IX. *On the Proof of the Law of Errors of Observations.*

By MORGAN W. CROFTON, *F.R.S.*

Received March 24,—Read April 22, 1869.

1. So much has been published upon the Theory of Errors, that some apology seems to be required from a new writer who does not profess to have arrived at any results which were unknown to his predecessors. Nevertheless, so great, as is well known, are the difficulties of the theory, whether we seek to form a correct estimate of the principles on which it rests, or to follow the subtle mathematical analysis which has been found indispensable in reasoning upon them, that any contribution which tends to simplify the processes, without weakening their logical exactness, will probably be considered of some value. My object in this paper is to give the mathematical proof, in its most general form, of the law of single errors of observations, on the hypothesis that an error in practice arises from the joint operation of a large number of independent sources of error, each of which, did it exist alone, would produce errors of extremely small amount as compared generally with those arising from all the other sources combined. Now this proof is contained in a process given for a different object, namely, POISSON'S generalization of LAPLACE'S investigation of the law of the mean results of a large number of observations, to be found in his 'Recherches sur la Probabilité des jugements,' and which is reproduced in Mr. TODHUNTER'S valuable 'History of the Theory of Probability.'

It is obvious that we should altogether restrict the generality of the proof, confining it merely to a few artificial and conventional cases, if we were to suppose each source of error to give positive and negative errors with equal facility, or to assume the law of error (even supposing it unknown) to be the same for all the sources. None of the processes, therefore, contained in the 4th chapter of the 'Théorie Analytique des Probabilités' are of sufficient generality for our purpose, though some writers have so employed them; nor will the method apply here which LESLIE ELLIS has given in his memoir "On the Method of Least Squares" (Camb. Phil. Trans. 1844), based upon FOURIER'S theorem, on account of the assumption of equal facility for positive and negative errors. The proof which follows will be found, I think, of full generality,—the only cases excluded being incompatible with the existence of the exponential law (see art. 7), and at the same time greatly simpler than POISSON'S, dispensing with his refined and difficult analysis*.

2. It is remarkable that the well-known exponential function which is now pretty

* The *length* of this communication may seem at variance with the statement that the proof here given is a simpler one than those of former writers. Still I think it will be found to be so on examination; the length of the paper arises from fuller explanations being given than is usually the case. I am persuaded that the doubts and misconceptions which have prevailed so extensively with relation to this subject have been in great part occasioned by the extreme brevity and scanty explanation of the great writers who have treated it.

generally received among mathematicians as expressing the law of frequency of single errors of observation, does not seem to have been distinctly given by any one of the three great philosophers LAPLACE, GAUSS, and POISSON (who may be called the founders of the Theory of Errors) as being, in their opinion, the expression of that law. It has been erroneously supposed, as LESLIE ELLIS points out, that GAUSS'S and LAPLACE'S proofs of the method of Least Squares depend upon that assumption. It is true that GAUSS'S first method, in the 'Theoria Motûs,' does require it; but he does not present that method as other than tentative and hypothetical: and later, in the 'Theoria Combinationis Observationum,' he says, speaking of the law of single errors, "plerumque incognita est."

As, however, this law of error seems in our day to have been adopted by general consent, some inquiry into the grounds on which its validity rests will be appropriate here. And first I would remark that it can scarcely be maintained that any attempt hitherto made to establish this law independently of the hypothesis I have named in art. 1 has been successful. We may pass by GAUSS'S proof in the 'Theoria Motûs,' which shows that the law must hold *if we take as an axiom* that the arithmetical mean of several observations is the most probable result. Now this really is not an axiom, but only a convenient rule which is generally near the truth: this we see by considering any case in which we are certain that the errors do not follow the exponential law; does the mind see here *à priori* that the rule does not give the most probable result? It seems certain that we should have just the same confidence in it here as in any case; yet GAUSS'S proof shows that it does not give the most probable result*. It should indeed be stated that GAUSS himself (as might have been expected from that acute and accurate mind) is very far from asserting the above assumption to be an axiom; consequently he does not give his proof as more than hypothetical. He only states that the rule is generally accepted—"axiomatis loco haberi solet hypothesis." A method of remarkable simplicity was given by Sir J. HERSCHEL in a very interesting review of QUETELET'S 'Letters on Probability†,' which conducts to the same law of error by means of one or two bold assumptions; but striking as the coincidence is, it can hardly be seriously viewed as a *demonstration*; nor is it formally so presented by its distinguished author. However, the methods both of GAUSS and Sir J. HERSCHEL are of great interest to the natural philosopher, as showing that certain *à priori* mathematical assumptions of a very simple kind lead to the same law of error which reasoning based on a study of the facts which surround us also points out as expressing, at least approximately, what generally does occur *in rerum naturâ*: though we can see no necessity that the facts

* See ELLIS, *loc. cit.* p. 207.

† Edinburgh Review, July 1850. See a criticism by LESLIE ELLIS in the Philosophical Magazine, vol. xxxvii. Also BOOLE (Edinb. Trans. vol. xxi.) and THOMSON and TAIT (Natural Philosophy), who speak more favourably. M. QUETELET'S 'Lettres' will amply repay a perusal; in connexion with our present inquiry, he points out that not only errors of observations, but the variations of many other fluctuating magnitudes, such as the stature of men, the temperature of the weather, &c. from their mean values, seem to follow the same law. If this be so, the inference seems legitimate that these divergences from the mean types, or *errors of Nature* herself, as they may be called, are produced in each case, not by one or two, but by a vast number of hidden coexisting causes.

should be so, it being quite easy to *conceive* a different economy of nature in which no such accordance would subsist*.

It is possible *à priori* to conceive that the law of single errors of observation might be of any form whatever, varying with each kind of observation: how far it is true that in practice one general law will be found to prevail, is essentially a question of facts—an inquiry, not into what *might be*, but what *is*. Now the hypothesis above mentioned,—namely, that errors *in rerum naturâ* result from the superposition of a large number of minuter errors arising from a number of independent sources,—when submitted to mathematical analysis, leads to the law which is generally received; as far therefore as this hypothesis is in accordance with fact, so far is the law practically true. Fully to decide how far this hypothesis does agree with facts is an extremely subtle question in philosophy, which would embrace not only an extended inquiry into the laws of the material universe, but an examination of the senses and faculties of man, which form an important element in the generation of error. Still, without pretending to enter on a demonstration of the truth of this hypothesis, a few reflections upon the facts, especially in the case of Astronomy (which is *par excellence* the science of observation, and where accordingly the lessons of experience are the clearest and most complete), will, I think, at least convince us of its *reasonableness* in certain large classes of errors of observations. Now if we attend to what has taken place in the history of astronomical observation, we find that the gross errors of the earlier observers proceeded mainly from three or four principal causes—for instance, refraction, imperfect measurement of time, and the use of the naked eye in pointing to objects. When these few capital occasions of error were removed (at least approximately), refraction being discovered and allowed for, and the pendulum and telescopic sights introduced, it was found that observations at once attained a high order of accuracy, showing that the principal sources of error had been eliminated. It would seem, in fact, that in coarse and rude observations the errors proceed from *a very few* principal causes; and in this case, consequently, our hypothesis will probably represent the facts only imperfectly, and the frequency of the errors will only approximate roughly and vaguely to the law which follows from it†. But when

* The extreme simplicity of the exponential relation itself, whether considered as expressing the law of single errors, or that of the mean results of a large number of observations, as contrasted with the long and difficult methods by which it was established, has naturally led to several attempts to dispense with or simplify the latter; in some the hypothesis we here adopt is taken as a basis; but, so far as the present writer is aware, every process given, except Poisson's, fails in generality. In a recent Memoir on the Law of Frequency of Error by Professor Tait (Edin. Trans. vol. xxiv.) (where, it should be stated, the learned author speaks with some hesitation, and only gives his method as an attempt), it is assumed that each of the elementary errors which are combined can be assimilated to the deviation from its most probable value of the number of white balls among a given large number of balls drawn from an urn, which contains white and black in a given proportion. It is then shown (as indeed is done in Laplace's 3rd chapter) that this error follows the exponential law. Thus the proof only applies to the combination of a number of elementary errors, each of which follows that law. But it is quite certain that many simple errors do not follow that law; hence the method is altogether deficient in generality.

† We cannot, however, assert this positively, if there is reason to believe that the error which arises from each principal cause is itself a composite error, which certainly is often the case. The "error in time," for

astronomers, not content with the degree of accuracy they had reached, prosecuted their researches into the remaining sources of error, they found that, not three or four, but a *great number* of minor sources of error, of nearly coordinate importance, began to reveal themselves, having been till then masked and overshadowed by the graver errors which had been now approximately removed*. It was as if a small number of forest trees had been cut down, leaving an innumerable growth of shrubs and brushwood at their feet, remaining to be cleared. There were errors of graduation, and many others, in the construction of instruments; other errors of their adjustments; errors (technically so called) *of observation*; errors from changes of temperature, of weather, from slight irregular motions and vibrations; in short, the thousand minute disturbing influences with which modern astronomers are familiar, and which it is superfluous to recapitulate here. Many of these are known and allowed for, or eliminated, at least approximately, in practical astronomy; still we seem to be justified in considering the error which remains as the result of a great number of yet minuter errors, each inconsiderable in itself. Thus a cursory view of the nature of astronomical errors, and the light which this throws on various cognate classes of observations, seem to lead to the conclusion that the above hypothesis will be found to hold, generally, in the case of refined and delicate observations. No doubt much more would be necessary to justify us in asserting

---

instance, is certainly not a simple error, but one resulting from the joint action of several causes, one or more of which we can conceive detected and allowed for, leaving the others in operation. An error may thus arise from the superposition of only three or four component errors, which at first sight are of simple origin, but in reality represent each a group of minor errors; and the hypothesis would then hold. It is questionable whether, among the causes which in practice vitiate any observation, any *simple* error ever does enter, of considerable magnitude and importance as compared with the others combined; such, for instance, as would be the error produced in the time (or through the time on some astronomical magnitude) by the pendulum being $\frac{1}{10}$ of an inch too long or too short, every thing else being pretty accurate. If it be said that ignorance or negligence might produce such a result, we may answer that such negligence or ignorance would make itself felt in other ways also: one such error would not stand alone. Isolated acts of neglect by a careful observer would come under the head of *occasional* errors, as explained further on.

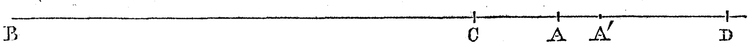
It seems very difficult to discern, *à priori*, the nature of the errors incurred in *estimating* magnitudes by the eye, or of errors arising from the imperfection of our senses, such as those incurred in pointing to a star with the naked eye. It is quite possible that such errors may arise each from *several* sources, though their nature be hidden from our view.

* A similar law to that mentioned above seems to prevail in many kindred cases. Thus in the successive improvements in artillery, machinery, &c., in proportion as the greater sources of imperfection and inaccuracy are understood and remedied, the *number* of minor disturbing influences which are thus rendered perceptible, and still vitiate the results, though to a less extent, increases rapidly. We may even trace a sort of analogy here in various phenomena both of the moral and material universe, which apparently have no bearing on the point we are considering. Thus the principal wants of human nature, the necessaries of life in fact, are very few; and so long as these are supplied with difficulty, minor wants are scarcely felt, as we see in uncivilized communities: but when the greater wants are satisfied, the number and variety of the *secondary* requirements of our nature are visible in the multitudinous productions of civilized life. The diseases which mainly operate in shortening human existence are very few in number; but could they be extirpated, the number of minor causes, of nearly coordinate importance, which still would influence the rate of mortality would be very large. The statistics of crime, and many other phenomena, would give rise to remarks of a similar nature.

this absolutely; thus it is not enough for our purpose to show, could we do so conclusively, that each error in practice is compounded of a large number of smaller errors; we must also show that they are *independent*, at least for the most part. Thus we may conceive one of the minute errors affecting an astronomical magnitude to be an error in the refraction proceeding from a rise in the general temperature, and another affecting the same observation to be an error of time arising from the expansion of the pendulum through the same cause; now these two minute errors are not independent, and would have to be mathematically combined in quite a different way from two that were independent; and, indeed, such a change of temperature would influence the actual error of the observation in other ways also. However, we may at least safely conclude that the hypothesis in question is not a mere arbitrary assumption, but a reasonable and probable account of what does in fact take place in the case of careful and refined observations.

3. In proceeding to submit this hypothesis to mathematical analysis, the minute simple errors which go to form the observed compound error will be assumed to follow each its own unknown law, expressed by different unknown functions of the utmost generality*: positive and negative values of each error will not be assumed equally possible; on the contrary, the cases will be included, as obviously ought to be done, of minute disturbing influences which always cause the observed magnitude to err in excess, and of others which cause it to err only in defect. I will exclude all mention of the term *probability*, and will consider solely the *frequency* or *density* of the error, viewed as a function of its magnitude.

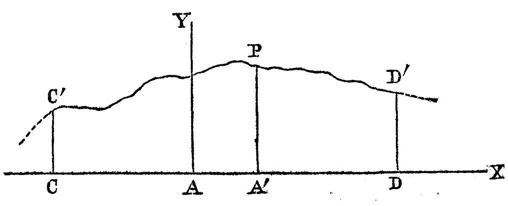
Let any magnitude which has to be determined by observations affected with some one cause of error (simple or compound) be represented by the line BA;



let a large number of such observations be made, and let the observed values be represented by a number of lengths BA′, measured from B: it will be found in general that in the neighbourhood of A the line will be dotted over with a multitude of points A′, the distance AA′ being the error in each case. These dots will begin at some point C, and end at some point D, which generally are on opposite sides of A, but may both be at the same side. Between C and D the dots will be distributed over CD with a variable density: this density, at any point A′, will represent the *frequency* or *density* of errors of magnitude AA′.

If at every point A′ we erect an ordinate A′P representing the density at A′, we shall thus trace out a locus or curve C′D′, whose equation we may call, taking A as origin,

$$y=\phi(x). \quad \ldots\ldots\ldots\ldots \quad (1)$$



* With regard to the limits or amplitudes of the errors, see note on art. 7.

This we may call the *curve* or *function* of Error*. It is of course generally discontinuous, as it is only to include values of x between the points C, D. The function $\phi(x)$ strictly speaking should vanish for all values of x beyond C and D; however, we shall not require any consideration of the analytical methods of expressing such functions. If N be the number of observations taken, and if we put $AD=a$, $AC=b$, then as ydx denotes the number of errors lying between x and $x+dx$,

$$N=\int_{-b}^{a}\phi(x)dx. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

It is well to notice that, if C be any constant, the equation

$$y=C\phi(x)$$

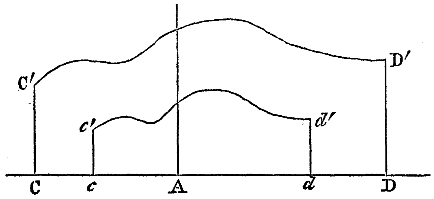
really is the same function of Error as (1), the number of observations only being altered.

4. In order not unduly to limit the generality of the investigation, it is necessary further to study the nature of the possible ways in which the dots we have spoken of as representing the observations may be scattered along the line CD, in the case of various unknown *simple* causes of error; noting also what becomes of the *function* $\phi(x)$ and the curve C′D′ in each case. And first, in many cases the dots will be distributed *continuously* along CD, thus giving a *curve* without gaps or intervals. It is by no means necessary that this curve should descend towards CD at its two extremities more than in the middle; in other words, the extreme values of a simple Error are not always less probable than the intermediate ones. There may be cases where the extreme values are the most probable; for instance, the Error occasioned by supposing a point fixed, which is in reality performing extremely minute and slow oscillations about its mean position. But besides the cases of continuous distribution, there are others, not only conceivable, but which we may be sure do actually occur, in which a function or curve does not assist our conceptions, and we shall do better merely to consider the points or dots themselves. There may be what is called a *constant* Error; that is, some cause which gives the observation always too great (or too small) by the same fixed minute amount: the distribution here is simply a group of N coincident points somewhere on CD. Or a certain cause may only admit of two or more definite values for the error; the distribution will be two or more groups of coincident points, the numbers in each group being equal or unequal. Again, an important class of Errors are those which may be called *occasional Errors*, that is, produced by intermittent causes not always in operation. In such a case, if N observations be made, a certain number of them (say n) are unaffected by the Error; the remaining $N-n$, made when the cause is in operation, we may suppose represented by dots continuously or discontinuously distributed; we have then a group of n coincident points at A, besides a number $N-n$ distributed in some way over CD. Errors of mistake or forgetfulness, and many others also, are of this description.

* The word "error" is sometimes used for shortness to express a source of error. To avoid confusion we may write it with a capital E, when used in this sense. Thus "an Error" will mean a source of error, or the assemblage of actual errors (or the curve or function symbolizing them) which that source produces in a large number of trials, and which form a visible manifestation or representation of it: "an error" will mean a particular magnitude.

5. If we alter the ordinate and abscissa of every point in the curve C′D′ in a given ratio, changing the limits a, $-b$ of the Error in the same ratio, we find the curve $c'd'$ represented by

$$\frac{y}{i}=\phi\left(\frac{x}{i}\right), \quad \ldots \ldots \ldots \ldots \quad (3)$$



which may be called a *similar Error* to (1) or C′D′. The number of observations will be different in the two cases, being represented by the areas of the two figures. We may find it convenient to suppose the number of observations the same; if so

$$y=\frac{1}{i}\phi\left(\frac{x}{i}\right) \quad \ldots \ldots \ldots \ldots \quad (4)$$

will be a *similar function of Error* to $y=\phi(x)$, the number of observations being the same for both, the limits of the error in (3) and (4) being ia, $-ib$.

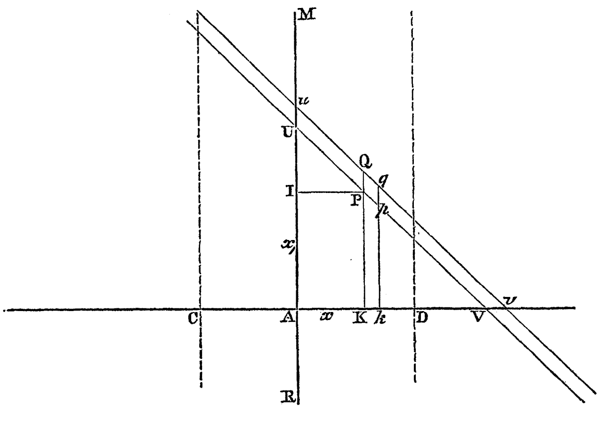
6. *To find the function of Error resulting from the combination of a given Error whose equation is*

$$y_1=f(x_1) \quad \ldots \ldots \ldots \ldots \quad (5)$$

(*the limits being* $\pm\infty$) *with another independent Error*

$$y=\phi(x), \quad \ldots \ldots \ldots \ldots \quad (6)$$

whose limits are a, $-b$.



We shall do this most clearly by help of a geometrical construction. Let the (N) values of the first Error be measured from A according to their signs along the indefinite line MR; likewise measure the (n) values of the second Error along CD, where AD$=a$, AC$=b$. Take any two values, AI$=x_1$ of the first, and AK$=x$ of the second; they give a value $x+x_1$ of the compound Error, to which will correspond a point P of the plane, whose *coordinates* are x, x_1. The number of such points contained within the element

$dx\,dx_1$, each point corresponding to a compound error, will be

$$yy_1\,dx\,dx_1 \text{ or } f(x_1)\phi(x)dS,$$

dS being the element of the area. Draw through P a line UV equally inclined to the axes, then $x+x_1$ is constant along this line; put $\xi=x+x_1=AV$, take $Vv=d\xi$, and draw uv parallel to UV; take $Kk=dx$, then the number of points within the elementary parallelogram $PQpq$ will be

$$f(\xi-x)\phi(x)d\xi\,dx.$$

Hence the whole number of points between the parallels UV and uv (that is, *the number of compound errors whose magnitudes lie between* ξ and $\xi+d\xi$) will be

$$d\xi\int_{-b}^{a} f(\xi-x)\phi(x)dx.$$

The total number of compound errors thus obtained will be Nn; however, for uniformity, we will suppose the number of observations taken, affected with the compound Error, to be N, the same as for (5). This will oblige us to divide by

$$n=\int_{-b}^{a}\phi(x)dx.$$

Thus if we represent the compound Error by a curve whose coordinates are (ξ, η), it will be

$$\eta=\frac{\int_{-b}^{a} f(\xi-x)\phi(x)dx}{\int_{-b}^{a}\phi(x)dx} \quad \ldots\ldots\ldots\ldots \quad (7)$$

Thus if we wish to find the Error resulting from the combination of the two Errors whose equations are

$$y=\frac{N}{\theta\sqrt{\pi}}e^{-\frac{(x-\alpha)^2}{\theta^2}},\quad y=\frac{N'}{\phi\sqrt{\pi}}e^{-\frac{(x-\beta)^2}{\phi^2}},$$

we have from formula (7) (N, N$'$ denoting the numbers of observations),

$$\eta=\frac{N}{\pi\theta\phi}\int_{-\infty}^{\infty} e^{-\frac{(\xi-x-\alpha)^2}{\theta^2}}\,e^{-\frac{(x-\beta)^2}{\phi^2}}\,dx,$$

whence

$$\eta=\frac{N}{\sqrt{(\theta^2+\phi^2)\pi}}e^{-\frac{(\xi-\alpha-\beta)^2}{\theta^2+\phi^2}}.$$

Hence it is easy to see that if any number of Errors of the forms

$$y=\frac{N}{\theta\sqrt{\pi}}e^{-\frac{(x-\alpha)^2}{\theta^2}},\quad y=\frac{N'}{\phi\sqrt{\pi}}e^{-\frac{(x-\beta)^2}{\phi^2}},\quad y=\frac{N''}{\psi\sqrt{\pi}}e^{-\frac{(x-\gamma)^2}{\psi^2}}, \text{ \&c.}$$

be combined, the resultant Error will be

$$y=\frac{N}{\sqrt{\pi(\theta^2+\phi^2+\psi^2+\ldots)}}e^{-\frac{(x-\alpha-\beta-\gamma-\ldots)^2}{\theta^2+\phi^2+\psi^2+\ldots}} \quad \ldots\ldots\ldots \quad (8)$$

Expanding $f(\xi-x)$ in formula (7), it becomes

$$\eta=f(\xi)-\alpha f'(\xi)+\frac{\lambda}{2}f''(\xi)-\frac{\sigma}{2.3}f'''(\xi)+\text{\&c.},$$

where

$$\alpha=\frac{\int_{-b}^{a} x\phi(x)dx}{\int_{-b}^{a}\phi(x)dx},\ \lambda=\frac{\int_{-b}^{a} x^2\phi(x)dx}{\int_{-b}^{a}\phi(x)dx},\ \sigma=\frac{\int_{-b}^{a} x^3\phi(x)dx}{\int_{-b}^{a}\phi(x)dx},\ \&c.,$$

α being the mean value of the Error $y=\phi(x)$, λ its mean square, σ its mean cube, &c.

7. In the problem of finding the law of error resulting from the superposition of a great number of Errors, each of very small importance by itself, we will consider each component Error as the *diminutive* of some Error of finite importance* (see art. 5). Thus if $y=\mathrm{F}(x)$ be some possible finite Error, and we reduce its dimensions in the ratio i, where i is infinitesimal, the diminished Error will be $\frac{y}{i}=\mathrm{F}\left(\frac{x}{i}\right)$; and if the mean value, mean square, mean cube, &c. of the former be called

$$\mathrm{E}_1,\ \mathrm{E}_2,\ \mathrm{E}_3,\ \ldots\ldots,$$

it is easy to see that the same means, for the reduced Error, will be

$$i\mathrm{E}_1,\ i^2\mathrm{E}_2,\ i^3\mathrm{E}_3,\ \ldots\ldots$$

Now adopting the usual axiom that *no function can represent a finite Error unless* E_1, E_2, E_3, ... *are finite*, it follows that the mean cube, mean 4th power, &c. of the

* Thus all conceivable cases of Errors whose extreme limits, or amplitude, are very small, are contained in the above method of proof; also those small Errors which, though their extreme amplitude be not very small, are merely possible finite Errors (of great or infinite amplitude) on a reduced scale. It is necessary, however, to observe, in examining the nature of all the minute simple Errors which our hypothesis in its generality comprises, that there are cases quite conceivable, and involving no absurdity, of simple Errors of trivial or infinitesimal importance which come under neither of these categories, and to which the method in the text will not apply. To give a simple instance, imagine an *occasional* source of Error, which rarely operates, but which, when it does, gives a fixed finite error k (thus we may conceive an observer to mistake, once in a thousand times, the succeeding division of his instrument for the true one). Let this happen on an average once for n times that the cause is not in operation (n being supposed very great); then the mean value of the Error is $\frac{k}{n+1}$, its mean square is $\frac{k^2}{n+1}$, &c. It is therefore of infinitesimal importance whether, with LAPLACE, we estimate the *importance* of an Error by its mean value (irrespective of sign), or, with GAUSS, by its mean square; but as its mean cube &c. cannot be rejected in comparison with the mean square, the above analysis cannot be applied to it. Minute *simple* Errors of such a description must then be excepted from those which are supposed to enter into the composition of the actual errors of observations. If an appreciable number of them did enter, the received exponential law could not hold for the compound Error. Thus were we to combine a large number of small Errors of the nature of the simple instance just cited, the resultant Error would be of a discontinuous nature, represented by groups of coincident points, with finite intervals between them.

Though it is necessary clearly to understand that the full generality of the hypothesis is restricted by the exceptions explained in this note, yet there seems every reason to suppose that such cases are too rare in practice to cause any sensible deviation from the exponential law of error, the great majority of the minute component Errors which jointly affect any observation *in rerum naturâ* having each, it is natural to suppose, a very minute range or amplitude.

diminutive Error may be rejected in comparison with its mean square*. We infer, therefore,

If $y=f(x)$ represent any Error of indefinite amplitude, and if a new Error, $y=\phi(x)$, of indefinitely small importance as compared with it, be superposed, the resulting compound Error will be represented by the equation

$$y=f(x)-\alpha\frac{d}{dx}f(x)+\frac{\lambda}{2}\frac{d^2}{dx^2}f(x), \quad \ldots \ldots \ldots \quad (9)$$

where α, λ are infinitesimal constants, viz. the mean value of the new Error and the mean value of its square†, the number of observations being supposed the same for the Error (9) as for the former, $y=f(x)$.

If we now conceive $y=\phi(x)$ in the above to be one of a large number of independent infinitesimal Errors, and $y=f(x)$ to be the compound finite Error resulting from the combination of all the others, we infer from (9) that each elementary Error $y=\phi(x)$ affects the law of the combined Errors in a manner which only involves (α) the mean value of the elementary Error, and (λ) its mean square. But if this be so, we may, for our present purpose, substitute for $y=\phi(x)$ any other Error whatever which has the same mean value and mean square (provided of course its mean cube &c. may be neglected in comparison with its mean square). We may therefore *for our purpose* replace $y=\phi(x)$ by‡

$$y=\frac{n}{\sqrt{2\pi(\lambda-\alpha^2)}}e^{-\frac{(x-\alpha)^2}{2(\lambda-\alpha^2)}}, \quad \ldots \ldots \ldots \quad (10)$$

which fulfils these conditions.

Likewise, if there be another elementary Error whose mean value is β and mean square μ, we may replace it by

$$y=\frac{n'}{\sqrt{2\pi(\mu-\beta^2)}}e^{-\frac{(x-\beta)^2}{2(\mu-\beta^2)}},$$

and so on, for all the elementary Errors. Hence (see equation 8) the Error compounded of any number of them will be

$$y=\frac{\text{N}}{\sqrt{2\pi(\lambda+\mu+\nu+\ldots-\alpha^2-\beta^2-\gamma^2-\ldots)}}e^{-\frac{(x-\alpha-\beta-\gamma-\ldots)^2}{2(\lambda+\mu+\ldots-\alpha^2-\beta^2-\ldots)}}.$$

* We cannot neglect the mean square as compared with the mean 1st power, as the latter is the algebraical sum of a number of positive and negative elements, which sum may be of any amount, however small, and may sometimes vanish altogether; whereas the former is the sum of a number of *positive* elements, and therefore cannot vanish.

† If the new Error be such as to give any discontinuous distribution of points (see art. 4), it is easy to satisfy ourselves, by the method of art. 6, that the above proposition still holds good. In fact, if the n values of the new Error be x_1, x_2, x_3, ..., we shall have, instead of the formula (7),

$$\eta=\frac{1}{n}\{f(\xi-x_1)+f(\xi-x_2)+f(\xi-x_3)+\&c.\},$$

which is true in all cases, whether the distribution be continuous or discontinuous, or a mixture of both; and hence the formula (9) will follow.

‡ This suggestion is due to Professor J. C. ADAMS, one of the Referees charged by the Royal Society with the duty of reporting upon the present Paper. The remainder of the proof, which was of a different nature in the Paper as originally presented, is much simplified thereby.

We conclude therefore that *if a great number of minute independent Errors be combined, and if we write*

$$\left.\begin{aligned} m&=\alpha+\beta+\gamma+\ldots=\text{sum of mean Errors,}\\ h&=\lambda+\mu+\nu+\ldots=\text{sum of mean squares of Errors,}\\ i&=\alpha^2+\beta^2+\gamma^2+\ldots=\text{sum of squares of mean Errors,}\end{aligned}\right\}^{*} \quad . \quad . \quad . \quad . \quad (11)$$

the resulting function of Error will be

$$y=\frac{\mathrm{N}}{\sqrt{2\pi(h-i)}}e^{-\frac{(x-m)^2}{2(h-i)}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

The Probability of an error being found to lie between x and $x+dx$ is of course

$$\frac{1}{\sqrt{2\pi(h-i)}}e^{-\frac{(x-m)^2}{2(h-i)}}\,dx\dagger.$$

If positive and negative errors in the observation are equally probable, as generally can be secured in practice, at least approximately, then $m=0$; that is, the sum of the mean values of the elementary component Errors vanishes, and the Probability is expressed by the usual value

$$\frac{1}{c\sqrt{\pi}}e^{-\frac{x^2}{c^2}}\,dx.$$

If we calculate by integration from equation (12) the mean value of the composite Error (or, as GAUSS calls it, *the constant part* of the Error) and the mean value of its square, we shall find

Mean Error$=m=$sum of mean values of component Errors,

Mean Square of Error$=h+m^2-i$.

We have thus a verification of the correctness of our analysis, as the same results may be found from independent algebraical computation‡.

8. Considering the celebrity of the question, it may not be superfluous to show how the result might have been obtained without any antecedent knowledge of the peculiar property of combination of the Errors in equation (8).

* We may observe that $h-i$ *is always positive*; for if we take any set of numbers, positive or negative, the mean of their squares is always greater than the square of the mean (see TODHUNTER's 'Algebra,' p. 407). Therefore

$$\lambda>\alpha^2,\ \text{also}\ \mu>\beta^2,\ \nu>\gamma^2,\ \&\text{c.}$$

Consequently $h>i$.

† This expression will be found to agree with POISSON's final result in the memoir already cited.

‡ If

$$\mathrm{U}=a+b+c+d+\&\text{c.},$$

where each of the quantities a, b, c, d, &c. may take any number (different for each quantity) of different independent values, adopting for shortness the symbol M(K) for "the mean value of K," it is not difficult to prove, by elementary algebra, that

$$\mathrm{M(U)}=\mathrm{M}(a)+\mathrm{M}(b)+\mathrm{M}(c)+\&\text{c.}=\Sigma\mathrm{M}(a),$$

$$\mathrm{M(U^2)}=\mathrm{M}(a^2)+\mathrm{M}(b^2)+\mathrm{M}(c^2)+\ldots+2\Sigma\{\mathrm{M}(a)\mathrm{M}(b)\},$$

or

$$\mathrm{M(U^2)}=\Sigma\mathrm{M}(a^2)+\{\Sigma\mathrm{M}(a)\}^2-\Sigma\{\mathrm{M}(a)\}^2.$$

Let us suppose all the infinitesimal simple Errors which it is proposed to combine to be successively superposed upon some assumed function of Error $y=f(x)$; then by equation (9) the new function arising from the first of them will be, putting $D=\frac{d}{dx}$,

$$y=\left(1-\alpha D+\frac{\lambda}{2}D^2\right)f(x).$$

If another be now superposed upon this, we shall have

$$y=\left(1-\beta D+\frac{\mu}{2}D^2\right)\left(1-\alpha D+\frac{\lambda}{2}D^2\right)f(x),$$

and finally the function arising from the superposition of all the given Errors upon the assumed Error $y=f(x)$ will be

$$y=\left(1-\alpha D+\frac{\lambda}{2}D^2\right)\left(1-\beta D+\frac{\mu}{2}D^2\right)\left(1-\gamma D+\frac{\nu}{2}D^2\right)\ldots.f(x). \quad . \quad . \quad (13)$$

But as α, λ are infinitesimals, we have, retaining the square of α,

$$1-\alpha D+\frac{\lambda}{2}D^2=e^{-\alpha D+\frac{1}{2}(\lambda-\alpha^2)D^2}.$$

Thus (13) will become

$$y=e^{-(\alpha+\beta+\gamma+\ldots.)D+\frac{1}{2}(\lambda-\alpha^2+\mu-\beta^2+\ldots.)D^2}f(x),$$

or, adopting the notation (11),

$$y=e^{\frac{1}{2}(h-i)D^2}e^{-mD}f(x)=e^{\frac{1}{2}(h-i)D^2}f(x-m). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

9. Let us now take as the assumed function of Error

$$y=f(x)=\frac{N}{\theta\sqrt{\pi}}e^{-\frac{x^2}{\theta^2}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

(where N is the number of observations), and imagine the whole given system of small Errors superposed upon it; the resulting function is

$$y=\frac{N}{\theta\sqrt{\pi}}e^{\frac{1}{2}(h-i)D^2}e^{-\frac{(x-m)^2}{\theta^2}}.$$

Now by a theorem in the Differential Calculus*,

$$e^{aD^2}e^{-kx^2}=\frac{1}{\sqrt{1+4ak}}e^{-\frac{kx^2}{1+4ak}};$$

* This theorem, which is new to the present writer, may be proved in various ways. Thus if we put

$$u=e^{aD^2}e^{-kx^2},$$

and differentiate with regard to a, we have

$$\frac{du}{da}=e^{aD^2}D^2e^{-kx^2}=e^{aD^2}(4k^2x^2-2k)e^{-kx^2}:$$

again,

$$\frac{du}{dk}=e^{aD^2}(-x^2e^{-kx^2});$$

we thus obtain the partial differential equation

$$\frac{du}{da}+4k^2\frac{du}{dk}+2ku=0,$$

hence

$$y=\frac{\mathrm{N}}{\sqrt{\pi}\sqrt{2(h-i)+\theta^2}}e^{-\frac{(x-m)^2}{2(h-i)+\theta^2}}.$$

Now we may here assume θ as small as we please*,—that is, we may assume the Error (15) upon which the given system was superposed, to be of as small importance as we please. We conclude, then, rejecting this Error altogether, that a system of very small Errors, when combined, give for the resulting function of Error

$$y=\frac{\mathrm{N}}{\sqrt{2\pi(h-i)}}e^{-\frac{(x-m)^2}{2(h-i)}}$$

as before.

---

the integral of which is

$$u=k^{-\frac{1}{2}}\phi\left(4a+\frac{1}{k}\right).$$

To determine the arbitrary function ϕ, we remark that if $a=0$, $u=e^{-kx^2}$,

$$\therefore \phi\left(\frac{1}{k}\right)=k^{\frac{1}{2}}e^{-kx^2},$$

hence

$$u=k^{-\frac{1}{2}}\left(4a+\frac{1}{k}\right)^{-\frac{1}{2}}e^{-x^2\left(4a+\frac{1}{k}\right)^{-1}}=(1+4ak)^{-\frac{1}{2}}e^{-\frac{kx^2}{1+4ak}}.$$

Another proof may be obtained by employing Poisson's ingenious transformation (Traité de Mécanique, tom. ii. p. 356), which gives

$$e^{a\mathrm{D}^2}\phi(x)=\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}e^{-\omega^2}\phi(x+2\omega\sqrt{a})d\omega.$$

* In order that we may retain the three first terms only in the expansion

$$y=f(x)-\alpha f'(x)+\frac{\lambda}{2}f''x-\&c.,$$

it is necessary to show that $f'''(x)$ and the succeeding differential coefficients are not infinite. Now they generally will be infinite in the case where $y=f(x)$ is an infinitesimal Error, as $f(x)$ will be of the form $\mathrm{K}\phi\left(\frac{x}{\epsilon}\right)$, where ϵ is infinitesimal; but in the case where

$$y=f(x)=\frac{\mathrm{N}}{\theta\sqrt{\pi}}e^{-\frac{x^2}{\theta^2}},$$

we may take θ as small as we please, and yet retain only the three first terms above, because the differential coefficients of y do not here become infinite; in fact it is easy to see that any differential coefficient $\frac{d^ny}{dx^n}$ will consist of a series of terms of the form

$$\mathrm{C}\frac{x^\rho}{\theta^r}e^{-\frac{x^2}{\theta^2}}:$$

now by the rules in the Differential Calculus for evaluating indeterminate forms, this quantity tends to zero as θ diminishes.