1 = b_2 = 0$, $v_1 = \sigma_1/h$ $(= u_1\gamma_1)$, $v_2 = \sigma_2/h$ $(= u_2\gamma_2)$ in Equations (8) to (13) we have (9), (11) and (13) identically satisfied, and (8), (10), and (12) become

$$z_1 + z_2 = 1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (31),$$

equality of component group-totals and of their standard-deviations. This equality seems less likely than equality of means and divergence of totals and standard-deviations. Should it exist, however, we fall back on a sub-case of the general case we have already dealt with. We need only, in Equations (8)–(13), put $z_1 = z_2$, $\gamma_1 = -\gamma_2$, $u_1 = u_2$, and we have

$$z_1 = z_2 = \tfrac{1}{2}, \quad \gamma_1^2 (1 + u_1^2) = \mu_2, \quad \gamma_1^4 (1 + 6u_1^2 + 3u_1^4) = \mu_4,$$

whence

$$\gamma_1 = \left\{\frac{3\mu_2^2 - \mu_4}{2}\right\}^{\frac{1}{4}} \quad u_1 = \left\{\frac{\sqrt{(2)}\mu_2}{\sqrt{(3\mu_2^2 - \mu_4)}} - 1\right\}^{\frac{1}{2}},$$

or,

$$c_1 = c_2 = \tfrac{1}{2}a,$$

$$b_1 = -b_2 = h\left\{\frac{3\mu_2^2 - \mu_4}{2}\right\}^{\frac{1}{4}},$$

$$\sigma_1 = \sigma_2 = h\left\{\sqrt{\left(\frac{3\mu_2^2 - \mu_4}{2}\right)}\left(\frac{\sqrt{(2)}\mu_2}{\sqrt{(3\mu_2^2 - \mu_4)}} - 1\right)\right\}^{\frac{1}{2}}$$

The possibility of the solution clearly depends on $3\mu_2^2$ being greater than μ_4.

The following is an example of this special case. Mr. MERRIMAN gives some results for American target practice, on page 14 of his Text Book on Least Squares. He does not seem to have noticed that the resulting-curve is very far from a normal-curve. I find that for these observations

$$\begin{aligned} \mu'_1 &= 6{\cdot}482 & \mu_1 &= 0 \\ \mu'_2 &= 44{\cdot}502 & \mu_2 &= 2{\cdot}486 \\ \mu'_3 &= 320{\cdot}582 & \mu_3 &= {\cdot}104 \\ \mu'_4 &= 2405{\cdot}094 & \mu_4 &= 15{\cdot}793. \end{aligned}$$

The smallness of μ_3 indicates general symmetry; assuming then that the shots were fired in two groups *with equal precision*, I find $c_1 = c_2$ and $b_1 = -b_2$ almost exactly.

We have accordingly

$$b_1 = -b_2 = 1{\cdot}082,$$

$$\sigma_1 = \sigma_2 = 1{\cdot}147,$$

[For the 1000 shots as a whole $\sigma = 1{\cdot}577$.]

Allowing for a uniform error of defective sighting amounting to ·482, we find a compound-curve fitting closely Mr. MERRIMAN's figure, and indicating that the gun was aimed *at the centres nearly of divisions* 5 *and* 7, *and not at that of* 6. Six was possibly white, 5 and 7 black. Like results of course would arise from a change of sighting about midfiring.

$$z_1 v_1^2 + z_2 v_2^2 = \mu_2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (32),$$

$$z_1 v_1^4 + z_2 v_2^4 = \tfrac{1}{3}\mu_4 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (33).$$

Clearly we require one more equation. At first sight it might seem that a fourth equation would come readily, from the fact that the mid-ordinate m of the frequency-curve is the sum of the mid-ordinates of the component probability-curves.

This leads to

$$\frac{c_1}{\sqrt{(2\pi)}\,\sigma_1} + \frac{c_2}{\sqrt{(2\pi)}\,\sigma_2} = m,$$

or

$$\frac{z_1}{\sqrt{v_1}} + \frac{z_2}{\sqrt{v_2}} = m' \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (34),$$

if

$$m' = \sqrt{(2\pi)}\, mh/\alpha.$$

But besides the disadvantage of throwing our solution back on the correctness with which we may have observed measurements of one size only, namely, the mean, the result of eliminating between (31)–(34) leads to an equation of the eighth order. To avoid this, it seems easier, as well as more accurate,* to take as the fourth equation that obtained from the sixth moment.

Let $\mu_6 \alpha h^6$ be the sixth moment of the given frequency-curve about its axis of symmetry, then†

$$\mu_6 \alpha h^6 = 15\sigma_1^6 c_1 + 15\sigma_2^6 c_2,$$

or,

$$z_1 v_1^6 + z_2 v_2^6 = \tfrac{1}{15}\,\mu_6 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (35).$$

The solution of (31), (32), (33), and (35) is easy.

Eliminating z_2 we have, writing $w_1 = v_1^2$, $w_2 = v_2^2$,

$$\begin{aligned} z_1(w_1 - w_2) &= \mu_2 - w_2, \\ z_1 w_1 (w_1 - w_2) &= \tfrac{1}{3}\mu_4 - \mu_2 w_2, \\ z_1 w_1^2 (w_1 - w_2) &= \tfrac{1}{15}\mu_6 - \tfrac{1}{3}\mu_4 w_2 \end{aligned}$$

whence

$$w_1 = \frac{\frac{1}{15}\mu_6 - \frac{1}{3}\mu_4 w_2}{\frac{1}{3}\mu_4 - \mu_2 w_2} = \frac{\frac{1}{3}\mu_4 - \mu_2 w_2}{\mu_2 - w_2}.$$

* Because our equation then depends on *all* the observations.

† Generally, if M_{2r} be the $2r$ moment of a probability-curve about its axis

$$M_{2r} = (2r-1)\,\sigma^2 M_{2r-2},$$

or,

$$M_{2r} = (2r-1)(2r-3)\ldots 5.3.1\sigma^{2r}c.$$

Thus

$$(\mu_4 - 3\mu_2^2)\, w_2^2 + (\mu_4\mu_2 - \tfrac{1}{5}\,\mu_6)\, w_2 - (\tfrac{1}{3}\,\mu_4^2 - \tfrac{1}{5}\,\mu_2\mu_6) = 0.$$

The two roots of this quadratic are clearly w_1 and w_2, so that the complete solution is

$$c_1 = \alpha\,\frac{\mu_2 - w_2}{w_1 - w_2}, \qquad c_2 = \alpha\,\frac{w_1 - \mu_2}{w_1 - w_2},$$

$$\sigma_1 = h\,\sqrt{w_1}, \qquad \sigma_2 = h\,\sqrt{w_2},$$

where w_1 and w_2 are roots of

$$(\mu_4 - 3\mu_2^2)\, w^2 + (\mu_2\mu_4 - \tfrac{1}{5}\mu_6)\, w - (\tfrac{1}{3}\mu_4^2 - \tfrac{1}{5}\mu_2\mu_6) = 0 \quad . \quad . \quad . \quad (36)$$

(12.) Now we may note several general points about these equations.

Let w_1 be the greater root, then if

(i.) μ_2 lie between w_1 and w_2, c_1 and c_2 are both positive, or the frequency-curve is the *sum* of two normal curves.

(ii.) $\mu_2 > w_1$, c_1 is positive and c_2 negative, or the greater component group is positive, we have then a real *difference* solution.

(iii.) $\mu_2 < w_2$, c_1 is negative and c_2 is positive, or again the greater component group is positive, or we have a real *difference* solution.

Obviously if $\mu_4 = 3\mu_2^2$, and $\mu_6 = 5\mu_2\mu_4$, the coefficients of the quadratic (36) all become zero, but these are just the conditions which would be satisfied if the frequency-curve were a true normal curve. This gives for all practical purposes a very sufficient test of whether a given symmetrical frequency-curve is a true normal curve

If μ_4 be not equal to $3\mu_2^2$, and μ_6 be not equal to $5\mu_2\mu_4$, then we have no right to assume that a symmetrical frequency-curve refers to homogeneous material. We must then investigate whether a better result cannot be obtained by treating it as two superposed normal curves having the same axis.

The quantities

$$\epsilon_1 = \frac{\mu_4 - 3\mu_2^2}{3\mu_2^2}, \quad \text{and} \quad \epsilon_2 = \frac{\mu_6 - 5\mu_2\mu_4}{5\mu_2^3},$$

I propose to call the *excess* and *defect* of the frequency-curve. The excess measures the excess of one-third of the fourth moment over the square of the second moment; the defect measures the defect of the fourth moment from one-fifth the ratio of the sixth moment to the second moment.* Here "excess" and "defect" are used in the algebraic sense, and may take either sign. They appear to be a good

* The introduction of the factor $1/\mu_2^2$ into both excess and defect is to preserve a *relative* as distinguished from an absolute measure of divergence.

measure for practical purposes of the divergence of a given symmetrical frequency-curve from the normal type.

We may now express the quadratic (36) in terms of ϵ_1 and ϵ_2, and analyze the results according to the character of the excess and defect.

The quadratic becomes

$$3\epsilon_1\left(\frac{w}{\mu_2}\right)^2 - \epsilon_2\frac{w}{\mu_2} + \epsilon_2 - 3\epsilon_1(1+\epsilon_1) = 0.$$

This gives

$$\frac{w}{\mu_2} = \frac{\epsilon_2 \pm \sqrt{\{(\epsilon_2 - 6\epsilon_1)^2 + 36\epsilon_1^3\}}}{6\epsilon_1} \quad \text{. . . . . (37).}$$

We have the following cases:

(i.) ϵ_1 and ϵ_2 both positive. Then the values of w are both real, but they must also be both positive, otherwise σ_1 and σ_2 would not be real. It is necessary, therefore, that

$$\epsilon_2 > \sqrt{\{(\epsilon_2 - 6\epsilon_1)^2 + 36\epsilon_1^3\}},$$

or

$$\epsilon_2 < 3\epsilon_1(1+\epsilon_1).$$

(ii.) ϵ_1 and ϵ_2 both negative. Then w will be real if, when

$$\sqrt{(-\epsilon_1)} < 1,$$

$(-\epsilon_2)$ does not lie between

$$6(-\epsilon_1)\{1+\sqrt{(-\epsilon_1)}\}$$

and

$$6(-\epsilon_1)\{1-\sqrt{(-\epsilon_1)}\}.$$

If

$$\sqrt{(-\epsilon_1)} > 1,$$

then we must have

$$(-\epsilon_2) > 6(-\epsilon_1)\{1+\sqrt{(-\epsilon_1)}\}.$$

Further, in order that w may have both values positive, we must have

$$(-\epsilon_2) > \{-\epsilon_2 - 6(-\epsilon_1)\}^2 - 36(-\epsilon_1)^3,$$

or

$$(-\epsilon_2) > 3(-\epsilon_1)\{1-(-\epsilon_1)\}.$$

This latter condition is clearly satisfied if

$$\sqrt{(-\epsilon_1)} > 1.$$

On the other hand, if

$$\sqrt{(-\epsilon_1)} < 1,$$

it is easy to see that

$$3(-\epsilon_1)\{1-(-\epsilon_1)\}$$

is less than

$$6(-\epsilon_1)\{1-\sqrt{(-\epsilon_1)}\}.$$

Hence, our final conditions are

$$\sqrt{(-\epsilon_1)} > 1,$$

then

$$(-\epsilon_2) > 6(-\epsilon_1)\{1+\sqrt{(-\epsilon_1)}\};$$

but if

$$\sqrt{(-\epsilon_1)} < 1,$$

then either

$$(-\epsilon_2) > 6(-\epsilon_1)\{1+\sqrt{(-\epsilon_1)}\},$$

or it must lie between

$$3(-\epsilon_1)\{1-(-\epsilon_1)\}$$

and

$$6(-\epsilon_1)\{1+\sqrt{(-\epsilon_1)}\}.$$

(iii.) ϵ_1 positive and ϵ_2 negative; if the values of w are real, one must be negative, and therefore the solution impossible.

(iv.) ϵ_1 negative and ϵ_2 positive; if the values of w are real, one must be negative, and therefore the solution impossible.

Thus we conclude:

If the excess and defect are not zero, the frequency-curve, although symmetrical, is not normal. If the excess and defect are of opposite signs, then the frequency-curve cannot be broken up into the sum or difference of *two* normal curves with common axis. The frequency-curve, if compounded of normal-curves at all, is of a higher and more complex character. If the excess and defect are of the same sign, then, provided certain relations hold between the numerical values of the excess and defect given in (i.) and (ii.) above, there is a real solution of the equation which resolves the frequency-curve into two components.

(13.) I propose to illustrate this discussion by the consideration of a numerical example. Professor WELDON has kindly complied with my request for the numerical details of the most symmetrical curve deduced from his measurements of Naples crabs by placing the following statistics for a shell measurement—No. 4 of his series —at my disposal. The resultant-curve and the corresponding normal curve are pictured in fig. 3 (Plate 3). Clearly, from the ordinary statistician's standpoint, we could not expect a more symmetrical result, or a closer graphical agreement, with the normal curve. But is this a real or merely an apparent agreement? The answer is, as we shall see, vital for the interpretation to be put on Professor WELDON's results.

CRAB MEASUREMENTS. No. 4. (Total Number of Crabs = 999.)

Abscissæ.	Ordinates (1 unit = 1 crab).	Abscissæ.	Ordinates (1 unit = 1 crab).
1	1	11	126
2	3	12	82
3	5	13	72
4	11	14	41
5	40	15	28
6	55	16	8
7	98	17	7
8	121	18	0
9	152	19	0
10	147	20	2

The first six moments were calculated exactly as in the previous case of § 9, by aid of Table I., except that α now equals 999, and we go a stage further to μ'_6 and μ_6. h equals unity as before. We have

$$\begin{aligned} \mu_1' &= 9{\cdot}684{,}684 & \mu_1 &= 0 \\ \mu_2' &= 101{\cdot}3022 & \mu_2 &= 7{\cdot}5092 \\ \mu_3' &= 1{,}129{\cdot}9971 & \mu_3 &= 3{\cdot}4751 \\ \mu_4' &= 13{,}334{\cdot}0710 & \mu_4 &= 176{\cdot}7280 \\ \mu_5' &= 165{,}488{\cdot}8438 & \mu_5 &= 271{\cdot}6007 \\ \mu_6' &= 2{,}150{,}845{\cdot}6867 & \mu_6 &= 7{,}919{\cdot}2781 \end{aligned}$$

These results give for the position of the centroid $d = \mu_1' = 9{\cdot}6847$, and for the standard-deviation $\sigma = \sqrt{\mu_2} = 2{\cdot}7403$. This gives the modulus $3{\cdot}874$, and the central ordinate of the normal curve $145{\cdot}44$. The modulus, as calculated from the mean error, is $3{\cdot}8634$, so that the agreement is very close. The normal curve in fig. 3 is constructed from the values $d = 9{\cdot}6847$, $\sigma = 2{\cdot}7403$, and $y_0 = 145{\cdot}44$ by aid of Table II.

The following additional quantities were now calculated :—

$$\begin{aligned} \mu_4 - 3\mu_2^2 &= 7{\cdot}5637 \\ \epsilon_1 &= {\cdot}044{,}712 \\ \lambda_4 &= -22{\cdot}6911 \\ \mu_5 - 10\mu_2\mu_3 &= 10{\cdot}6485 \\ \lambda_5 &= -31{\cdot}9455 \\ \mu_6 - 5\mu_2\mu_4 &= 1283{\cdot}8486 \\ \epsilon_2 &= {\cdot}606{,}45 \end{aligned}$$

If we had a perfect probability-curve, μ_3, μ_5, $\mu_4 - 3\mu_2^2$, and $\mu_6 - 5\mu_2\mu_4$ ought to be zero. This, of course, we should not expect in any actual set of observations, but the comparative smallness of μ_3, μ_5, λ_4, λ_5, ϵ_1, and ϵ_2 shows a very fair approximation to the symmetry of the normal curve in these results.

Since $\epsilon_2 > 3\epsilon_1(1 + \epsilon_1)$, we see that the roots (37) of our p. 94 are both positive, and accordingly it is possible to break up the observation-curve into two normal curves with coincident axes.

Calculating the two values of w we have

$$\frac{w_1}{\mu_2} = 3\cdot50971, \qquad \frac{w_2}{\mu_2} = 1\cdot01148,$$

whence from p. 93 :

$$c_1 = -\alpha \times \cdot0046, \qquad c_2 = \alpha \times 1\cdot0046,$$
$$\sigma_1 = \sqrt{(\mu_2 \times 3\cdot50971)}, \qquad \sigma_2 = \sqrt{(\mu_2 \times 1\cdot01148)}$$

or

$$c_1 = -5,^* \qquad c_2 = 1004,$$
$$\sigma_1 = 5\cdot134, \qquad \sigma_2 = 2\cdot756.$$

For all practical purposes the second group gives the normal curve ($c = 999$, $\sigma = 2\cdot740$) of the set of observations; that a half per cent. of Crabs have been removed by selection about the same mean is not large enough to be significant in measurements of the kind we are here dealing with. So far, then, we may say that No. 4 of Professor WELDON'S measurements cannot be treated as the sum or difference of two normal curves having their axes coincident with any substantial improvement on the normal curve peculiar to the original group.

(14.) Hitherto we have used "Crab Measurements No. 4" to illustrate the dissection of symmetrical frequency-curves, but a little consideration shows at once that this judging of symmetry by the eye is very likely to be fallacious, and No. 4 may, after all, break up into two normal curves with non-coincident axes. Should these two curves correspond to practically the same groups as in the case of the "Foreheads," then we shall have demonstrated that the asymmetry of that frequency-curve is in all probability due to a mixture of two families in the Naples Crabs and not a result of differentiation going on in one homogeneous species. The *apparent* symmetry of No. 4 weighs nothing in the balance, as may be readily tested by adding together two normal curves with not widely divergent axes or totals.

What we have been investigating, therefore, in § 13 is really only the special case in which the method of our first investigation would fail, owing to the coincidence of the axes of the component normal curves—a coincidence which is improbable *à priori*.

I, therefore, proceeded to form the nonic for No. 4, a result which requires only the values of μ_3, λ_4, and λ_5 already given.†

The nonic being

$$p_2^9 + a_2 p_2^7 + a_3 p_2^6 + a_4 p_2^5 + a_5 p_2^4 + a_6 p_2^3 + a_7 p_2^2 + a_8 p_2 + a_9 = 0,$$

* The nearest whole number is here taken for the Crabs in each group.

† The arithmetic throughout was of course of a most laborious character.

the coefficients were—

$$
\begin{aligned}
a_2 &= \quad 26{\cdot}47295. \\
a_3 &= \quad 18{\cdot}11448. \\
a_4 &= \quad 325{\cdot}54964639. \\
a_5 &= \quad 1604{\cdot}777825{,}114. \\
a_6 &= \quad 977{\cdot}342{,}6614. \\
a_7 &= -\,3154{\cdot}2006888. \\
a_8 &= -\,4412{\cdot}284{,}2437. \\
a_9 &= -\,1761{\cdot}180374.
\end{aligned}
$$

Writing $p_2 = -\chi$, we have for the nonic $f(\chi)$ and its first derived function* $f_1(\chi)$ the following expressions—

$$
\begin{aligned}
f(\chi) = \chi^9 &+ 26{\cdot}472{,}95\chi^7 - 18{\cdot}114{,}48\chi^6 \\
&+ 325{\cdot}549{,}646\chi^5 - 1604{\cdot}777{,}825\chi^4 \\
&+ 977{\cdot}342{,}661\chi^3 + 3154{\cdot}200{,}689\chi^2 - 4412{\cdot}284{,}244\chi \\
&+ 1761{\cdot}180{,}374 = 0,
\end{aligned}
$$

and

$$
\begin{aligned}
f_1(\chi) = \chi^8 &+ 20{\cdot}590{,}07\chi^6 - 12{\cdot}076{,}32\chi^5 \\
&+ 180{\cdot}860{,}915\chi^4 - 713{\cdot}234{,}589\chi^3 \\
&+ 325{\cdot}780{,}887\chi^2 + 700{\cdot}933{,}486\chi \\
&- 490{\cdot}253{,}805.
\end{aligned}
$$

The STURM's functions were now formed, and with the following results—

	$\chi = \infty$.	$\chi = 0$.	$\chi = -\infty$.
$f(\chi) =$	+	+	−
$f_1(\chi) =$	+	−	+
$f_2(\chi) =$	−	−	+
$f_3(\chi) =$	+	+	+
$f_4(\chi) =$	+	−	−
$f_5(\chi) =$	+	−	+
$f_6(\chi) =$	+	−	−
$f_7(\chi) =$	−	−	−
$f_8(\chi) =$	−	+	+
$f_9(\chi) =$	+	+	+
Totals	4 changes	4 changes	5 changes.

Thus the nonic has one root of χ between 0 and $-\infty$, and no roots between 0 and $+\infty$. In other words it has 8 imaginary roots and only 1 real one.

* Divided by the factor 9.

This root was now localized. Putting $p_2 = \frac{1}{10}/\chi'$ in the original nonic, I easily found χ' to lie between 0 and 1, then between ·15 and ·16, and by a succession of approximations to be ·1533, and finally ·15326.

Thus

$$p_2 = 1{\cdot}5326.$$

p_3 was then ascertained from equation (27) of p. 84, and finally $p_1 = p_3/p_2$ was found to be 2·17245. The quadratic (28) for γ was then :

$$\gamma^2 - 2{\cdot}17245\gamma + 1{\cdot}5326 = 0,$$

which has both its roots imaginary.

Thus, considerably to my surprise, but greatly to my satisfaction, it was demonstrated that there is no solution whatever of the problem of breaking up the curve of No. 4 measurements into two normal components.

All nine roots of the fundamental nonic lead to imaginary solutions of the problem. The best and most accurate representation of No. 4 is the normal curve of fig. 3.

The result of this investigation seems to me most important. Professor WELDON'S material *is homogeneous*, and the asymmetry of the "forehead" curve points to a real differentiation in that organ, and not to a mixture of two families having been dredged up.

On the other hand, I cannot think that for the problem of evolution the dissection of the most symmetrical curve given by the measurements is unnecessary. There will always be the problem : Is the material homogeneous and a true evolution going on, or is the material a mixture ? To throw the solution on the judgment of the eye in examining the graphical results is, I feel certain, quite futile.

Whenever in measuring a series of organs the results give an asymmetrical curve, we must accordingly proceed as follows :—

Stage (*i*).—Break up this asymmetrical curve into components ; if there are several solutions, the theory of correlation or the test of the sixth moment will, perhaps, enable us to say which is the most satisfactory.

Stage (*ii*).—Endeavour to break up the most symmetrical curve ; if it cannot be broken up, either into normal components with non-coincident axes or normal components with coincident axes, the material is homogeneous and the asymmetrical curve points to a true differentiation in the organ to which it refers. If, on the other hand, the most symmetrical frequency-curve does break up, then if the numbers in its component groups be the same (or practically the same) as in those corresponding to the asymmetrical curve, we are really dealing with a mixture of heterogeneous material, and we shall have ascertained the proportions of the mixture. If the numbers should not be the same, then we cannot assert that we have a mixture, but we have found a case of differentiation in both organs at the same time.*

* BERTILLON has found a double-humped frequency-curve for the height of the inhabitants of the

These stages seem to represent the mathematical treatment of this portion of the problem of evolution.

(15.) Although the nonic corresponding to "Crabs No. 4," has no real negative root, I found on tracing its value for values of p_2 between 0 and -2, that near $p_2 = -\cdot 82$ it reached a minimum value of about 199 as compared with about 1761 at 0 < 1254 at -2. Here then was, as it were, a tendency towards a root, and the question occurred to me whether this "tendency" in any way corresponded to the groups into which the "foreheads" were differentiated. I therefore investigated the root of the first derived function of the nonic lying about $-\cdot 82$, and found it to be $-\cdot 8497$. This led to p_1 from equation (27) being $-5\cdot 2521$, whence

$$\gamma^2 + 5\cdot 2521\gamma - \cdot 8497 = 0,$$

or

$$\gamma_1 = \cdot 15705, \qquad \gamma_2 = -5\cdot 40915.$$

Whence nearly

$$z_1 = \cdot 972, \qquad z_2 = \cdot 028,$$

or the numbers in the two groups are

$$c_1 = 971 \quad \text{and} \quad c_2 = 28.$$

Clearly even this "tendency to a root" in no way fits either solution of the "forehead" case, and No. 4 measurements neither break up, nor have they even a tendency to break up, in the same manner as the "foreheads." Since the nonic must always have a "tendency" to two real roots at a time, we may note that the other root to which it may be said to tend, or for which $f(p_2)$ is a minimum, lies between $-\cdot 9$ and -1, and is just as insignificant as that investigated above. We may say that not only is the material of No. 4 homogeneous, but it has not even a "tendency" towards heterogeneity.

III.

(16.) The object of the present paper being solely to illustrate a general method for the reduction frequency-curves to normal types, and not a biological investigation, it might suffice to stop at this point, when the rules for the reduction of symmetrical and asymmetrical curves have been given and illustrated. But it must be remembered that the method depends upon the solution of a nonic, and that the variety presented

department of the Doubs. Mr. BATESON has found a double-humped curve for the claspers of Earwigs. Without the investigation of measurements of another organ, it seems impossible to say whether the inhabitants of the Doubs, as BERTILLON supposes, are a mixture of races, or Mr. BATESON's earwigs were really homogeneous. In either case our methods of investigation would show the proportions belonging to each group of the mixture, or to each group of the differentiating species.

by the roots of this equation suggests very considerable divergences and peculiarities as likely to arise, when a considerable number of frequency-curves are dealt with.

The discussion of the case of Crabs must not be taken as indicating that the incidents of this case will be generally true for other groups of biological measurements, until a very great variety of such groups of measurements have been mathematically analyzed.

In order to throw more light on the general question, I have added the following analysis for the case of Prawns, the measurements for which were kindly placed at my disposal by Mr. H. THOMPSON, who has been making elaborate measurements of 1,000 specimens in the Zoological Laboratory of University College, London.

Palæmon serratus.—Measurements in 998 ♀ specimens (adult) from penultimate to hindmost tooth on the carapace.

Measurements reduced to thousandths of body length.	Number of specimens.	Measurements reduced to thousandths of body length.	Number of specimens.
27	1	49	25
28	0	50	17
29	0	51	11
30	0	52	8
31	1	53	4
32	0	54	1
33	3	55	0
34	3	56	0
35	4	57	1
36	11	58	1
37	24	59	0
38	38	60	0
39	56	61	0
40	80	62	0
41	105	63	0
42	121	64	0
43	117	65	1
44	108	66	0
45	77	67	0
46	69	68	0
47	62	69	1
48	48		

The novel and somewhat remarkable feature in these results are the "giants" at 65 and 69. To neglect these giants, as in some degree anomalous, would, no doubt be convenient, so far as the analysis is concerned, and would lead to a simpler reduction of the group. They have, however, been retained as among the data given to me, and their presence affords an interesting illustration of the various singularities which may arise in the solution of the fundamental nonic.

(17.) The curve (see fig. 4) given by the observed numbers will be at once seen to be distinctly asymmetrical. Adopting the carapace length 31 as the origin of coordinates, and using the same notation as before, we have the following results :—*

$$\begin{array}{llll} \mu'_1 = d\,(=q) = & 16\cdot191{,}382{,}8 & \mu_1 = 0 & \\ \mu'_2 = & 276\cdot277{,}555 & \mu_2 = & 14\cdot116{,}678{,}13 \\ \mu'_3 = & 4{,}963\cdot876{,}753{,}5 & \mu_3 = & 33\cdot424{,}02673 \\ \mu'_4 = & 94{,}386\cdot734{,}469 & \mu_4 = & 1{,}288\cdot640{,}094{,}26 \\ \mu'_5 = & 1{,}920{,}725\cdot520{,}040 & \mu_5 = & 16{,}752\cdot563{,}9961 \end{array}$$

$$\lambda_4 = -\ \ 2072\cdot394{,}903$$

$$\lambda_5 = -\ 36{,}102\cdot605{,}1706.$$

The standard-deviation of the group as a whole is given by $\sigma = \sqrt{\mu_2}$, or

$$\sigma = 3\cdot7572.$$

The mean error† obtained from σ . . . . . $= 2\cdot9978$

" " " directly . . . . $= 2\cdot8776$.

(In the case of the "foreheads" of Crabs, the mean error from σ was 3·8028, and directly 4·4087. This divergence between the mean error, as found practically from second and first moments, is a very good test of the asymmetry of the frequency-curve. In the very symmetrical measurements of "Crabs No. 4," the modulus, as calculated from the standard-deviation and from the mean error, had the near values 3·874 and 3·863.)

The curve obtained from the observations as a single group (*i.e.*, $d = 16\cdot1914$ and $\sigma = 3\cdot7572$) is given in fig. 4 (Plate 4).

Taking $\chi = \frac{1}{10}p_2$ we have for the fundamental nonic and its first differential

$$\begin{array}{ll} f(\chi) = \chi^9 & f'(\chi) = 9\chi^8 \\ \quad +\ \ 24\cdot177{,}940{,}535\chi^7 & \quad +\ \ 169\cdot245{,}583{,}743\chi^6 \\ \quad +\ \ \ \ 1\cdot675{,}748{,}344\chi^6 & \quad +\ \ \ 10\cdot054{,}490{,}066\chi^5 \\ \quad +\ 299\cdot620{,}303{,}770\chi^5 & \quad +\ 1498\cdot101{,}518{,}851\chi^4 \\ \quad -\ 943\cdot393{,}909{,}962\chi^4 & \quad -\ 3773\cdot575{,}639{,}850\chi^3 \\ \quad -\ 864\cdot540{,}147{,}350\chi^3 & \quad -\ 2593\cdot620{,}442{,}052\chi^2 \\ \quad -\ 274\cdot750{,}163{,}918\chi^2 & \quad -\ \ 549\cdot500{,}327{,}835\chi \\ \quad -\ \ 34\cdot486{,}278{,}563\chi & \quad -\ \ \ 34\cdot486{,}278{,}563 \\ \quad -\ \ \ \ 1\cdot394{,}286{,}418 = 0. & \end{array}$$

* These results were calculated to a higher degree of accuracy than in the case of the Crabs, a result rendered necessary by the apparent sensitiveness of the roots in this case to a slight change in the value of the coefficients of the nonic.

† Mean error is here used, not in GAUSS's sense, but in the sense of arithmetically mean error, $= \cdot7979\,\sigma$ theoretically.

Clearly there is only one positive root. This was found to be

$$\chi = 2\cdot5868658.$$

This gave

$$p_2 = 25\cdot868{,}658,$$

whence I found

$$p_1 = 9\cdot669{,}970.$$

Consequently the roots of

$$\gamma^2 - p_1\gamma + p_2 = 0$$

were imaginary and no solution involving the difference of two normal components was possible.

The next stage was to find the negative roots. These were easily demonstrated to lie between 0 and 1, and then it was shown that the value of $f(\chi)$ only changed sign twice between these values. Thus the nonic was proved, without calculating STURM's functions, to have only three real roots. The two negative roots are :—

$$\chi_1 = -\cdot154{,}481{,}14$$

and

$$\chi_2 = -\cdot078{,}262{,}95.$$

These roots lead to the following solutions :—

(A.) *First additive Solution for Carapace of Prawns.*

$$p_2 = -\ 1\cdot544{,}8114,$$
$$p_1 = 26\cdot758{,}0108,$$

$$\gamma_1 = -\cdot057{,}6086, \qquad \gamma_2 = 26\cdot815{,}6194,$$
$$z_1 = \cdot997{,}856, \qquad z_2 = \cdot002{,}144.$$

1st Component.	2nd Component.
$c_1 = 995{,}860,$	$c_2 = 2\cdot140,$
$b_1 = -\cdot057{,}6086,$	$b_2 = 26\cdot815{,}6194,$
$\sigma_1 = 3\cdot5595,$	$\sigma_2 = 5\cdot7626\sqrt{-1}$
$y_1 = 111\cdot6142.$	$y_2 =$ imaginary.

(B.) *Second additive Solution for Carapace of Prawns.*

$$p_2 = -\cdot782{,}6295,$$
$$p_1 = 5\cdot163{,}5907,$$

$$\gamma_1 = -\cdot147{,}3614, \qquad \gamma_2 = 5\cdot310{,}9521,$$
$$z_1 = \cdot973{,}0024, \qquad z_2 = \cdot026{,}9976.$$

1st Component.	2nd Component.
$c_1 = 971{\cdot}0564,$	$c_2 = 26{\cdot}9436,$
$b_1 = -\ {\cdot}147{,}3614,$	$b_2 = 5{\cdot}310{,}9521,$
$\sigma_1 = 3{\cdot}389{,}672,$	$\sigma_2 = 8{\cdot}932{,}996,$
$y_1 = 114{\cdot}28698.$	$y_2 = 1{\cdot}203{,}280.$

To these solutions we may add :—

(C.) *Parameters of Normal Curve deduced from entire group of observations.*

$$d = 16{\cdot}191{,}383,$$
$$c = 998,$$
$$\sigma = 3{\cdot}7572,$$
$$y = 105{\cdot}968{,}04.$$

(D.) *Parameters of Normal Curve deduced by excluding two "giants" from observations.*

$$d = 16{\cdot}14357 \quad (b = -{\cdot}04781),$$
$$c = 996,$$
$$\sigma = 3{\cdot}6051,$$
$$y = 110{\cdot}21786.$$

The curves corresponding to (A), (B), (C), and (D) as well as the observation-curve are given in figs. 4 and 5, and I shall now proceed to discuss several important points with regard to them.

(18.) The first point to be noted is the existence of the dwarf, carapace 27, and the giants, carapaces 65 and 69.

The normal curve has a standard-deviation 3·7572, and the mean carapace being about 43, we have no less than *three* measurements deviating by more than four times the standard-deviation from the mean ; two of them, indeed, differ by nearly six times the standard-deviation from the mean. We might expect three such deviations of over four times the standard-deviation to occur in the measurement of 50,000 Prawns, but they are extremely improbable in the measurement of 1000 prawns. That two should occur in the measurement of 1000 Prawns, with a deviation six times the standard, is so improbable that it ought to lead us to reject the normal curve as a representation of the measurements. We are either dealing with a mixed population of Prawns, or possibly there are a few deformed individuals amid a normal population.*

There is another point, however, in which the normal curve, based on the total

* I exclude the possibility of any serious error of measurement, having reason to believe in the great care with which the determinations were made.

observations, diverges considerably from the observational result, namely (see fig. 4), in the defect of carapaces about 45. This defect largely contributes to the asymmetrical appearance of the curve. I felt very confident that by neglecting the eccentric group of "giants" I could find two components, whose resultant would fit the curve of observation as closely as the resultant-curves found for the similar case of the forehead of Crabs. I was peculiarly interested, however, in ascertaining whether the method of resolution by aid of the nonic would pay more attention to the outlying giants or to the less improbable defect of individuals about 45. I even imagined that out of the nine possible solutions some might be solutions for the giants and some for the 45 defect. As a matter of fact, the two solutions which have any meaning are entirely taken up with the very improbable outlying eccentricities of the observations. These eccentricities must first be removed from the observations before the method will be of service in resolving the asymmetry of the bulk of the observation-curve.

The method in which the nonic deals with the abnormalities is very characteristic, and I venture to think highly suggestive.

In fig. 4 the normal curve excluding the two giants is given. It fits the observation-curve, as far as *appearances* go, slightly better than the true normal curve. But the first solution of the nonic tells us not to absolutely reject the giants. It gives us two components, the first of which fits the observations slightly better than the normal curve D (giants excluded). It has practically the same area (995·86 as compared with 996), a slightly less standard-deviation (3·5595 as compared with 3·6051), and consequently an increased maximum ordinate. This, with a slightly shifted axis, gives a somewhat better fit. In addition to this first component we have a second component with an area of 2·140, and a mean of 70 for the carapace. This component corresponds closely to the *two* giants with a mean of 67. It has, however, an *imaginary* standard-deviation. Clearly the addition of two to the first component, if distributed really, could make no sensible change in its appearance, and we may then sum up the first solution of the nonic in the following words :—

It does not absolutely reject the two giants, but places an imaginary distribution of 2·14 in their neighbourhood, and thus obtains for the other component and the resultant-curve (which must be practically identical with it) a better approach to the observation-curve than if the giants had been rejected.

It would appear, therefore, that our method of dissection offers, by means of small components with imaginary distributions, a means of obtaining better results than by simply rejecting (or, perhaps, even weighting) anomalous observations.

The second method by which the nonic attempts to account for the eccentricities of these carapace measurements, is by mixing a small population of about 2·7 per cent. of giants with the normal population. These giants have a mean carapace of 48·5, while the rest of the population has a mean of only 43. This population of giants, however, has a very large standard-deviation, *i.e.*, 8·9330 as compared with the 3·3897 of the rest of the population. It is clear that this population of giants is an *unstable* population, *i.e.*, a very small disturbance would largely change its centre. That it accounts for and covers the dwarf and two giant anomalies is clear, and the resultant-curve, based on the addition of the two components, is a fairly close approach to the observation-curve—far closer indeed than that provided by the first solution, and a great advance on the normal-curve C, resulting from the observations as a whole (see fig. 5). I am inclined, accordingly, to suspect that the family of Prawns was *not homogeneous*, but contained between 2 and 3 per cent. of a giant population with a large standard deviation. Possibly the theory of correlations may settle whether this is the real state of the case, or whether the anomalies referred to ought to be rejected and a new investigation made to dissect the asymmetrical curve for the carapaces when the outlying parts, which control the nonic at present, are removed.

The investigation of this case, however, with all the observations included, shows the great variety of solutions which may be suggested by the dissection of various anomalous and asymmetrical frequency-curves.

TABLE I.—Powers of the Natural Numbers.

Powers.					
First.	Second.	Third.	Fourth.	Fifth.	Sixth.
1	1	1	1	1	1
2	4	8	16	32	64
3	9	27	81	243	729
4	16	64	256	1,024	4,096
5	25	125	625	3,125	15,625
6	36	216	1,296	7,776	46,656
7	49	343	2,401	16,807	117,649
8	64	512	4,096	32,768	262,144
9	81	729	6,561	59,049	531,441
10	100	1,000	10,000	100,000	1,000,000
11	121	1,331	14,641	161,051	1,771,561
12	144	1,728	20,736	248,832	2,985,984
13	169	2,197	28,561	371,293	4,826,809
14	196	2,744	38,416	537,824	7,529,536
15	225	3,375	50,625	759,375	11,390,625
16	256	4,096	65,536	1,048,576	16,777,216
17	289	4,913	83,521	1,419,857	24,137,569
18	324	5,832	104,976	1,889,568	34,012,224
19	361	6,859	130,321	2,476,099	47,045,881
20	400	8,000	160,000	3,200,000	64,000,000
21	441	9,261	194,481	4,084,101	85,766,121
22	484	10,648	234,256	5,153,632	113,379,904
23	529	12,167	279,841	6,436,343	148,035,889
24	576	13,824	331,776	7,962,624	191,102,976
25	625	15,625	390,625	9,765,625	244,140,625
26	676	17,576	456,976	11,881,376	308,915,776
27	729	19,683	531,441	14,348,907	387,420,489
28	784	21,952	614,656	17,210,368	481,890,304
29	841	24,389	707,281	20,511,149	594,823,321
30	900	27,000	810,000	24,300,000	729,000,000

TABLE II.—Ordinates of Normal Curve.

D = Deviation. S = Standard Deviation.

F = Frequency. P = Maximum Frequency $\left(\frac{c}{\sigma\sqrt{2\pi}}\right)$.

D/S.	F/P.	D/S.	F/P.
0	1	1·6	·2780
0·1	·9950	1·7	·2357
0·2	·9802	1·8	·1979
0·3	·9560	1·9	·1645
0·4	·9231	2	·1353
0·5	·8825	2·2	·0889
0·6	·8353	2·4	·0561
0·7	·7827	2·6	·0340
0·8	·7262	2·8	·0198
0·9	·6670	3	·0111
1	·6065	3·2	·0060
1·1	·5467	3·4	·0031
1·2	·4868	3·6	·0015
1·3	·4286	3·8	·0007
1·4	·3753	4	·0003
1·5	·3246	5	·000,004

[NOTE, added February 10, 1894.—(1.) The importance of breaking up asymmetrical frequency-curves into normal components has been recognized for a long time by anthropologists and biologists. Attempts at a solution have been made by R. LIVI, '*Sulla statura degli Italiani*,' Firenze, 1883 (see also '*Archivio per l'Antropologia e l'Etnologia*,' vol. 13, Firenze, 1883, and '*Annali di Statistica*,' vol. 8, 1883, pp. 119–56). Also by O. AMMON in his recent work '*Die natürliche Auslese beim Menschen*,' Jena, 1893. These attempts can hardly be looked upon as serious. Professor LEXIS and Dr. VENN have pointed out that the curve of deaths for each year for 1000 persons born in the same year—the true mortality-curve—is also in all probability a compound curve.

Since writing the above memoir I have succeeded in resolving this mortality-curve into components which are not, however, all of the normal type, but become, as we approach infantile mortality, of the skew form (see p. 74 above).

O. AMMON, in the volume cited above, endeavours to demonstrate an evolution in the length-breadth index of the skull of South-Germans since primitive times. He does this by comparison of the index as obtained from measurements on skulls from the Row-Graves and on modern skulls. He has not, however, noticed that the frequency-curve for Row-Grave skulls is *asymmetrical.* I have succeeded in breaking it up into two components, one of which practically coincides in mean and standard-deviation with the frequency-curve for the skulls of modern South-

Germans. In other words, the Row-Graves contain a mixed population, one element of which corresponds closely to the modern South-German population. AMMON's statement, therefore, that an evolution has taken place in this particular skull index appears to fall to the ground. The whole problem of the compound nature of skull frequency-curves, both in England and Germany, is a very interesting and difficult one, and I do not wish at present to anticipate results, which I hope when my investigations are complete to publish as a whole. The above may suffice to indicate the range of problems to which a resolution of asymmetrical frequency-curves into normal components may be applied.

(2.) With regard to the method adopted in the memoir itself, I am very conscious of the defects under which it suffers—the laborious character of the arithmetic involved, and the question of what may be the probable error of the solution obtained by the method of higher moments. But I had to deal with the fact that the problem is one which urgently needed a solution in the case of both economic and biological statistics. Better solutions than mine may be ultimately found, but although more than one mathematically trained statistician has for some time recognized the importance of the problem, no solution, so far as I am aware, has hitherto been forthcoming.

With regard to the amount of error introduced by the use of higher moments, a word may be said. I have not been able to work out the general problem suggested to me by Professor GEORGE DARWIN: "Given the probable error of every ordinate of a frequency-curve, what are the probable errors of the elements of the two normal curves into which it may be dissected?"

I can, however, indicate the sort of differences which are likely to occur in results based on high or on low moments. Suppose the distribution of an organ in a group of animals actually does follow a normal frequency-curve. Then it is obvious that in selecting 1000 of these animals at random and measuring their organs, an error of the same magnitude in the frequency of an organ of a given size is more likely to occur in a size near the mean than in a size far from the mean. Now a low moment pays greater attention than a high moment to an error in the frequency near the mean and less attention than a high moment to one far off. In other words, a frequency-curve calculated from low moments fits best near the centre; one calculated from high moments fits best near the tails of the observation-curve. The problem is accordingly the following: an error in frequency near the tail is not as probable as an equal error in frequency near the mean; but if it does occur a high moment pays much more attention to it than a low moment; on the other hand, the low moment pays more attention than the high moment to more probable errors in frequency. Which tendency on the whole will prevail?

Turning to the result in the foot-note, p. 92, we have for the $2r^{\text{th}}$ moment—

$$M_{2r} = (2r - 1)(2r - 3) \ldots 5.3.1\, \sigma^{2r} c,$$

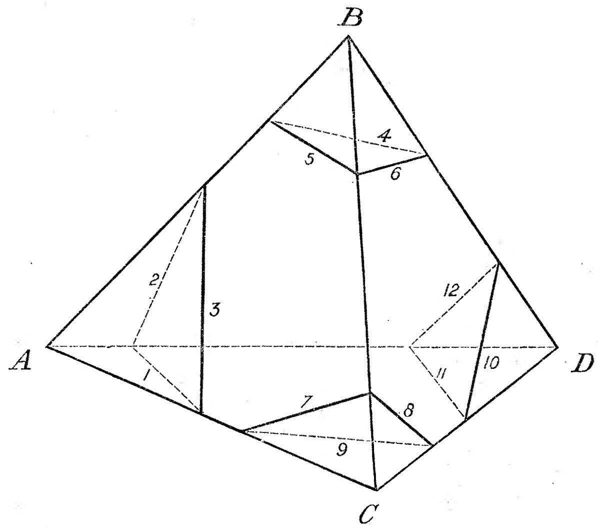
and

$$M_{2r} = S\,(x^{2r} y\, \delta x).$$

Then the equation (A) takes the form

$$xyzu = (x - a\mathrm{T})(y - b\mathrm{T})(z - c\mathrm{T})(u - d\mathrm{T}) \quad . \quad . \quad . \quad . \quad . \quad (\mathrm{D}),$$

and it represents, besides the plane T, the cubic surface passing through the twelve straight lines, which are represented in the annexed figure, as well as three other straight lines which are not represented in the figure.



The equations of the lines may be written as follows :—

$$\begin{array}{cccc}
 & \left.\begin{array}{l} x = a\mathrm{T} \\ y = 0 \end{array}\right\}(1) & \left.\begin{array}{l} x = a\mathrm{T} \\ z = 0 \end{array}\right\}(2) & \left.\begin{array}{l} x = a\mathrm{T} \\ u = 0 \end{array}\right\}(3) \\
\left.\begin{array}{l} y = b\mathrm{T} \\ x = 0 \end{array}\right\}(6) & & \left.\begin{array}{l} y = b\mathrm{T} \\ z = 0 \end{array}\right\}(4) & \left.\begin{array}{l} y = b\mathrm{T} \\ u = 0 \end{array}\right\}(5) \\
\left.\begin{array}{l} z = c\mathrm{T} \\ x = 0 \end{array}\right\}(8) & \left.\begin{array}{l} z = c\mathrm{T} \\ y = 0 \end{array}\right\}(9) & & \left.\begin{array}{l} z = c\mathrm{T} \\ u = 0 \end{array}\right\}(7) \\
\left.\begin{array}{l} u = d\mathrm{T} \\ x = 0 \end{array}\right\}(10) & \left.\begin{array}{l} u = d\mathrm{T} \\ y = 0 \end{array}\right\}(11) & \left.\begin{array}{l} u = d\mathrm{T} \\ z = 0 \end{array}\right\}(12) &
\end{array}$$

and

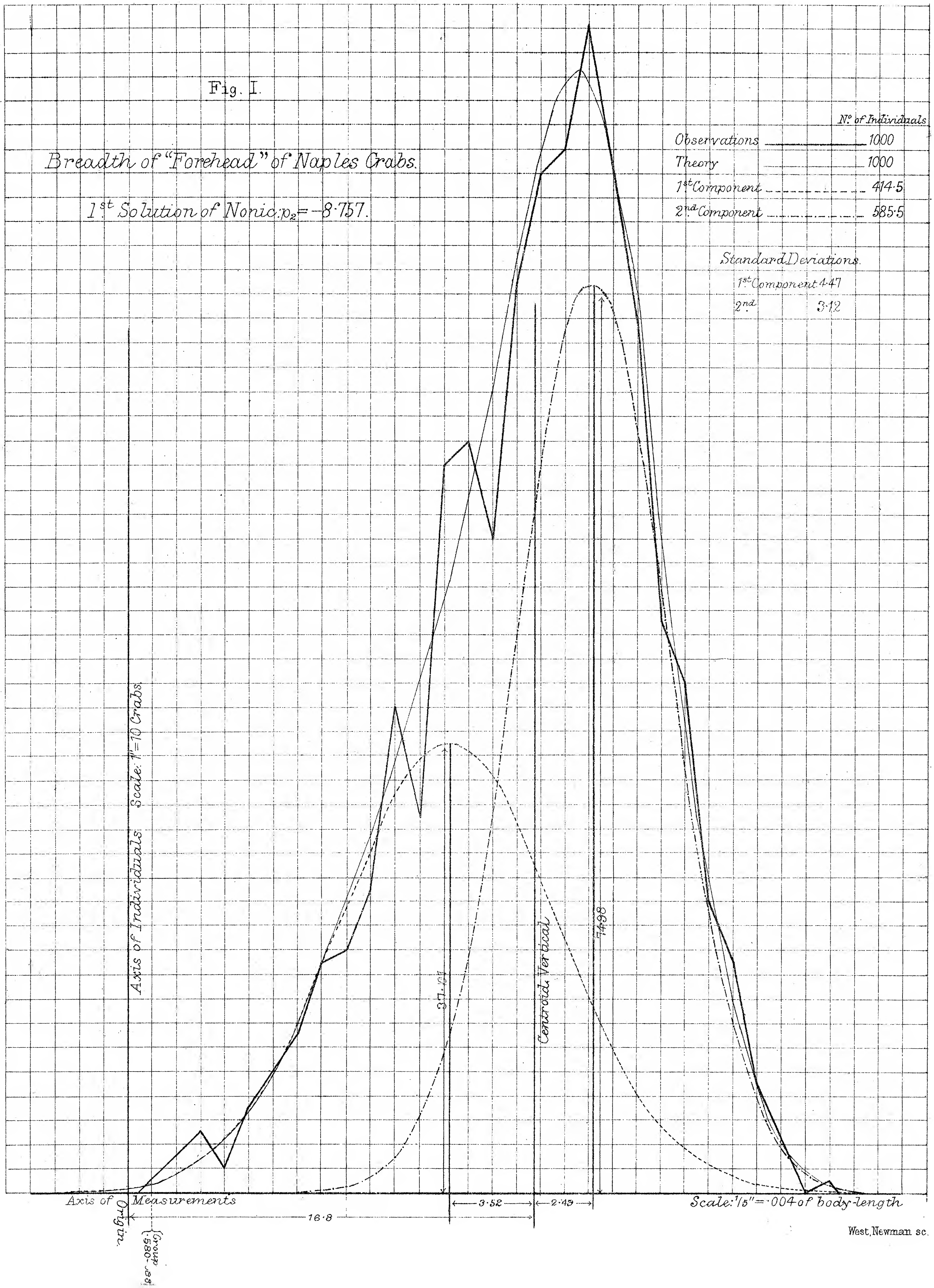
$$\left.\begin{array}{l} \dfrac{x}{a} + \dfrac{u}{d} - \mathrm{T} \\[2ex] \dfrac{y}{b} + \dfrac{z}{c} - \mathrm{T} \end{array}\right\} (13),$$

which meets (3), (4), (9), and (10) ;

moments respectively, I notice the following values for the standard-deviation of "Crabs No. 4," as calculated from the second, fourth, and sixth moments—

$$\sigma_2 = 2{\cdot}74,$$
$$\sigma_4 = 2{\cdot}77,$$
$$\sigma_6 = 2{\cdot}84.$$

Practically, it would be difficult to say which of these results gives the best fitting theoretical curve. For statistics of this kind they are sensibly the same. Thus, till another method of attacking the problem of the resolution of asymmetrical frequency-curves is propounded, I think there is not sufficient evidence against the use of higher moments to lead us to discard a method based upon them as essentially likely to lead to large errors.—K. P.]



West, Newman sc.

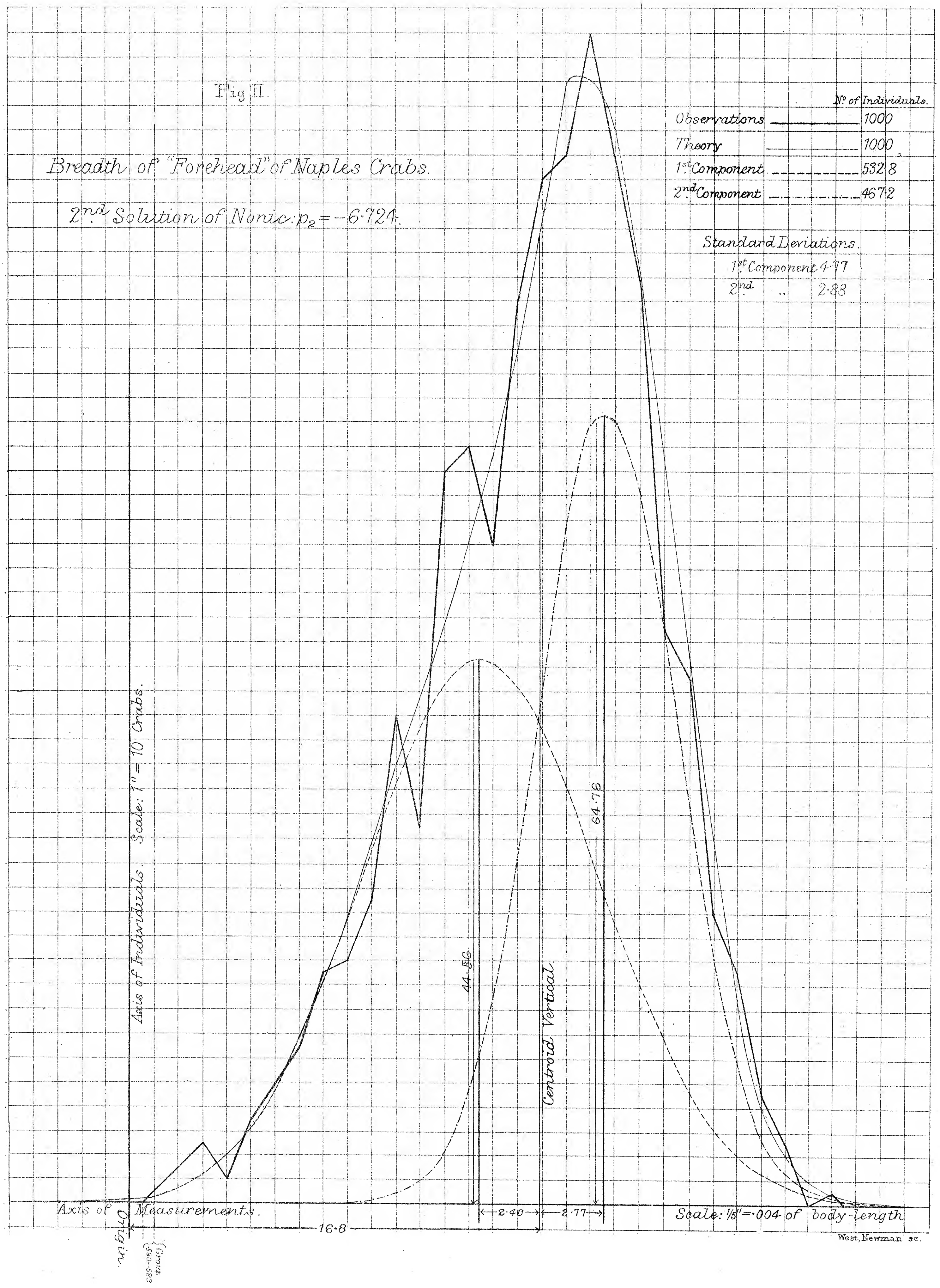
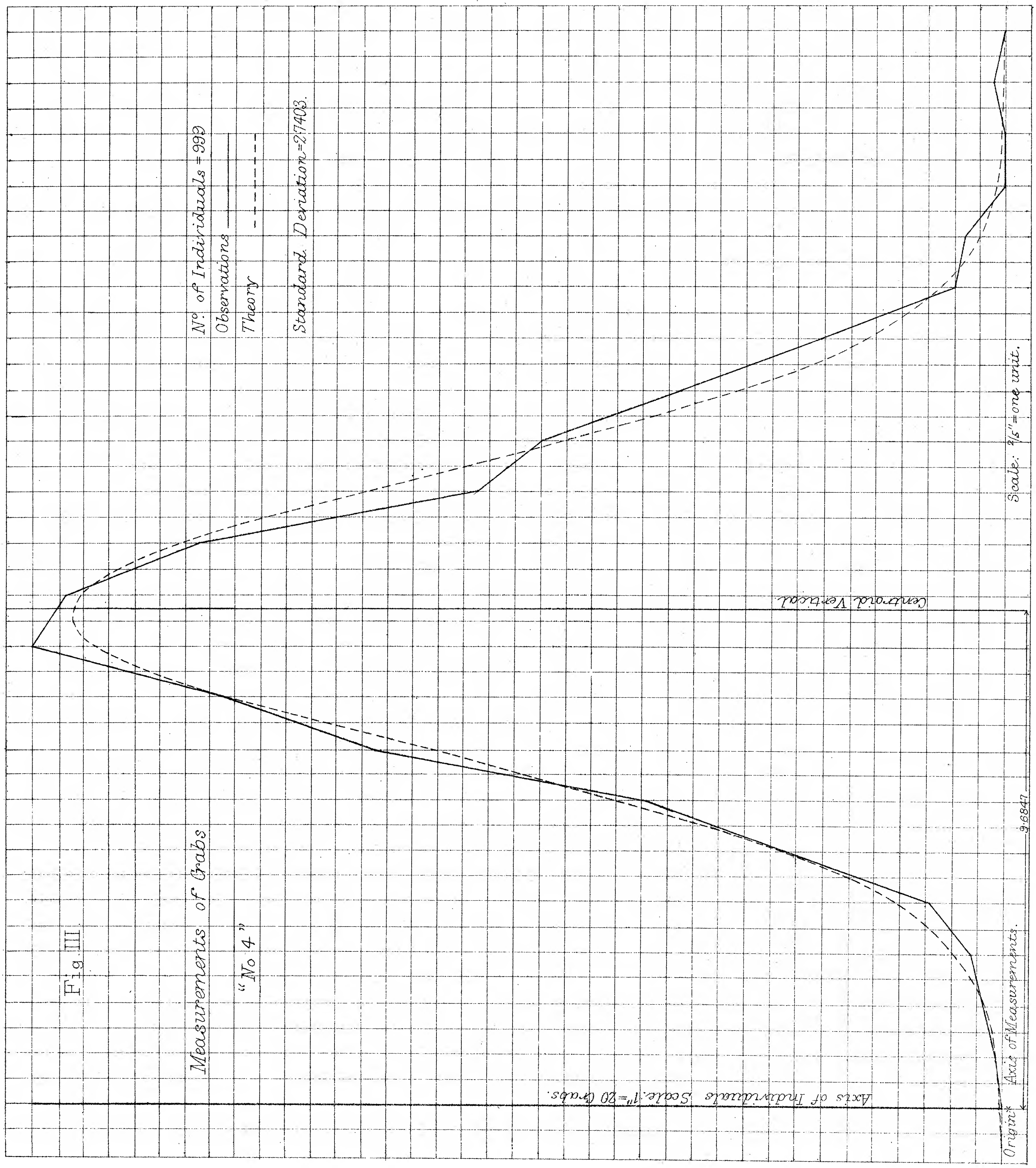


Fig II.
Breadth of "Forehead" of Naples Crabs.
2nd Solution of Nonic: $p_2 = -6{\cdot}724$.
No. of Individuals.
Observations 1000
Theory 1000
1st Component 532·8
2nd Component 467·2
Standard Deviations.
1st Component 4·77
2nd .. 2·83
Axis of Individuals Scale: 1″ = 10 Crabs.
44·56
64·76
Centroid Vertical
Axis of Measurements.
Origin.
Group 580–583
16·8
2·48
2·77
Scale: ⅕″ = ·004 of body-length
West, Newman sc.



West, Newman sc.

Pearson.

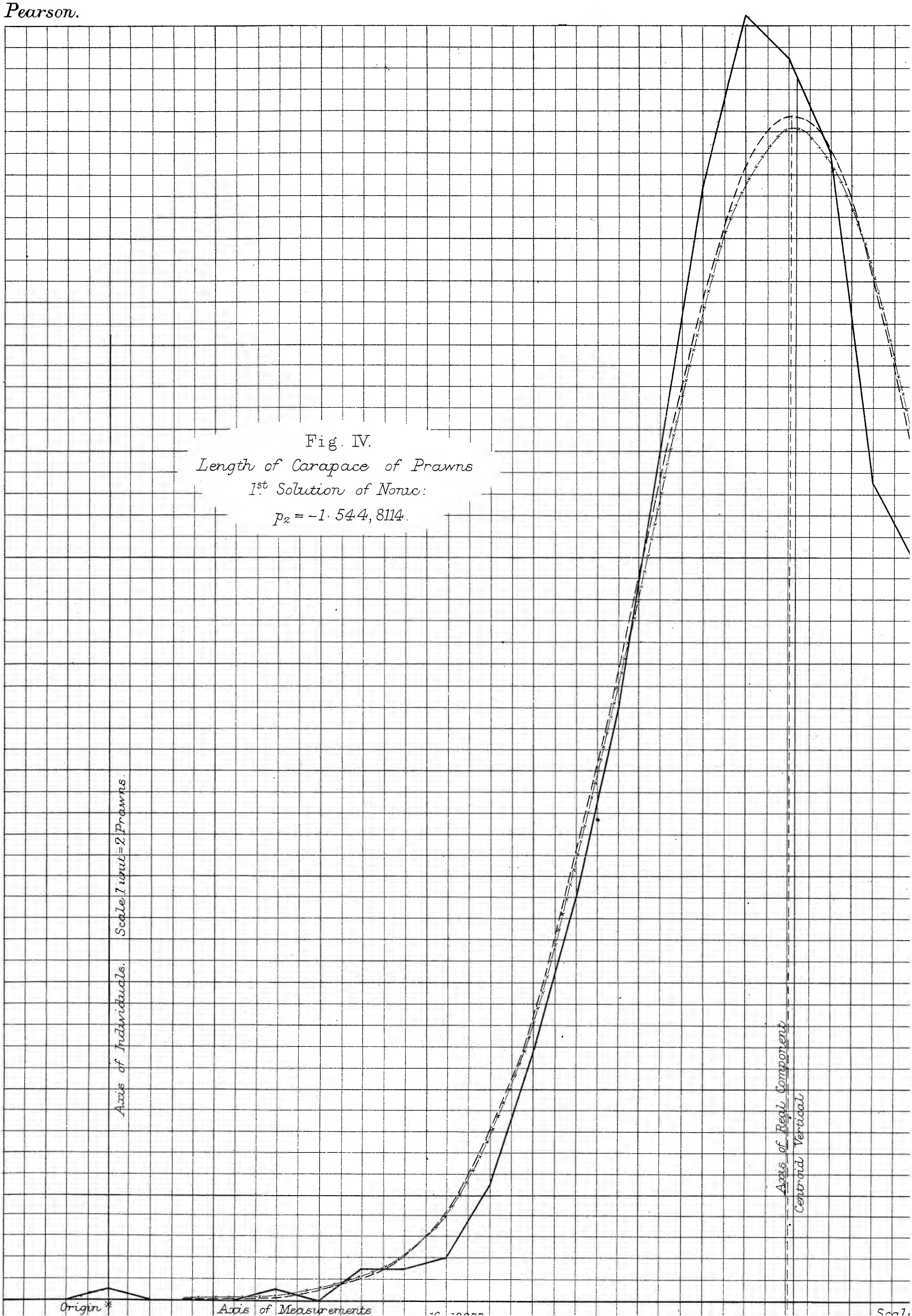
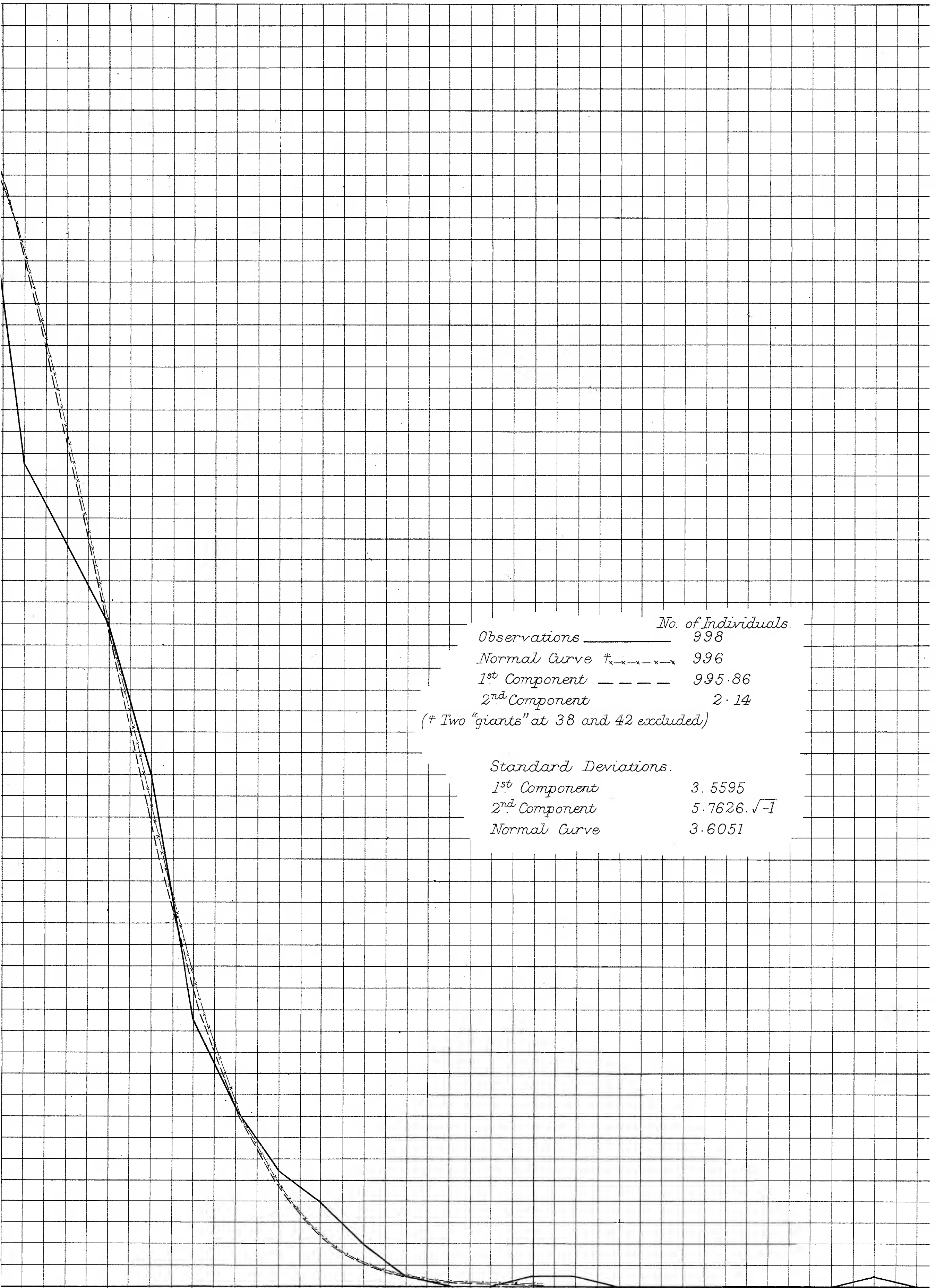
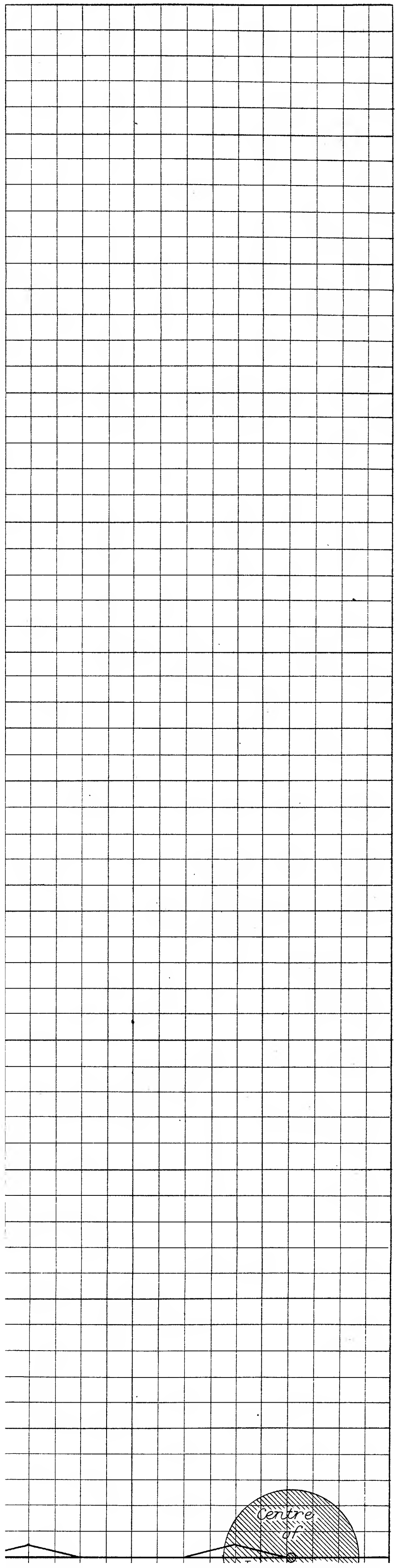
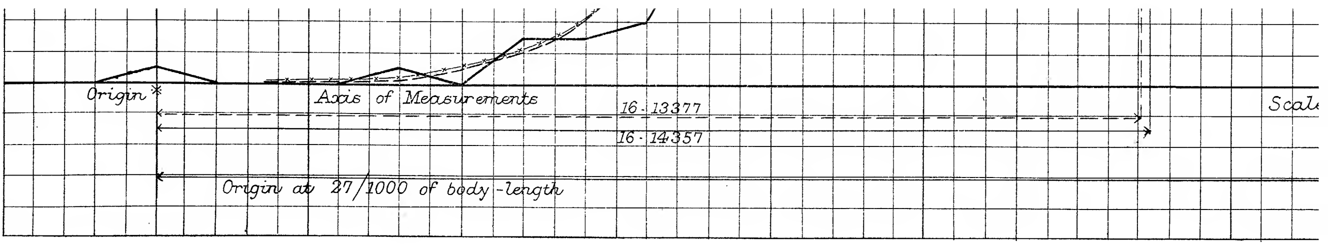


Fig. IV.
Length of Carapace of Prawns
1st Solution of Nonic:
$p_2 = -1 \cdot 544, 8114$.
Axis of Individuals.
Scale 1 unit = 2 Prawns.
Axis of Real Component
Centroid Vertical
Origin
Axis of Measurements

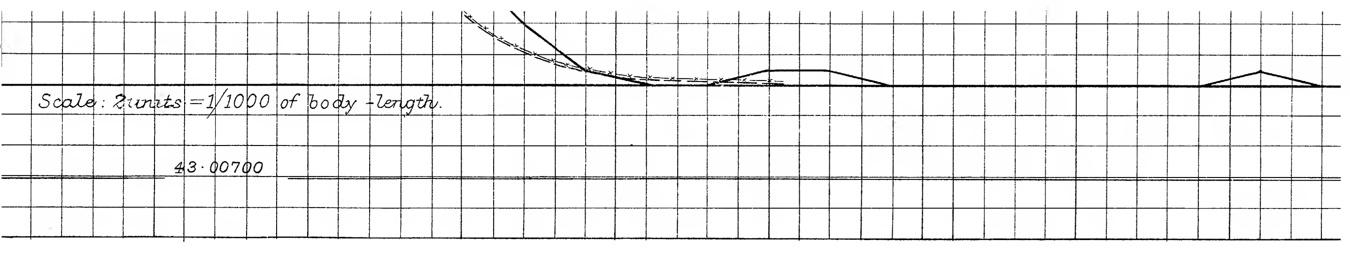


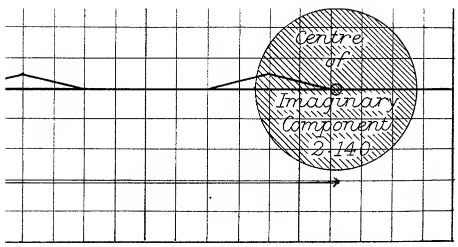
Scale: 2 units = 1/1000 of body-length


Centre
of



Origin
Axis of Measurements
16 · 13377
16 · 14357
Origin at 27/1000 of body-length
Scal


Scale: 2 units = 1/1000 of body-length.
43·00700



West, Newman, lith.

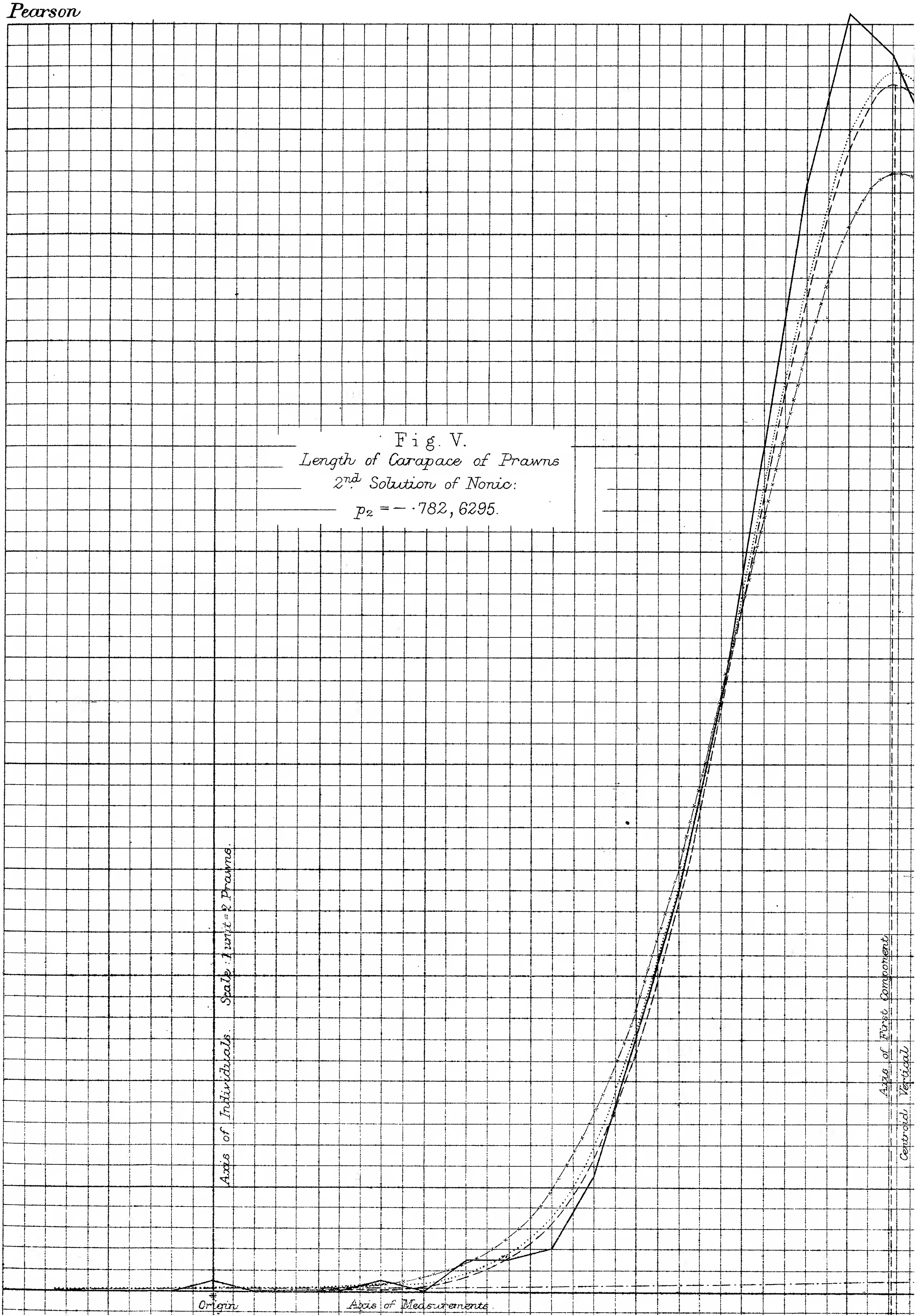
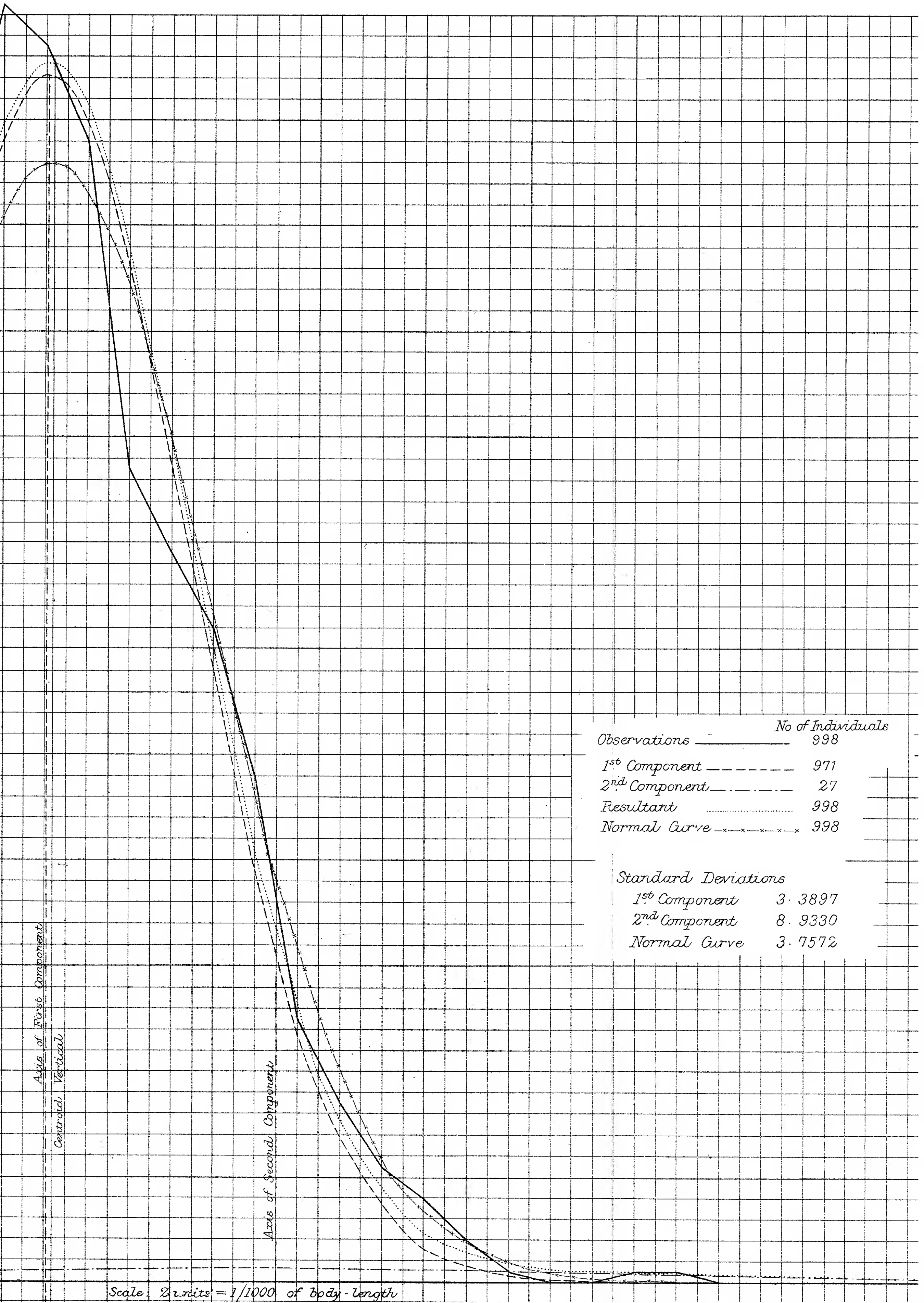
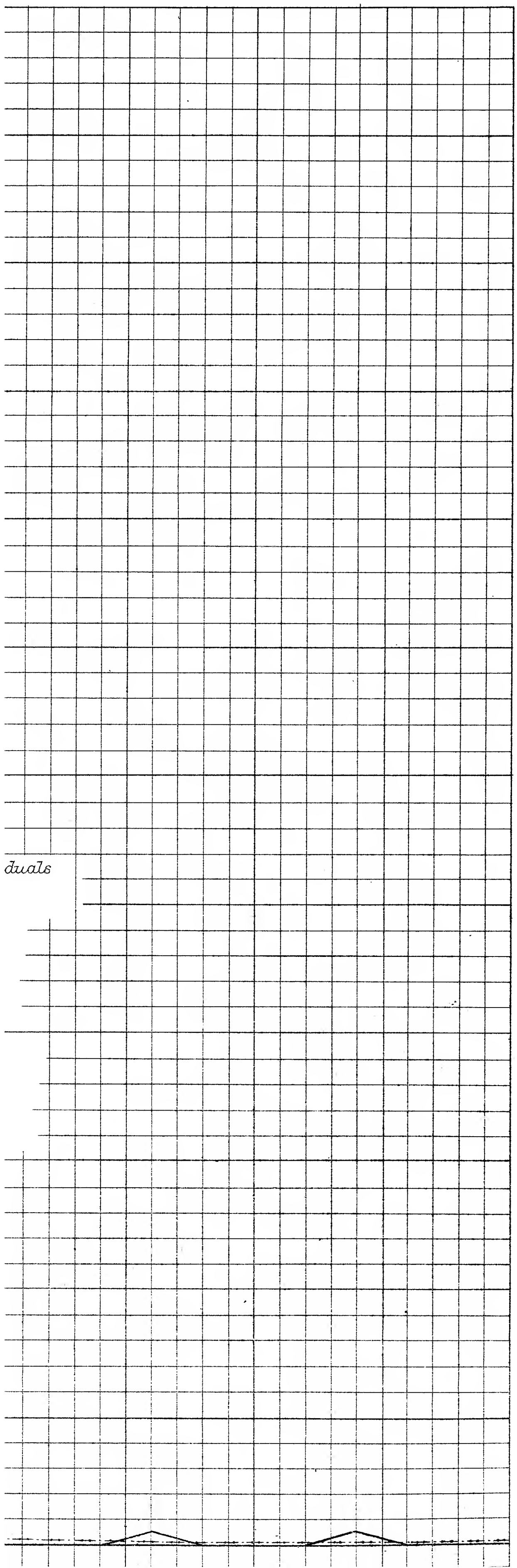
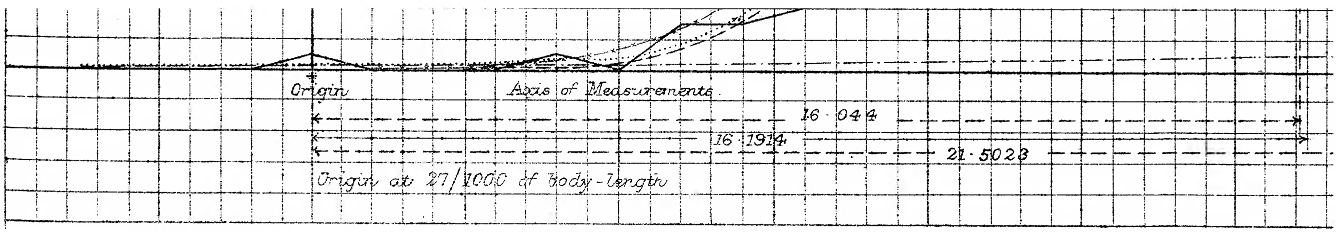


Fig. V.
Length of Carapace of Prawns
2nd Solution of Nonic:
$p_2 = -\cdot 782, 6295.$
Axis of Individuals. Scale: 1 unit = 2 Prawns.
Origin
Axis of Measurements
Axis of First Component
Centroid Vertical

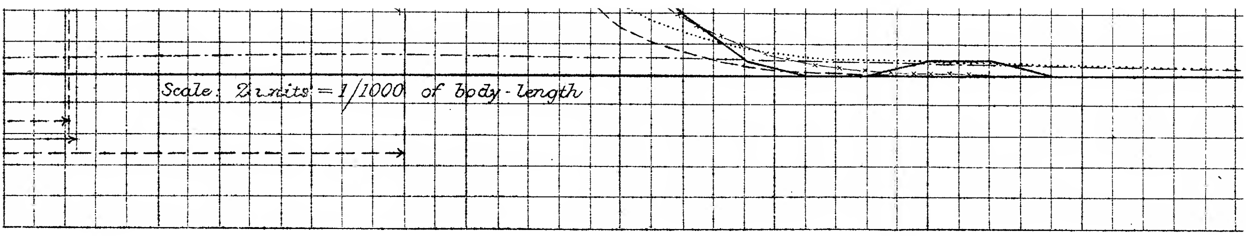


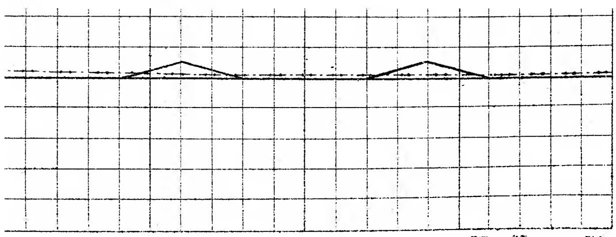
No of Individuals
Observations 998
1st Component 971
2nd Component 27
Resultant 998
Normal Curve 998
Standard Deviations
1st Component 3·3897
2nd Component 8·9330
Normal Curve 3·7572
Axis of First Component
Centroid Vertical
Axis of Second Component
Scale: 2 units = 1/1000 of body-length


duals



Origin
Axis of Measurements
16·044
16·1914
21·5023
Origin at 27/1000 of Body-length


Scale: 2 units = 1/1000 of body-length



West, Newman, lith.

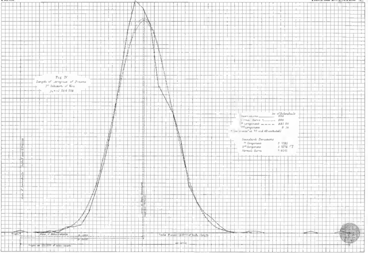

No. of Individuals
Normal Curve
Standard Deviations
Normal Curve

Fig IV

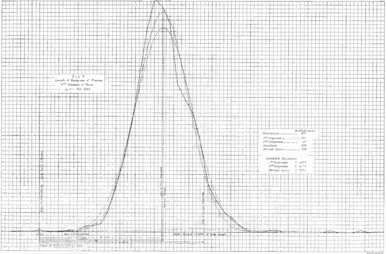


Fig V
Length of Carapace of Prawns
2nd Selection of Naples
Observations
1st Component
2nd Component
Resultant
Normal Curve
No. of Individuals
Standard Deviations
1st Component
2nd Component
Normal Curve