CHAPTER XXXVIII.

ERRORS OR UNCERTAINTIES OF OBSERVATIONS.

1744. Suppose a large number of observations to be made to ascertain the measurement of some physical element. To fix the ideas take one of the simplest kind, the distance between two marked points A and B on a straight rod. Suppose the distance AB to be roughly known to be 10 feet long, but that its true value T is unknown to the observers, of whom there are many, but known to some other person. And suppose that as great accuracy as possible is required. Out of a large number of observations by careful observers, it is clear that there will be none of them which differ very much from the true value T. The more care is taken, and the more accurate the means of measurement at disposal, the closer will the estimates be together. And it is a matter of experience that slight over estimates are as likely as under estimates, and occur with equal frequency. Absolute "mistakes" of counting feet or inches, or of registration of units, or of the use of the instruments we are not considering. In fact we eliminate from this explanation any errors which are of the class of careless "blunders."

It will be found by the person who knows the true value T, that very few of the estimates differ from T by as much as $\frac{1}{2}$ an inch either way; fewer still by $\frac{3}{4}$ of an inch, still fewer by a whole inch, whilst errors of 4 or 5 inches would not occur in the tabulated results of the observations at all. And if the *number* of observations which give an error between x and $x+dx$ be represented graphically, it will be found that the graph takes the form of a curve symmetrical about the y-axis, having a maximum ordinate at the origin, falling rapidly to the x-axis, the ordinate speedily becoming insensibly small (see Fig. 586).

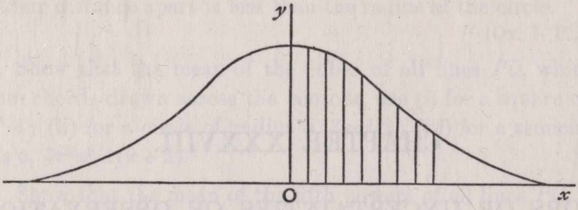

Fig. 586.

1745. It follows, therefore, that for the existence of an error of magnitude lying between x and $x+dx$, there will be a far greater probability when x is small than when x is large; *i.e.* a far greater number of errors of observation will fall between x and $x+dx$ for small values of x than for larger ones. Let $\phi(x)dx$ be that number. We wish to examine the nature of this function $\phi(x)$. And about it we know that

(i) it decreases very rapidly as x increases;

(ii) it must be such as to become insensibly small within a short range of values of x;

(iii) it must be an even function of x, as errors of excess or defect are equally numerous within corresponding limits;

(iv) it must contain some constant or constants depending upon the goodness of the observation, the training and competence of the observer, the accuracy of the instruments used, and the circumstances under which the observation is made;

(v) the number of observations must be $\int_{-\infty}^{\infty} \phi(x)dx$, and supposing N be this number, the chance that the error of any particular observation lies between x and $x+dx=\phi(x)dx/N=\psi(x)dx$, say.

1746. Laplace's Investigation.

Starting with the hypothesis that an error in an observation is due to no one single cause, but is the aggregate of the cumulative effects of a large number of causes, each producing its own separate effect, and that these effects are extremely small, and as likely to be positive as negative, Laplace has shown by a very laborious and difficult investigation that the chance that the error lies in magnitude between x and $x+dx$, viz. $\psi(x)dx$, is $\sqrt{\frac{\omega}{\pi}}e^{-\omega x^2}dx$ for some value of ω which depends upon the goodness of the observation. The argument is of such length that we must refer the reader to Laplace's original work (*Théorie Analytique des Probabilités*). We therefore assume the law as our fundamental hypothesis in what follows. A good idea of the principal steps in the process, which avoids the obscurity of the original work of Laplace, will be found in Airy's *Theory of Errors of Observation*, pages 7 to 15. Todhunter's *History of Probability*, Arts. 1001 onwards, may be consulted, also a paper by Leslie Ellis (*Trans. Camb. Phil. Soc.*, viii.), and a paper by Merriman (*Trans. Conn. Acad.*, iv.).

1747. **The Frequency Law.**

The law $\psi(x)=\sqrt{\frac{\omega}{\pi}}e^{-\omega x^2}$ is termed the law of "Facility" or "Frequency" of Errors. It will be noticed at once that this is a probable law, for it answers all the requirements laid down in Art. 1745. It has a maximum at $x=0$, it is an even function of x, it contains an arbitrary constant ω, it diminishes with great rapidity as x increases, and speedily becomes of insensible magnitude, and

$$\int_{-\infty}^{\infty}\phi(x)\,dx=N\int_{-\infty}^{\infty}\sqrt{\frac{\omega}{\pi}}e^{-\omega x^2}dx=N.$$

1748. **Weight and Modulus.**

The constant ω is called the *weight* of the observation. It is sometimes replaced by $\frac{1}{c^2}$. Then c or $\frac{1}{\sqrt{\omega}}$ is called the *modulus*. The weight ω measures the care, skill and precision of the observer, the goodness of his instruments and the excellence of the conditions under which the observation is made.

1749. The ordinary method of estimating the value of a physical element of which a number of presumably equally good measurements have been made is to take the arithmetical mean of the result. As a matter of experience this gives good results, and therefore this mean is frequently adopted as giving the best estimate available, and regarded as the most likely value. If we might assume this, the above law of Facility of Errors easily follows.

Let T be the true value of the measured quantity, T being unknown. Let $z_1, z_2, \dots z_n$ be n independent results of observation; $\phi(x)$ the law of Facility.

Then $z_1 - T$, $z_2 - T$, ... $z_n - T$ are the actual errors, some positive, some negative, and the *à priori* probability of the coexistence of these errors is proportional to the product

$$P \equiv \phi(z_1 - T)\,\phi(z_2 - T) \dots \phi(z_n - T).$$

Then, by the principles of inverse probability, the probability that the true value lies between T and $T + dT$ is $P\,dT \Big/ \int P\,dT$, the limits being such that the integration is conducted over all values of T which it is capable of assuming. That is, after the observations were made, the probability that T is the true value is also proportional to the product P, and therefore this expression is to be made a maximum by variation of T. Taking logarithms and differentiating, we have $\sum_1^n \phi'(z_r - T)/\phi(z_r - T) = 0$.

Now, if we take for T the arithmetic mean of the observations, this equation is to hold when $nT = \sum_1^n z_r$. To find the form of ϕ which will satisfy these requirements, take the case $z_2 = z_3 = \dots = z_n = z_1 - n\tau$. Then

$$nT = z_1 + (n-1)\,z_2 = z_1 + (n-1)(z_1 - n\tau) = nz_1 - n(n-1)\,\tau,$$

i.e. $\quad z_1 - T = (n-1)\,\tau, \quad z_2 - T = (z_2 - z_1) + (z_1 - T) = -\tau,$

$$z_3 - T = -\tau, \text{ etc.};$$

$$\therefore \frac{\phi'(z_1 - T)}{\phi(z_1 - T)} + (n-1)\frac{\phi'(z_2 - T)}{\phi(z_2 - T)} = 0$$

or

$$\frac{\phi'(n-1)\,\tau}{(n-1)\,\tau\phi\,(n-1)\,\tau} = \frac{\phi'(-\tau)}{(-\tau)\,\phi(-\tau)},$$

which is independent of n; and this is to be true for all positive integral values of n.

This will be satisfied if ϕ be such that $\frac{1}{u}\frac{\phi'(u)}{\phi(u)}=\text{const.}=C$;

whence $\log\phi(u)=C\frac{u^2}{2}$ and $\phi(u)=Ae^{C\frac{u^2}{2}}$.

And since $\phi(u)$ is to decrease as u increases, C must be negative. Let $C=-\frac{2}{c^2}$. Then $\phi(u)=Ae^{-\frac{u^2}{c^2}}$. Again, if N be the total number of observations,

$$N=\int_{-\infty}^{\infty}\phi(u)\,du=\int_{-\infty}^{\infty}Ae^{-\frac{u^2}{c^2}}du=Ac\sqrt{\pi};\quad\therefore\ A=N/c\sqrt{\pi},$$

i.e.
$$\phi(x)=\frac{N}{c\sqrt{\pi}}e^{-\frac{x^2}{c^2}},$$

which establishes the law of facility under the hypothesis specified as to the Arithmetic mean.

This remark is made by Dr. Glaisher in the solutions of the *Senate H. Problems* for 1878, pages 167, 168, where there will also be found a concise account of the allied subject of the principle of "Least Squares." [See also Todhunter, *Hist.*, Art. 1014.]

1750. **Mean of the Errors, Mean of the Squares, Error of Mean Square, Probable Errors.**

The following facts will now appear:

(1) The mean of all the positive errors

$$=\frac{\int_0^{\infty}x\frac{1}{c\sqrt{\pi}}e^{-\frac{x^2}{c^2}}dx}{\int_0^{\infty}\frac{1}{c\sqrt{\pi}}e^{-\frac{x^2}{c^2}}dx}=\frac{c}{\sqrt{\pi}}=\frac{1}{\sqrt{\pi\omega}}.$$

(2) The mean of all the negative errors with their signs changed is also $\frac{c}{\sqrt{\pi}}=\frac{1}{\sqrt{\pi\omega}}$.

(3) The mean of all the errors taken positively is $\frac{c}{\sqrt{\pi}}=\frac{1}{\sqrt{\pi\omega}}$.

(4) The mean of the squares of all the errors

$$=\frac{\int_{-\infty}^{\infty} x^2 \frac{1}{c\sqrt{\pi}} e^{-\frac{x^2}{c^2}} dx}{\int_{-\infty}^{\infty} \frac{1}{c\sqrt{\pi}} e^{-\frac{x^2}{c^2}} dx}=\frac{c^2}{2}=\frac{1}{2\omega}.$$

(5) The "Error of Mean Square," *i.e.* the square root of the mean of the squares of the errors, $=\frac{c}{\sqrt{2}}=\frac{1}{\sqrt{2\omega}}$. This is the abscissa of the point of inflexion on the Probability Curve $y=e^{-\frac{x^2}{c^2}}$.

(6) The "Probable Error," which is such that the number of positive errors which are greater than itself is equal to the number which are less, is given by the value of p, where

$$\int_0^p \frac{1}{c\sqrt{\pi}} e^{-\frac{x^2}{c^2}} dx=\frac{1}{2}\int_0^{\infty} \frac{1}{c\sqrt{\pi}} e^{-\frac{x^2}{c^2}} dx=\frac{1}{4}.$$

Let $x=cz$. Then $\frac{1}{\sqrt{\pi}}\int_0^{\frac{p}{c}} e^{-z^2} dz=0{\cdot}25.$

Tables have been calculated for the values of this integral for various values of the upper limit [Kramp's *Refractions; Encyc. Metropol.*, "Theory of Probabilities"], and interpolation from them gives $\frac{p}{c}={\cdot}476948\ldots$ Hence the "Probable Error" $={\cdot}476948\ldots c$ or ${\cdot}476948\ldots/\sqrt{\omega}$.

1751. **Kramp's Table** is given by Airy (*Th. of Errors*, p. 22), also by De Morgan (*Diff. Calc.*, p. 657). We reproduce Airy's abstract of this table for convenience for other purposes.

Integral tabulated, $I\equiv\frac{1}{\sqrt{\pi}}\int_0^x e^{-z^2} dz.$

x	I	x	I	x	I	x	I
0·0	0·000000	1·0	0·421350	2·0	0·497661	3·0	0·499988
0·1	0·056232	1·1	0·440103	2·1	0·498510		
0·2	0·111351	1·2	0·455157	2·2	0·499068		
0·3	0·164313	1·3	0·467004	2·3	0·499428		
0·4	0·214196	1·4	0·476143	2·4	0·499655		
0·5	0·260250	1·5	0·483053	2·5	0·499796		
0·6	0·301928	1·6	0·488174	2·6	0·499881		
0·7	0·338901	1·7	0·491895	2·7	0·499932		
0·8	0·371051	1·8	0·494545	2·8	0·499962		
0·9	0·398454	1·9	0·496395	2·9	0·499979	∞	0·500000

1752. Relative Magnitude of Probable Error, Mean Error, Error of Mean Square, Modulus.

To sum up, we have

Probable Error $= \cdot 476948 \ldots / \sqrt{\omega}$;

Mean Error $= 1/\sqrt{\pi\omega} = \cdot 564189 \ldots / \sqrt{\omega}$;

Error of Mean Square $= 1/\sqrt{2\omega} = \cdot 707107 \ldots / \sqrt{\omega}$;

Modulus $= 1/\sqrt{\omega}$;

in each case varying inversely as the square root of the weight, *i.e.* directly as the modulus; and obviously, when any one of these is found the rest may be deduced. They are arranged in ascending order of magnitude.

Taking the x-axis as the axis of magnitude of errors and the y-axis as the axis of frequency, Fig. 587 will exhibit to the eye the relative magnitude of these errors and the fall in frequency. The figure is that given by Airy (*loc. cit. sup.*). The abscissa is the ratio of the magnitude of an error to the modulus. The points P, M in the figure indicate respectively the abscissae for Probable and Mean Error.

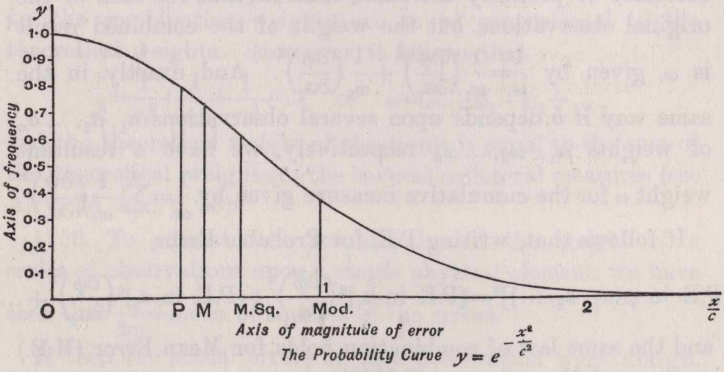

Fig. 587.

1753. Several Observations. Resultant Weight.

Suppose there to be a result b dependent upon two observations a_1 and a_2 of weights ω_1, ω_2 respectively, say $b = \phi(a_1, a_2)$. To find the weight of the result.

Let x_1, x_2 be the actual errors and z the consequent error in b; all being small quantities of the first order, then to that order $z = \frac{\partial\phi}{\partial a_1} x_1 + \frac{\partial\phi}{\partial a_2} x_2 = \phi_{a_1} x_1 + \phi_{a_2} x_2$, say.

The chance of the co-existence of errors in a_1 and a_2 respectively between x_1 and x_1+dx_1 for the one and x_2 and x_2+dx_2 for the other is

$$\sqrt{\frac{\omega_1}{\pi}}e^{-\omega_1 x_1^2}dx_1 . \sqrt{\frac{\omega_2}{\pi}}e^{-\omega_2 x_2^2}dx_2.$$

Therefore writing $\frac{1}{\omega}\equiv\frac{1}{\omega_1}\phi_{a_1}^2+\frac{1}{\omega_2}\phi_{a_2}^2$, and $A=\frac{\omega}{\omega_1}\phi_{a_1}$, the chance of an error in b lying between z and $z+dz$ is

$$\frac{\sqrt{\omega_1\omega_2}}{\pi}\int_{-\infty}^{\infty}dx_1\left[\frac{dz}{\phi_{a_1}}e^{-\omega_1 x_1^2-\omega_2\left(\frac{\phi_{a_1}}{\phi_{a_2}}\right)^2\left(\frac{z}{\phi_{a_1}}-x_1\right)^2}\right],$$

that is, $$=\frac{\sqrt{\omega_1\omega_2}}{\pi}e^{-\omega z^2}\frac{dz}{\phi_{a_2}}\int_{-\infty}^{\infty}e^{-\frac{\omega_1\omega_2}{\omega\phi_{a_2}^2}(x-Az)^2}dx$$

$$=\frac{\sqrt{\omega_1\omega_2}}{\pi}e^{-\omega z^2}\frac{dz}{\phi_{a_2}}.\sqrt{\frac{\pi\omega\phi_{a_2}^2}{\omega_1\omega_2}}=\sqrt{\frac{\omega}{\pi}}e^{-\omega z^2}dz.$$

The law of facility for the compound result $\phi(a_1, a_2)$ is therefore of precisely the same form as that for each of the original observations, but the weight of the combined result is ω, given by $\frac{1}{\omega}=\frac{1}{\omega_1}\left(\frac{\partial\phi}{\partial a_1}\right)^2+\frac{1}{\omega_2}\left(\frac{\partial\phi}{\partial a_2}\right)^2$. And exactly in the same way if b depends upon several observations $a_1, a_2, \ldots a_n$ of weights $\omega_1, \omega_2, \ldots \omega_n$ respectively, we have a resultant weight ω for the cumulative measure given by $\frac{1}{\omega}=\sum_1^n\frac{1}{\omega_r}\left(\frac{\partial\phi}{\partial a_r}\right)^2$.

It follows that, writing P.E. for Probable Error,

$$[\text{P.E. in } \phi(a_1, a_2, \ldots)]^2=(\text{P.E. in } a_1)^2\left(\frac{\partial\phi}{\partial a_1}\right)^2+(\text{P.E. in } a_2)^2\left(\frac{\partial\phi}{\partial a_2}\right)^2+\ldots,$$

and the same law of combination holds for Mean Error (M.E.) or Error of Mean Square (E.M.S.).

1754. For example, if we require the weight of the Arithmetic Mean of n observations of equal weights ω_1,

$$b=\sum_1^n a_r/n \quad\text{and}\quad \frac{1}{\omega}=\frac{1}{\omega_1}\Sigma\frac{1}{n^2}=\frac{1}{n\omega_1}, \textit{ i.e. } \omega=n\omega_1.$$

That is the weight of the combination is n times the weight of any of the original observations, and

the Probable Error in $b=(\text{P.E. in any of the } a\text{'s})/\sqrt{n}$, etc.

Similarly the weight of a resultant $pa_1+qa_2+ra_3+\dots$ is given by

$$\frac{1}{\omega}=\frac{p^2}{\omega_1}+\frac{q^2}{\omega_2}+\frac{r^2}{\omega_3}+\dots;$$

and if $\omega_1=\omega_2=\omega_3=\dots$, $\frac{1}{\omega}=\frac{p^2+q^2+r^2+\dots}{\omega_1}$.

1755. If observations be taken upon a single physical element, and the *weights* and *probable errors* of the several observations $(a_1, a_2, a_3, \dots)$ be respectively $(\omega_1, \omega_2, \omega_3, \dots)$ and $(\epsilon_1, \epsilon_2, \epsilon_3, \dots)$, whilst ω and ϵ are those of a *resultant* formed according to the law $\Sigma p_r a_r/\Sigma p_r$, which is the usual form adopted, where $(p_1, p_2, p_3, \dots)$ are certain constant multipliers, called "combination weights," to be so determined as to give a minimum probable error in that resultant, we have

$$\epsilon^2=\epsilon_1^2\left(\frac{p_1}{\Sigma p_r}\right)^2+\epsilon_2^2\left(\frac{p_2}{\Sigma p_r}\right)^2+\dots;$$

and differentiating with regard to $p_1, p_2, p_3, \dots$,

$$p_1\epsilon_1^2=p_2\epsilon_2^2=p_3\epsilon_3^2=\dots=\Sigma p_r^2\epsilon_r^2/\Sigma p_r,$$

i.e. $\quad p_1/\omega_1=p_2/\omega_2=p_3/\omega_3=\dots,$

i.e. the combination weights are to be proportional to the theoretical weights. Moreover, it follows that

$$\frac{1}{\epsilon^2}=\frac{1}{\epsilon_1^2}+\frac{1}{\epsilon_2^2}+\frac{1}{\epsilon_3^2}+\dots \quad \text{or} \quad \omega=\omega_1+\omega_2+\omega_3+\dots,$$

and the theoretical weight of the result is equal to the sum of the theoretical weights of the several collateral measures (see Airy, *Th. Err.*, p. 56).

1756. To estimate the actual value of the weight of a series of observations upon a single physical element, we have seen that $\frac{1}{2\omega}=$ mean of squares of the errors.

If then the actual errors of each observation were known, we should have a rule to determine ω. But the exact measurement of the quantity upon which the observations are made is rarely known. Let T be its true value, $A_1, A_2, \dots A_n$ the observed values. Then A_1-T, A_2-T, etc., are the *actual* errors, and $\frac{1}{2\omega}=\frac{1}{n}\sum_1^n(A_r-T)^2$. But T being unknown, we have to *approximate.* Let us adopt the arithmetical mean of the observations as the value of T, and write $T=\frac{1}{n}\sum_1^n A_r$, which is known as the "apparent value," but is not necessarily the *true* one. This gives as an approximation

$$\frac{n}{2\omega}=A_1^2+A_2^2+\ldots+A_n^2-2T(A_1+A_2+\ldots)+nT^2=\sum_1^n A_r^2-nT^2,$$

i.e. as an *approximation* we have $\frac{1}{2\omega}=\frac{1}{n}\sum_1^n A_r^2-\frac{1}{n^2}\left(\sum_1^n A_r\right)^2$

$$=\begin{pmatrix}\text{Mean of squares}\\ \text{of observations}\end{pmatrix}-\begin{pmatrix}\text{Square of mean}\\ \text{of observations}\end{pmatrix}.$$

1757. **Determination of the "Error of Mean Square," "Probable Error," etc., of a Measurement of an Element from the Apparent Errors.**

Since the true value of the measured element is rarely or never known, we have to devise a method of obtaining the Error of Mean Square, etc., by some way other than as being $1/\sqrt{2\omega}$, which would require a knowledge of ω. Let $A_1, A_2, A_3, \ldots$ be the actual results of n independent observations on the single physical element in question, $a_1, a_2, a_3, \ldots$ the actual errors, T the true value; then $A_1=T+a_1$, $A_2=T+a_2$, etc.

Let M and m be the arithmetic means of the A's and of the a's. Then

$$a_r-m=A_r-T-\frac{1}{n}\sum_1^n(A_r-T)=A_r-\frac{1}{n}\Sigma A_r=A_r-M.$$

The difference a_r-m, viz. the difference between the actual error and the mean of the actual errors, is called the "Apparent Error." And the sum of the squares of the Apparent Errors

$$=\sum_1^n(a_r-m)^2=\Sigma a_r^2-2m\,.\,nm+nm^2=\Sigma a_r^2-\frac{1}{n}(\Sigma a_r)^2.$$

Therefore, if $Q\equiv\Sigma(A_r-M)^2$, we have $Q=\Sigma a_r^2-\frac{1}{n}(\Sigma a_r)^2$.

Now let ϵ be the error of mean square of each measure.

Then (Art. 1750, 5) $\epsilon^2=\frac{1}{n}\sum_1^n a_r^2$, *i.e.* $\sum_1^n a_r^2=n\epsilon^2$.

Again, the square of Σa_r = sq. of error in ΣA_r

$$=(\text{Error of mean square in } \Sigma A_r)^2$$

$$=\sum_1^n(\text{Error of mean square in } A_r)^2$$

$$=n\epsilon^2 \quad \text{(Art. 1753);}$$

$\therefore$ sum of squares of Apparent Errors $=n\epsilon^2-\frac{1}{n}n\epsilon^2=(n-1)\epsilon^2$.

Hence $\epsilon=\sqrt{\frac{Q}{n-1}}$; and Q being known, this determines ϵ.

Since the Error of mean square $=1/\sqrt{2\omega}$, we have

$$\omega=(n-1)/2Q.$$

Also Mean Error $=\frac{1}{\sqrt{\pi\omega}}=\sqrt{\frac{2}{\pi}\frac{Q}{n-1}}$;

Probable Error $=\frac{0{\cdot}476948}{\sqrt{\omega}}=0{\cdot}476948\ldots\sqrt{\frac{2Q}{n-1}}$.

1758. Again, since the Error of mean square *of the mean* of n independent measures of a physical quantity

$=\frac{1}{\sqrt{n}}\times$ Error of mean square of any one measure (Art. 1754)

$=\frac{1}{\sqrt{n}}\epsilon=\sqrt{\frac{Q}{n(n-1)}}$, we also have

$$\left.\begin{array}{r}\text{Mean Error}\\ \text{of the mean}\end{array}\right\}=\sqrt{\frac{2}{\pi}\frac{Q}{n(n-1)}},$$

$$\left.\begin{array}{r}\text{Probable Error}\\ \text{of the mean}\end{array}\right\}=0{\cdot}476948\ldots\sqrt{\frac{2Q}{n(n-1)}}.$$

1759. **Case of a System of Physical Elements.**

Suppose next that it is required to discover the values of a certain set of physical elements $\xi, \eta, \zeta, \ldots$, and that observations upon certain connected groups of them have been taken giving results of the form

$$\phi_1(\xi, \eta, \zeta, \ldots)=N_1, \quad \phi_2(\xi, \eta, \zeta, \ldots)=N_2, \text{ etc.},$$

the forms of ϕ_1, ϕ_2, etc., being known, and all the constants involved being known from theoretical or other considerations, whilst $N_1, N_2, \ldots$ are the results of observation, and therefore subject to small errors.

Theoretically, if the number (m) of observations be the same as the number (μ) of elements to be found, there will be a definite number of sets of solutions of these equations depending upon the degrees of the several functions. If, however, the number of observations exceed the number of elements, it will not in general be possible to satisfy all the equations by the same values of ξ, η, ζ, etc., and it becomes important to examine a method of finding their most probable values under the circumstances.

1760. **Reduction of the Equations to Linear Form.**

The observed quantities N_1, N_2, etc., will not differ largely from those which would give true values to ξ, η, ζ, etc., and if we solve μ of these equations we shall obtain close approximations to the values of ξ, η, ζ, etc., or in some cases such close approximations may be otherwise available. Let these approximate values be α, β, γ, etc., and x, y, z, etc., the small residuals of the true values of ξ, η, ζ, etc., so that $\xi=\alpha+x$, $\eta=\beta+y$, etc., and these residuals being small their second and higher powers and products may be rejected, and each equation of form $\phi_i(\xi, \eta, \zeta, \ldots)=N_i$ may be regarded as reduced after expansion of $\phi_i(\alpha+x, \beta+y, \ldots)$ by Taylor's theorem to the type

$$a_i x+b_i y+c_i z+\ldots=n_i,$$

such equations being m in number. Now n_i being itself the result of the subtraction of $\phi(\alpha, \beta, \gamma, \ldots)$ and various second and higher order small quantities from N_i depends upon the observations, and is a small quantity subject to error, whilst a_i, b_i, c_i, ... are supposed known from theoretical or other considerations.

1761. **The Equations of Condition.**

We therefore have m linear equations connecting μ unknowns x, y, z, etc., μ being $<m$. Let a typical equation be $a_i x+b_i y+c_i z+\ldots-n_i=0$, where $i=1, 2, 3, \ldots m$. We need not for the moment consider x, y, z, ... to be small.

These m equations are not in general capable of being satisfied by the same values of x, y, z, ..., but we have to obtain the most probable values of x, y, z, ... from them; that is, as good an approximation as we can under the circumstances.

These equations are called the "Equations of Condition."

1762. **Standardisation of the Equations.**

As to the several results of observation, n_1, n_2, n_3, ..., let us suppose that they are each the result of several separate and independent observations; *e.g.* taking the typical case n_i, suppose it to have been formed as the arithmetic mean of ω_i observations upon the value of $a_i x+b_i y+\ldots$, and suppose all these ω_i observations to be equally good observations. Then the weight of this observation is proportional to ω_i.

Therefore, unless the number of observations in forming $n_1, n_2, n_3, \ldots$ has been the same and the individual observations equally good, some of the Equations of Condition will have greater importance than others.

If n_i be found by ω_i observations, each with the same probable error ϵ, the probable error in n_i is $\epsilon/\sqrt{\omega_i}$, and the probable error in $n_i . \sqrt{\omega_i}$ is ϵ.

Hence, if we multiply the Equations of Condition by $\sqrt{\omega_1}$, $\sqrt{\omega_2}$, $\sqrt{\omega_3}$, etc., we get another group in which the probable errors of the right-hand sides are each ϵ.

We shall suppose our m Equations of Condition to have been already subjected to this preparation, and therefore suppose that the quantities $n_1, n_2, n_3, \ldots$ which occur are subject to the same probable error ϵ.

1763. Principle of Least Squares.

If $x_0, y_0, z_0, \ldots$ be the most probable values of $x, y, z, \ldots$ respectively, then, by the nature of the case,

$$a_i x_0 + b_i y_0 + c_i z_0 + \ldots - n_i$$

is a small quantity of the nature of an error. Call it v_i. Then the probability of the occurrence of the error v_i being $\sqrt{\frac{\omega_i}{\pi}} e^{-\omega_i v_i^2} dv_i$, the probability of the co-existence of errors $v_1, v_2, \ldots v_i \ldots v_m$ is $\prod_1^m \sqrt{\frac{\omega_i}{\pi}} e^{-\omega_i v_i^2} dv_i$ and as these errors have occurred through taking $x_0, y_0, z_0, \ldots$, etc., as the true values of $x, y, z, \ldots$, etc., the probability that x_0, y_0, etc., are the true values is $\prod_1^m \sqrt{\frac{\omega_i}{\pi}} e^{-\omega_i v_i^2} dv_i \Big/ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \prod_1^m \sqrt{\frac{\omega_i}{\pi}} e^{-\omega_i v_i^2} dv_i$, in which the denominator is a definite constant; and, supposing the Conditional Equations to have been prepared as described in the preceding article, the ω's occurring are all equal.

But in any case we have to determine x_0, y_0, etc., so that this probability shall be as great as possible; and this will be achieved by making $\sum_1^m \omega_i v_i^2$ a minimum; or, if the ω's are equal, $\Sigma v_i^2 =$ a minimum. The method of procedure is therefore called the method of "Least Squares."

1764. The "Normal" Equations.

The primary condition for a minimum is

$$\sum_1^m \omega_i v_i (a_i\, dx_0 + b_i\, dy_0 + \dots) = 0,$$

and therefore, on equating to zero the coefficients of dx_0, dy_0, ..., we have m linear equations to determine x_0, y_0, z_0, ..., viz

$$\Sigma\omega_i a_i v_i = 0, \quad \Sigma\omega_i b_i v_i = 0, \quad \Sigma\omega_i c_i v_i = 0, \quad \text{etc.};$$

or in the case when the equations have been prepared beforehand, so that the weights are equal,

$$\Sigma a_i v_i = 0, \quad \Sigma b_i v_i = 0, \quad \Sigma c_i v_i = 0, \quad \text{etc., } i.e.;$$

$$\left.\begin{array}{l} \Sigma a^2 . x_0 + \Sigma ab . y_0 + \Sigma ac . z_0 + \dots = \Sigma an, \\ \Sigma ba . x_0 + \Sigma b^2 . y_0 + \Sigma bc . z_0 + \dots = \Sigma bn, \\ \Sigma ca . x_0 + \Sigma cb . y_0 + \Sigma c^2 . z_0 + \dots = \Sigma cn, \\ \qquad \text{etc.,} \end{array}\right\} \text{which are known as the "Normal" Equations.}$$

The very compact notation $[ab]$, $[aa]$, etc., is often used for Σab, Σa^2, etc., but we adopt the sigma notation as a little easier to write.

These equations determine the values of x_0, y_0, etc., so as to give the most probable values of x, y, etc., to satisfy the original group of Conditional Equations in which the n's are subject to small errors.

1765. Before proceeding further, let us examine the m prepared equations of type $a_i x + b_i y + c_i z + \dots = n_i$ from another point of view.

Multiply the several equations by $p_1, p_2, \dots, p_m$ and add; then by $q_1, q_2, \dots q_m$ and add; then by $r_1, r_2, \dots r_m$ and add; and so on; viz. by μ groups of multipliers, m in each group. We obtain μ equations,

$$\left.\begin{array}{l} x\Sigma a_i p_i + y\Sigma b_i p_i + z\Sigma c_i p_i + \dots = \Sigma n_i p_i, \\ x\Sigma a_i q_i + y\Sigma b_i q_i + z\Sigma c_i q_i + \dots = \Sigma n_i q_i, \\ x\Sigma a_i r_i + y\Sigma b_i r_i + z\Sigma c_i r_i + \dots = \Sigma n_i r_i, \\ \qquad \text{etc.} \end{array}\right\} \dots\dots\dots\dots(1)$$

Again multiply these by $\lambda_1, \lambda_2, \dots, \lambda_\mu$ and add, and choose the λ's so as to remove the terms $y, z, \dots$, *i.e.*

$$\left.\begin{array}{l} \lambda_1 \Sigma b_i p_i + \lambda_2 \Sigma b_i q_i + \lambda_3 \Sigma b_i r_i + \dots = 0, \\ \lambda_1 \Sigma c_i p_i + \lambda_2 \Sigma c_i q_i + \lambda_3 \Sigma c_i r_i + \dots = 0, \\ \qquad \text{etc.} \end{array}\right\}$$

Then $$x=\frac{\lambda_1\Sigma n_i p_i+\lambda_2\Sigma n_i q_i+\lambda_3\Sigma n_i r_i+\dots}{\lambda_1\Sigma a_i p_i+\lambda_2\Sigma a_i q_i+\lambda_3\Sigma a_i r_i+\dots}=\frac{\Sigma n_i k_i}{\Sigma a_i k_i},$$

where $k_i=\lambda_1 p_i+\lambda_2 q_i+\lambda_3 r_i+\dots$; whilst $\Sigma b_i k_i=0$, $\Sigma c_i k_i=0$, etc., and the new constant multipliers k_1, k_2, k_3, ... replace the p's, q's, r's, etc., and λ's.

Let ω be the weight of each of the observations n_1, n_2, ... n_m, by supposition prepared to be of equal weight, and let ω_x, ω_y, ω_z, ... be the weights of the deduced values of x, y, z, ...

Then $$\frac{1}{\omega_x}=\frac{\Sigma k_i^2}{(\Sigma a_i k_i)^2}\frac{1}{\omega},\quad \text{Art. 1753.} \quad \dots\dots\dots\dots\dots(2)$$

And if ϵ be the error of mean square, or the probable error in each of the n's, and ϵ_x, ϵ_y, ϵ_z, ... the resulting error of mean square, or the probable error in the deduced values of x, y, z, ..., we therefore have $\epsilon_x^2=\frac{\Sigma k_i^2}{(\Sigma a_i k_i)^2}\epsilon^2$, and we have to make this error of mean square, or this probable error, as small as possible with the conditions $\Sigma b_i k_i=0$, $\Sigma c_i k_i=0$, etc.

1766. To do this we have the k's at our disposal. Their number is m and their connecting equations number $\mu-1$, which is $< m$. It will be observed that the expression ϵ_x contains only the *ratios* of the k's, and when their ratios to any particular standard k have been fixed ϵ_x becomes determinate. We shall therefore in no way alter the value of ϵ_x by the addition of some one additional linear equation amongst the k's. For convenience we take that relation as $\Sigma a_i k_i=1$, which will give $x=\Sigma n_i k_i$. We then have to make $\epsilon_x^2=\Sigma k_i^2\,.\,\epsilon^2$ a minimum with the μ conditions $\Sigma a_i k_i=1$, $\Sigma b_i k_i=0$, $\Sigma c_i k_i=0$, etc. We obtain at once $\Sigma k_i dk_i=0$, $\Sigma a_i dk_i=0$, $\Sigma b_i dk_i=0$, etc., and by Lagrange's method of undetermined multipliers

$$k_1=Aa_1+Bb_1+\dots,\ k_2=Aa_2+Bb_2+\dots,\ \dots k_m=Aa_m+Bb_m+\dots,$$

whence $$\Sigma k_i^2=A\Sigma a_i k_i=A.$$

Also $$\left.\begin{aligned} A\Sigma a^2+B\Sigma ab+C\Sigma ac+\dots&=\Sigma a_i k_i=1,\\ A\Sigma ba+B\Sigma b^2+C\Sigma bc+\dots&=\Sigma b_i k_i=0,\\ A\Sigma ca+B\Sigma cb+C\Sigma c^2+\dots&=\Sigma c_i k_i=0,\\ \text{etc.},&\end{aligned}\right\}\dots\dots\dots\dots(3)$$

whence $A = \begin{vmatrix} \Sigma b^2, & \Sigma bc, & \ldots \\ \Sigma cb, & \Sigma c^2 & \ldots \\ \ldots & \ldots & \ldots \end{vmatrix} \Big/ \begin{vmatrix} \Sigma a^2, & \Sigma ab, & \Sigma ac & \ldots \\ \Sigma ba, & \Sigma b^2, & \Sigma bc & \ldots \\ \Sigma ca, & \Sigma cb, & \Sigma c^2 & \ldots \\ \ldots & \ldots & \ldots & \ldots \end{vmatrix}$ and is known,

and $A = \Sigma k_i^2$. Therefore $\epsilon_x^2 = A\epsilon^2$ and $\epsilon_x = \epsilon\sqrt{A}$; and A is essentially positive, being the sum of a number of squares of real quantities. The weight of the deduced value for x is $\omega_x = \dfrac{1}{\Sigma k_i^2}\omega = \dfrac{1}{A}\cdot\omega$, and if we take ω as unity, $\omega_x = \dfrac{1}{A}$.

1767. The symmetry of the work shows that the same process will give us a minimum error of mean square, or a minimum probable error also for y or for z, etc., and that the weight of y so deduced may be found by solving equations of the same form as those in group (3), but with the 1 now replaced by 0 in the first equation and the 0 by 1 in the second; and so on for the weights of z, etc.

1768. Again it will be noticed that if we choose our preliminary multipliers, viz. the p's, q's, r's, etc., as the coefficients of the original prepared conditional equations, viz. $p_i = a_i$, $q_i = b_i$, $r_i = c_i$, etc., we have $k_i = \lambda_1 a_i + \lambda_2 b_i + \lambda_3 c_i + \ldots$, and for this choice

$$\Sigma k_i^2 = \Sigma(\lambda_1 a_i + \lambda_2 b_i + \ldots)k_i = \lambda_1 \Sigma a_i k_i + \lambda_2 \Sigma b_i k_i + \ldots = \lambda_1 = A.$$

That is, substituting for the p's, q's, r's, ... in equations of group (1), the equations which will give a value of x with the least error of mean square, or least probable error for x are the "normal" equations arrived at in Art. 1764 otherwise, and the symmetry shows that the values of y, z, etc., will also be determined by the same equations with the least error. But as these equations are the same as those arrived at by making $\Sigma(a_i x + b_i y + \ldots - n_i)^2$ a minimum by variation of $x, y, z, \ldots$, this is a convenient way of reproducing the equations for these unknowns. And the result is the same as that arrived at in Art. 1764, the weights of the several observations having been made equal by preparation of the conditional equations.

1769. If the conditional equations are left unprepared, we arrive at the proper equations for the values of x, y, z, etc., by making $\Sigma\omega_i(a_i x + b_i y + \ldots - n_i)^2$ a minimum.

1770. Reality of $\sqrt{A}$.

The determinants occurring in Art. 1766 are essentially positive. For such a determinant as

$$\begin{vmatrix} \Sigma a^2, & \Sigma ab, & \Sigma ac, \ldots \\ \Sigma ba, & \Sigma b^2, & \Sigma bc, \ldots \\ \Sigma ca, & \Sigma cb, & \Sigma c^2, \ldots \\ \ldots & \ldots & \ldots \end{vmatrix} \text{ occurs in squaring the rectangular array } \begin{vmatrix} a_1, & a_2, \ldots a_m \\ b_1, & b_2, \ldots b_m \\ c_1, & c_2, \ldots c_m \\ \ldots & \ldots \end{vmatrix},$$

in which the number of rows (μ) is less than the number of columns (m), and is therefore expressible as the sum of the squares of all possible determinants which can be formed from the array by taking μ columns (Burnside & Panton, *Th. of Eq.*, p. 260). Such a determinant is therefore essentially positive.

1771. To complete the theory we must examine how the quantity ϵ is to be found from the details before us; that is, we are to do for the case of measurements upon a system of physical elements what was done in Art. 1757 for the measurement of a single element. We have used ϵ indifferently in Art. 1765, etc., for either the error of mean square, the probable error or the mean error. We shall now define the letter as standing definitely for the "error of mean square" in the measure of an observation. Let v_i be the residual error in $a_ix+b_iy+\ldots-n_i$, when the values $x_0, y_0, z_0, \ldots$ obtained from the "normal" equations have been substituted for $x, y, z, \ldots$. Then we shall show that the equation $\epsilon=\sqrt{\dfrac{\Sigma v_i^2}{m-\mu}}$ replaces that of Art. 1757.

Let the true values of $x, y, z, \ldots$ be $x_0+\delta x$, $y_0+\delta y$, $z_0+\delta z$, etc., and let

$$a_i(x_0+\delta x)+b_i(y_0+\delta y)+\ldots-n_i=u_i \quad (i=1 \text{ to } i=m).$$

Multiply by a_i and add the system. Then

$$\left.\begin{aligned} &\Sigma a^2.\,x_0+\Sigma ab.\,y_0+\Sigma ac.\,z_0+\ldots-\Sigma an \\ &\quad+\Sigma a^2.\,\delta x+\Sigma ab.\,\delta y+\Sigma ac.\,\delta z+\ldots &&=\Sigma au; \\ \therefore\ &\Sigma a^2.\,\delta x+\Sigma ab.\,\delta y+\Sigma ac.\,\delta z+\ldots &&=\Sigma au. \\ \text{Similarly }\ &\Sigma ba.\,\delta x+\Sigma b^2.\,\delta y+\Sigma bc.\,\delta z+\ldots &&=\Sigma bu, \\ &\Sigma ca.\,\delta x+\Sigma cb.\,\delta y+\Sigma c^2.\,\delta z+\ldots &&=\Sigma cu, \text{ etc.,} \end{aligned}\right\}$$

which, as in Arts. 1765, 1766, give $\delta x=\Sigma ku$.

1772. Equations of type $a_ix_0+b_iy_0+\ldots-n_i=v_i (i=1 \text{ to } i=m)$, multiplied by v_i and added, give $\Sigma v_i^2=-\Sigma n_iv_i$, since

$$\Sigma av=0, \quad \Sigma bv=0, \quad \Sigma cv=0, \quad \text{etc.}$$

And equations of type $a_i(x_0+\delta x)+b_i(y_0+\delta y)+\ldots-n_i=u_i$ give in the same way $\Sigma u_iv_i=-\Sigma n_iv_i$.

Hence $\Sigma v_i^2=\Sigma u_iv_i=-\Sigma n_iv_i$.

1773. Equations $a_ix_0+b_iy_0+c_iz_0+\ldots-n_i=v_i$, multiplied by u_i and added, give

$$\Sigma a_iu_i\,.\,x_0+\Sigma b_iu_i\,.\,y_0+\ldots-\Sigma n_iu_i=\Sigma v_iu_i=\Sigma v_i^2.$$

Equations $a_i(x_0+\delta x)+b_i(y_0+\delta y)+\ldots-n_i=u_i$, multiplied by u_i and added, give

$$\Sigma a_iu_i\,.\,x_0+\Sigma b_iu_i\,.\,y_0+\ldots-\Sigma n_iu_i$$
$$+\Sigma a_iu_i\,\delta x+\Sigma b_iu_i\,.\,\delta y+\ldots \qquad =\Sigma u_i^2.$$

Hence $\Sigma u_i^2=\Sigma v_i^2+\Sigma a_iu_i\,.\,\delta x+\Sigma b_iu_i\,.\,\delta y+\Sigma c_iu_i\,.\,\delta z+\ldots$.

And, since Σu_i^2 is the sum of the squares of the true errors of the observations, $\Sigma u_i^2=m\epsilon^2$.

Now, in the terms $\Sigma a_iu_i\,.\,\delta x+\Sigma b_iu_i\,.\,\delta y+\ldots$, we must necessarily approximate.

Take for them their mean values. Then

$$\Sigma a_iu_i\,.\,\delta x=(a_1u_1+a_2u_2+a_3u_3+\ldots)(k_1u_1+k_2u_2+k_3u_3+\ldots),$$

whose mean value is that of $a_1k_1u_1^2+a_2k_2u_2^2+a_3k_3u_3^2+\ldots$; for, remembering that the errors $u_1, u_2, u_3, \ldots$ may have either sign, all products involving errors with unequal suffixes will disappear in taking the mean. And the mean values of $u_1^2, u_2^2, u_3^2, \ldots$ are each ϵ^2.

Hence $\Sigma a_iu_i\,.\,\delta x$ will be replaced by $\Sigma a_ik_i\,.\,\epsilon^2$, that is ϵ^2.

Similarly $\Sigma b_iu_i\,.\,\delta y$, $\Sigma c_iu_i\,.\,\delta z, \ldots$ will be replaced by ϵ^2.

Therefore $m\epsilon^2=\Sigma v_i^2+\mu\epsilon^2$, μ being the number of unknowns $x, y, z, \ldots$.

Hence $$\epsilon^2=\frac{\Sigma v_i^2}{m-\mu}.$$

1774. If there be but one unknown, *i.e.* when the observations are made upon a single physical element, we have $\epsilon^2=\dfrac{\Sigma v_i^2}{m-1}$. (Art. 1757.)

1775. **Effect of Exact Co-existent Relations.**

If, in addition to the m conditional equations of type

$$a_i x + b_i y + \ldots - n_i = 0,$$

there be p $(<\mu)$ exact equations of type

$$\alpha_i x + \beta_i y + \ldots - \nu_i = 0,$$

these latter equations may be regarded as determining p of the unknowns in terms of the other $\mu - p$. Upon substitution of these in the conditional equations, we have a system of m conditional equations amongst $\mu - p$ unknowns. Hence the error of mean square ϵ will in this case be given by $\epsilon = \sqrt{\dfrac{\Sigma v_i^2}{m+p-\mu}}$, where v_i is, as before, $a_i x_0 + b_i y_0 + \ldots - n_i$, and the summation is from $i=1$ to $i=m$.

If μ be large, or if there be several exact equations, a different process is usually employed to reduce the labour of the elimination. (For this see Chauvenet, *Astron.*, p. 552, Vol. II.)

1776. Finally, if $\epsilon_x, \epsilon_y, \epsilon_z, \ldots$ be the errors of mean square in $x_0, y_0, z_0, \ldots$, and if $X, Y, Z, \ldots$ be the respective weights of $x_0, y_0, z_0, \ldots$, then $\epsilon_x = \dfrac{\epsilon}{\sqrt{X}}$, $\epsilon_y = \dfrac{\epsilon}{\sqrt{Y}}$, etc., and the values of $X, Y, Z, \ldots$ are to be determined as follows (Art. 1766):

For X:
$$\left.\begin{aligned} \Sigma a^2 \cdot \frac{1}{X} + \Sigma ab \cdot \frac{1}{Y'} + \Sigma ac \cdot \frac{1}{Z'} + \ldots &= 1, \\ \Sigma ba \cdot \frac{1}{X} + \Sigma b^2 \cdot \frac{1}{Y'} + \Sigma bc \cdot \frac{1}{Z'} + \ldots &= 0, \\ \Sigma ca \cdot \frac{1}{X} + \Sigma cb \cdot \frac{1}{Y'} + \Sigma c^2 \cdot \frac{1}{Z'} + \ldots &= 0, \\ \text{etc.}; \end{aligned}\right\}$$

For Y:
$$\left.\begin{aligned} \Sigma a^2 \cdot \frac{1}{X''} + \Sigma ab \cdot \frac{1}{Y} + \Sigma ac \cdot \frac{1}{Z''} + \ldots &= 0, \\ \Sigma ba \cdot \frac{1}{X''} + \Sigma b^2 \cdot \frac{1}{Y} + \Sigma bc \cdot \frac{1}{Z''} + \ldots &= 1, \\ \Sigma ca \cdot \frac{1}{X''} + \Sigma cb \cdot \frac{1}{Y} + \Sigma c^2 \cdot \frac{1}{Z''} + \ldots &= 0, \\ \text{etc.}, \end{aligned}\right\}$$

the accented unknowns of each group not being required; and such equations may obviously be written down from the normal equations.

Hence we obtain X, *i.e.* the value of $\frac{1}{A}$ (Art. 1766), etc.

Moreover, in cases where the values of x_0, y_0, z_0, ... are expressed in terms of letters and not numerically, their weights may be obtained more readily, as in Art. 1753, by differentiation.

This completes the details of the process to obtain the Mean Square error for each element, and the Probable and Mean errors may be at once deduced.

1777. Order of Procedure.

To sum up, the order of procedure is as follows:

I. Given the m conditional equations amongst μ unknowns ($m > \mu$) of type $a_ix+b_iy+c_iz+\ldots-n_i=0$, let each have been prepared by multiplication by the square root of its weight, viz. $\sqrt{\omega_i}$.

II. From these prepared equations, or by differentiating

$$\Sigma(a_ix+b_iy+\ldots-n_i)^2,$$

deduce the normal equations and find x_0, y_0, z_0, ...

III. Form $\Sigma v_i^2 \equiv \Sigma(a_ix_0+b_iy_0+\ldots-n_i)^2$.

IV. Find ϵ, the error of Mean Square of an observation from $\epsilon=\sqrt{\frac{\Sigma v_i^2}{m-\mu}}$.

V. Then to find ϵ_x, ϵ_y, ϵ_z, etc., in the normal equations replace Σan, Σbn, Σcn, ... by 1, 0, 0, etc., and solve for x, say $x=\frac{1}{X}$; then replace Σan, Σbn, Σcn, ... by 0, 1, 0, ..., etc., and solve for y, say $y=\frac{1}{Y}$, and so on; then X, Y, Z, ... are the several weights of x_0, y_0, z_0, ..., and the errors of Mean Square are $\epsilon_x=\frac{\epsilon}{\sqrt{X}}$, $\epsilon_y=\frac{\epsilon}{\sqrt{Y}}$, ...

These values may also be obtained by Art. 1753 without the trouble of solving the normal equations when the results of the observations are given in letters instead of numerical quantities.

VI. Having found ϵ, ϵ_x, ϵ_y, ϵ_z, ..., we may then deduce the Probable Error or the Mean Error by Art. 1752.

1778. For further information, the reader may consult the appendix to Vol. II. of Chauvenet's *Sph. and Practical Astronomy*.

For those interested in the Bibliography of the subject, reference may be made to

Legendre, *Nouvelles Méthodes pour la détermination des orbites des Comètes*, 1806.

Gauss, *Theoria Motus Corporum Coelestium*, 1809.

Disquisitio de Elementis Ellipt. Palladis, 1811, etc.

Bertrand, *Méthode des moindres carrées*, 1855.

Encke, *Ueber der Meth. d. Klein. Quad.*, Berlin (*Astr. Year Book*, 1834, etc.).

Laplace, *Théorie analytique des Probabilités*.

Poisson, *Sur la probabilité des resultats moyens des observations* (*Connaisance des Temps*, 1827).

Bessel, *Astron. Nach.* (357, 358, 399).

Hansen, Do. (192, 292, etc.).

Peirce, *Astron. Journal* (Camb. Mass., Vol. II.).

Liagre, *Calc. des Prob.*, Brussels, 1852.

And other references have been made to the works of Airy, Glaisher and Merriman in the course of this chapter.

1779. Illustrative Examples.

1. *Suppose O a central station on a plain, and A, B, C, D four distant points. Let the angles AOB, BOC, COD, DOA be respectively estimated by p, q, r, s, equally good measurements to be* α, β, γ, δ; *and suppose that after all due care has been taken* $\alpha+\beta+\gamma+\delta$ *falls a little short of* 360°, *say by* k''. *It is required to find the corrections to be applied to the several observations.*

Suppose the true values of the several angles to be $\alpha+x''$, $\beta+y''$, $\gamma+z''$, $\delta+w''$.

Then $x+y+z+w=k$ is an exact equation. The equations of condition are $x=0$, $y=0$, $z=0$, $x+y+z-k=0$, which cannot be satisfied simultaneously. Making $px^2+qy^2+rz^2+s(x+y+z-k)^2$ a minimum, we have $px=qy=rz=-s(x+y+z-k)=\lambda$, say. These are the Normal Equations.

Thus $x=\frac{\lambda}{p}$, $y=\frac{\lambda}{q}$, $z=\frac{\lambda}{r}$ and $\lambda\left(\frac{1}{p}+\frac{1}{q}+\frac{1}{r}\right)=k-\frac{\lambda}{s}$; *i.e.* $\lambda=\frac{k}{\Sigma\frac{1}{p}}$;

whence $$x=\frac{k}{p}\Big/\Sigma\frac{1}{p},\quad y=\frac{k}{q}\Big/\Sigma\frac{1}{p},\quad z=\frac{k}{r}\Big/\Sigma\frac{1}{p},\quad w=\frac{k}{s}\Big/\Sigma\frac{1}{p},$$

which give the probable values of x, y, z, w.

2. *Let p observations of the zenith distance of a circumpolar star be made at the upper culmination, and q at the lower. It is required to find the co-latitude of the place.* [Airy, p. 42, *Errors of Observation*.]

Let a and b be the means of the two sets of observations. Then $z_1=a$ and $z_2=b$ are the estimated zenith distances at the two culminations. And we are to find the probable error in $\frac{1}{2}(a+b)$, which would be the true co-latitude if the means of the observations were accurate.

Let ω be the weight of any of the original observations, all supposed of equal value; ω' the weight of $\frac{1}{2}(a+b)$. Then

$$\frac{1}{\omega'}=\frac{1}{4}\frac{1}{p\omega}+\frac{1}{4}\frac{1}{q\omega}=\frac{1}{4\omega}\frac{p+q}{pq}.$$

Hence if ϵ and ϵ' be the probable errors of an observation and of the deduced co-latitude, $\epsilon'=\frac{1}{2}\sqrt{\frac{p+q}{pq}}\,\epsilon$, with the same formula connecting the errors of mean square and the mean errors.

3. *Consider a rod, whose accurate weight is h grammes, to be broken into three random pieces of unknown weights x, y, z grammes; y and z are weighed together l times; z and x, m times; x and y, n times, and the means of the three sets of weighings are a, b and c grammes, and all the weighings are equally good observations so far as is known. It is required to find the most probable weights of the several parts and the probable error in each.*

[MATH. TRIP., 1876.]

Here $x+y+z=h$, (1); $y+z=a$, (2); $z+x=b$, (3); $x+y=c$, (4).

Equation (1) is exact. The others are subject to error. Let ω be the "weight" of any one observation. The "weights" of the means are $l\omega$, $m\omega$, $n\omega$. The equations (2), (3), (4) may be written $h-x-a=0$, $h-y-b=0$, $h-z-c=0$, and we are to make

$$l(h-x-a)^2+m(h-y-b)^2+n(h-z-c)^2$$
$$=\text{a minimum with condition } x+y+z=h.$$

Thus, $l(h-x-a)dx+\,+\,=0$, $dx+\,+\,=0$, whence $l(h-x-a)=\ldots=\ldots=\lambda$,

i.e. $3h-(x+y+z)-a-b-c=\lambda\left(\frac{1}{l}+\frac{1}{m}+\frac{1}{n}\right)$, *i.e.* $2h-a-b-c=\lambda\left(\frac{1}{l}+\,+\right)$,

i.e. $\qquad x=h-a-(2h-a-b-c)\dfrac{mn}{mn+nl+lm}$, $\quad y=\text{etc.}$, $z=\text{etc.}$

If ω_x be the "weight" of this expression for x,

$$\frac{1}{\omega_x}=\frac{1}{\omega l}\left(\frac{\partial x}{\partial a}\right)^2+\frac{1}{\omega m}\left(\frac{\partial x}{\partial b}\right)^2+\frac{1}{\omega n}\left(\frac{\partial x}{\partial c}\right)^2=\text{etc.}=\frac{1}{\omega}\frac{m+n}{mn+nl+lm}.$$

Now h being known exactly, $2h-a-b-c$ is a known error, and it is the only known error, and if Ω be the "weight" of this expression $\frac{1}{\Omega}=\frac{1}{\omega l}+\frac{1}{\omega m}+\frac{1}{\omega n}$ (Art. 1753), and $\frac{1}{2\Omega}=(2h-a-b-c)^2$ (Art. 1750). The latter equation is the approximative one for Ω. Hence

$$\frac{1}{\omega}=\frac{2lmn}{mn+nl+lm}(2h-a-b-c)^2,\quad \frac{1}{\omega_x}=\frac{2lmn(m+n)}{(mn+nl+lm)^2}(2h-a-b-c)^2.$$

The probable error for x, viz. p, is such that

$$\sqrt{\frac{\omega_x}{\pi}}\int_0^p e^{-\omega_x x^2}dx=\frac{1}{4}\quad\text{and}\quad p=\frac{\cdot4769\ldots}{\sqrt{\omega_x}},$$

i.e. $$p=\cdot4769\ldots\times\frac{\sqrt{2lmn(m+n)}}{mn+nl+lm}(2h-a-b-c).$$

Suppose in the same example that h were not known, but that the several observations are $(a_1, a_2, \ldots a_l)$, $(b_1, b_2, \ldots b_m)$, $(c_1, c_2, \ldots c_n)$.

We then have l equations of type $y+z-a_r=0$, m of type $z+x-b_r=0$, n of type $x+y-c_r=0$.

Then x, y, z are to be found from making

$$\sum_1^l (y+z-a_r)^2+\sum_1^m (z+x-b_r)^2+\sum_1^n (x+y-c_r)^2 \text{ a minimum;}$$

from which

$$\left.\begin{aligned}(m+n)x_0+ny_0+mz_0&=\sum_1^m b_r+\sum_1^n c_r,\\ nx_0+(n+l)y_0+lz_0&=\sum_1^n c_r+\sum_1^l a_r,\\ mx_0+ly_0+(l+m)z_0&=\sum_1^l a_r+\sum_1^m b_r,\end{aligned}\right\} x_0, y_0, z \text{ being the values which give the minimum.}$$

We then have as an approximation

$$\frac{1}{2\omega}=\frac{\sum_1^l (y_0+z_0-a_r)^2+\sum_1^m (z_0+x_0-b_r)^2+\sum_1^n (x_0+y_0-c_r)^2}{l+m+n}.$$

4. *A, B, C, D are four points in order on a straight line; AB, BC, CD, AC, BD, AD are measured respectively $\alpha, \beta, \gamma, \delta, \epsilon, \zeta$ times with mean respective measurements a, b, c, d, e, f. Find the most probable value of AB; and if $\alpha=\beta=\gamma=\delta=\epsilon=\zeta$, find its probable error.* (MATH. TRIP., 1878.]

Let $AB=x$, $BC=y$, $CD=z$; then we are to find a minimum for

$$\alpha(x-a)^2+\beta(y-b)^2+\gamma(z-c)^2+\delta(x+y-d)^2+\epsilon(y+z-e)^2+\zeta(x+y+z-f)^2$$

The conditions are:

$$\left.\begin{aligned}&\alpha(x-a)+\delta(x+y-d)+\zeta(x+y+z-f)=0,\\ &\beta(y-b)+\delta(x+y-d)+\epsilon(y+z-e)+\zeta(x+y+z-f)=0,\\ &\gamma(z-c)+\epsilon(y+z-e)+\zeta(x+y+z-f)=0,\end{aligned}\right\} \text{which determine } x, y, z.$$

In the case $\alpha=\beta=$ etc., these become

$$3x+2y+z=a+d+f, \quad 2x+4y+2z=b+d+e+f, \quad x+2y+3z=c+e+f;$$

whence

$$x=\tfrac14(2a-b+d-e+f);\ y=\tfrac14(-a+2b-c+d+e);\ z=\tfrac14(-b+2c-d+e+f),$$

i.e.

$$\begin{aligned}x-a&=\tfrac14(-2a-b+d-e+f), & x+y-d&=\tfrac14(a+b-c-2d+f),\\ y-b&=\tfrac14(-a-2b-c+d+e), & y+z-e&=\tfrac14(-a+b+c-2e+f),\\ z-c&=\tfrac14(-b-2c-d+e+f), & x+y+z-f&=\tfrac14(a+c+d+e-2f),\end{aligned}$$

and the sum of the squares of these six expressions is, say K.

We also have

$$\frac{1}{\omega_x}=\frac{1}{16}(4+1+1+1+1)\frac{1}{\omega}, \quad \frac{1}{\omega_y}=\frac{1}{16}(1+4+1+1+1)\frac{1}{\omega},$$

$$\frac{1}{\omega_z}=\frac{1}{16}(1+4+1+1+1)\frac{1}{\omega},$$

i.e. $\omega_x=2\omega$, $\omega_y=2\omega$, $\omega_z=2\omega$ by (Art. 1753), or they may be derived as in Art. 1776.

Now $\frac{1}{\omega}=\sqrt{\frac{K}{6-3}}=\sqrt{\frac{K}{3}}$ (Art. 1773); $\therefore \frac{1}{\omega_x}=\frac{1}{\omega_y}=\frac{1}{\omega_z}=\frac12\sqrt{\frac{K}{3}}$;

whence the Mean Errors, Mean Square Errors and Probable Errors of x, y, z may be at once written down.

[See *Sol. S.H. Prob.*, Glaisher, 1878, p. 165.]

PROBLEMS.

1. In a plane triangle the angles A, B, C are respectively measured m, n and p times, and the means of these measurements are respectively α, β and γ, and $\alpha+\beta+\gamma=\pi+\epsilon$. The separate measurements are equally good. Show that if $\alpha+x$, $\beta+y$, $\gamma+z$ be the true values of the angles, the probable values of x, y, z are

$$-np\epsilon/\delta, \quad -pm\epsilon/\delta, \quad -mn\epsilon/\delta, \quad \text{where } \delta=np+pm+mn.$$

2. In the plane triangle ABC, the side b is to be determined in terms of a from the measured values of B and C. Find the actual error in the determination of b in terms of the actual errors of measurement of B and C, and the probable error of b in terms of the probable error of any measurement supposed to be the same for the measurement of any angle. Show that of all the directions in which the side b can be drawn, that gives the probable error of the determination of its length a minimum for which the angle C satisfies the equation

$$ab(2a^2+3b^2)(1+\cos^2 C)=(a^4+7a^2b^2+2b^4)\cos C.$$

[MATH. TRIPOS.]

3. At Pine Mount, a station in the U. S. Coast Survey, the angles subtended by four surrounding stations A, B, C, D were observed as follows:

AB, weight 3, $65°\ 11'\ 52''\cdot500$; CD, weight 3, $87°\ 2'\ 24''\cdot703$;

BC, weight 3, $66°\ 24'\ 15''\cdot553$; DA, weight 1, $141°\ 21'\ 21''\cdot757$.

The five points are in one plane. It is required to estimate the corrected values of these angles. The result is that the several results in the seconds should be $53''\cdot4145$, $16''\cdot4675$, $25''\cdot6175$, $24''\cdot5005$, the degrees and minutes being unaltered.

[CHAUVENET, *Astron.*, II., p. 551.]

4. Taking the equations

$$x-y+2z-3=0, \qquad 4x+y+4z-21=0,$$
$$3x+2y-5z-5=0, \qquad -x+3y+3z-14=0,$$

show that (1) the probable values of x, y, z are $2\cdot470$, $3\cdot551$, $1\cdot916$ respectively;

(2) the weights of x, y, z are $24\cdot597$, $13\cdot648$, $53\cdot927$;

(3) the error of mean square of an observation, *i.e.* of the numbers 3, 5, 21, 14, is $0\cdot284$;

(4) the errors of mean square of x, y, z are $0\cdot057$, $0\cdot077$, $0\cdot039$;

(5) the probable errors of an observation and of x, y, z are respectively 0·192, 0·038, 0·052, 0·026.

[GAUSS, *Th. Motus*; CHAUVENET, II., p. 521.]

5. In finding the latitude of a place by observation of two meridian altitudes of a circumpolar star, p observations are made at the upper transit, q at the lower. Taking the probable error of each observation at the upper transit as ϵ_1, and at the lower as ϵ_2, and all astronomical and instrumental corrections to have been applied, show that the probable error in the determination of the latitude is $\sqrt{p\epsilon_2^2 + q\epsilon_1^2}/\sqrt{2pq}$.

6. If the altitudes of the upper and lower transits of several circumpolar stars be observed and H_1, H_2, H_3, ... be the harmonic means of the numbers of observations at the upper and lower transits for the several stars, and all observations be equally trustworthy, with a common probable error ϵ, supposing all astronomical and instrumental corrections to have been applied, show that the probable error in the determination of the latitude will be $\frac{\epsilon}{\sqrt{2}}[\Sigma H]^{-\frac{1}{2}}$.

7. At three stations P, Q, R on the same meridian, the zenith distances of n_1 stars are observed at each of the stations P, Q, R; n_2 at P and Q; n_3 at Q and R; n_4 at R and P. It is required to determine the amplitude of the portion PQ of the meridian. Show that there are four independent modes of determining that arc; and on the supposition that the probable error of each observation is the same and $=\epsilon$, show how to determine the combination weights of the four measures. If $n_1 = n_2 = n_3 = n_4 = n$, show that the square of the probable error in the result $= \frac{4}{5}\frac{\epsilon^2}{n}$.

8. State the criterion for the selection of the combination weights of n independent measures of a magnitude. Determine the probable error of the result in terms of the probable errors of the n measures.

In the observation of the zenith distances of stars for the amplitude of a meridian divided into four sections by three stations intermediate between the extreme stations, a stars are observed at the first, second, third only; b at the second, third, fourth; c at the third, fourth, fifth; and the probable error of every observation is ϵ. Show that there are only three independent modes of measuring the whole arc, and obtain equations for determining the combination weights of the three measures. In the case where $a = b = c$, prove that the square of the probable error of the result is $10\epsilon^2/3a$.

[MATH. TRIP.]

9. If $a, b, c, \ldots$ be the actual errors in n measures of a physical element, the apparent error of each measure is defined as the difference of each measure from the mean.

Let Q be the sum of the squares of the apparent errors. Then prove that (i) the Probable error of a measure, (ii) the Mean error of a measure, (iii) the Probable error of the Mean and (iv) the Mean error of the Mean are respectively

$$0{\cdot}674506\sqrt{\frac{Q}{n-1}},\quad 0{\cdot}797885\sqrt{\frac{Q}{n-1}},$$

$$0{\cdot}674506\sqrt{\frac{Q}{n(n-1)}},\quad 0{\cdot}797885\sqrt{\frac{Q}{n(n-1)}}.$$

10. If we have any number of sets of n observations of the value of a physical element, all of which are *à priori* supposed to be equally good, and if the difference between any observation and the mean of the set of n observations to which it belongs be called the apparent error of that observation, then, assuming the usual law of frequency of errors, prove that the mean of the squares of the apparent errors $=\frac{n-1}{n}m^2$, where m^2 is the mean value of the square of an actual error of observation. [SMITH'S PRIZES.]

11. A rod of known length l is broken into four portions. The lengths x, y, z, w of these portions are measured respectively p, q, r, s times under the same circumstances and with the same care. The means of these several measurements are $\alpha, \beta, \gamma, \delta$. Show that the probable length of x is $\alpha + {\cdot}6745\frac{l-(\alpha+\beta+\gamma+\delta)}{\Sigma p^{-1}}\sqrt{\frac{1}{p}\left(\frac{1}{q}+\frac{1}{r}+\frac{1}{s}\right)}$.

12. The angles of a geodetic triangle of known spherical excess are measured, and the probable errors of the several measurements are $\epsilon_1, \epsilon_2, \epsilon_3$ respectively. It is found that the sum of the three measurements needs a correction of θ''. Show that if $\alpha'', \beta'', \gamma''$ be the corrections to be applied to the angles,

$$\alpha/\epsilon_1{}^2=\beta/\epsilon_2{}^2=\gamma/\epsilon_3{}^2=\theta/(\epsilon_1{}^2+\epsilon_2{}^2+\epsilon_3{}^2).$$