PREMIER SUPPLEMENT.
SUR L'APPLICATION DU CALCUL DES PROBABILITÉS A LA PHILOSOPHIE NATURELLE

Pierre Simon Laplace*

1816. OC **7** pp. 497–530

The phenomena of nature are most often enveloped in so many strange circumstances, so great a number of perturbing causes mix their influence, that it is very difficult to recognize them. We are able to arrive there only by multiplying the observations or experiments, so that the strange effects coming to be destroyed reciprocally, the mean results set into evidence these phenomena and their diverse elements. The more the observations are numerous and the less they deviate among themselves, the more their results approach to the truth. We fulfill this last condition by the choice of methods, by the precision of the instruments and by the care that we take to observe well. Next we determine through the theory of probabilities the most advantageous mean results or those which lay us less open to error. But this does not suffice; it is more necessary to estimate the probability that the errors of these results are comprehended within some given limits. Without this, we have only an imperfect knowledge of the degree of exactitude obtained. Some formulas proper to this object are therefore a true perfection of the method of the sciences, and so it is quite important to add to this method. The analysis that they require is the most delicate and the most difficult of the theory of probabilities. This is one of the things that I have had principally in view in my Work, in which I am arrived to some formulas of this kind, which have the remarkable advantage to be independent of the law of probabilities of the errors and to contain only quantities given by the observations themselves and by their expressions. I am going to recall here the principles. [497] [498]

Each observation has for analytic expression a function of the elements that we wish to determine; and, if these elements are nearly known, this function becomes a linear function of their corrections. By equating it to the same observation, we form that which we name the *equation of condition*. If we have a great number of similar equations, we combine them, in a manner to obtain as many final equations as there are elements of which we determine next the corrections, by resolving these equations. But what is the most advantageous manner to combine the equations of condition in order to obtain the final equations? What is the law of the errors of which the elements that we deduce from it are yet susceptible? It is this that the theory of probabilities makes known. The formation of a final equation, by means of the equations of con-

*Translated by Richard J. Pulskamp, Department of Mathematics & Computer Science, Xavier University, Cincinnati, OH. January 6, 2014

dition, revert to multiplying each of these by an indeterminate factor and to reunite these products; but it is necessary to choose the system of factors which give the least error to fear. Now it is clear that, if we multiply the possible errors of an element by their respective probabilities, the most advantageous system will be the one in which the sum of these products, all taken positively, is a minimum; because a positive or negative error must be considered as a loss. By forming therefore this sum of products, the condition of the minimum will determine the system of factors that it is necessary to choose. We find thus that this system is the one of the coefficients of the elements in each equation of condition, so that we form a first final equation by multiplying respectively each equation of condition by its coefficient of the first element and by reuniting all these equations thus multiplied. We form a second final equation by employing likewise the coefficients of the second element, and thus consecutively. In this manner, the elements and the laws of the phenomena, contained in the compilation of a great number of observations, are developed with the most evidence. I have given, in § 21 of Book II of my *Théorie analytique des Probabilités*, the expression of the mean error to fear respecting each element. This expression gives the probability of the errors of [499]
which the element is further susceptible, and which is proportional to the number of which the hyperbolic logarithm is unity, raised to a power equal to the square of the error taken to less and divided by the square of the double of this expression and by the ratio of the circumference to the diameter. The coefficient of the negative square of the error in this exponent is able therefore to be considered as the modulus of the probability of the errors, since the error remaining the same, the probability decreases with rapidity when it increases, so that the result obtained weighs, if I am able to say so, toward the truth, so much more as this modulus is greater. I will name, for this reason, this modulus *weight* of the result. By a remarkable analogy of these weights with those of bodies compared to their common center of gravity, it happens that, if one same element is given by diverse systems, each composed of a great number of observations, the most advantageous mean result of them altogether is the sum of the products of each partial result by its weight, this sum being divided by the sum of all the weights. Moreover the total weight of the diverse systems is the sum of their partial weights, so that the probability of the mean result of them altogether is proportional to the number which has unity for hyperbolic logarithm, raised to a power equal to the square of the error, taken to less and multiplied by the sum of all the weights. Each weight depends, in truth, on the law of probability of the errors in each system, and nearly always this law is unknown; but I am happily arrived to eliminate the factor which contains it, by means of the sum of the squares of the deviations of the observations of the system, from their mean result. It would be therefore to desire, in order to complete our understandings on the results obtained by the totality of a great number of observations, that we wrote beside each result the weight which corresponds to it. In order to facilitate the calculation of these weights, I develop its analytic expression, when we have no more than three elements to determine. But, this expression becoming more and more complicated in measure as the number of elements increase, I give a quite simple way in order to determine the weight of a result, whatever be the number [500]
of elements. When we have obtained thus the exponential which represents the law of probability of the errors, we will have the probability that the error of the result is comprehended within some given limits, by taking, within these limits, the integral of

the product of this exponential by the differential of the error and by multiplying it by the square root of the weight of the result divided by the circumference of which the diameter is unity. Thence it follows that, for one same probability, the errors of the results are reciprocals to the square roots of their weights, that which is able to serve to compare their respective precision.

In order to apply this method with success, it is necessary to vary the circumstances of the observations or of the experiments, in a manner to avoid the constant causes of the error. It is necessary that the observations be numerous and that they be so many more as there are more elements to determine: because the weight of the mean result increases as the number of the observations divided by the number of elements. It is further necessary that the elements follow, in these observations, a different march; because, if the march of the two elements were rigorously the same, that which would render their coefficients proportionals in the equations of condition, these elements would form only a single unknown, and it would be impossible to distinguish them by these observations. Finally it is necessary that the observations be precise. This condition, the first of all, increases much the weight of the result, of which the expression has for divisor the sum of the squares of their deviations from this result. With these precautions, we will be able to make use of the preceding method and to measure the degree of confidence which the results deduced from a great number of observations merit.

§ 1. A great advantage of this method, which permits evaluating from it the expressions numerically, is, as we have said, to be independent of the law of probability of the errors of the observations. The factor $\frac{2k''}{k}a^2s$, which depends on this law, has been eliminated from the formulas of §§ 19 and 21 of Book II, by observing that this factor which is the sum of the squares of all the possible errors of the observations, multi- [501]
plied by their respective probabilities, and which expresses thus the true mean of these squares, is very probably equal to the sum of the squares of the rest of the equations of condition, when we have substituted the elements determined by the most advantageous method. The importance of this method in natural philosophy requires that the uncertainty that it is able to permit is dissipated, and the only one which remains yet is relative to the equality of which I just spoke. I will first clarify this delicate point of the theory of the probabilities and show that the preceding equality is able to be employed without sensible error.

The sum of the squares of the errors of the observations, of which the number is s, being supposed equal to $\frac{2k''}{k}a^2s + a^2r\sqrt{s}$, the probability that the value of r is comprehended within the given limits is, by § 19 cited,

$$\frac{1}{2\sqrt{\pi}}\int \beta' dr\, c^{-\frac{\beta'^2 r^2}{4}},$$

the integral being taken within these limits. Let us represent the general equation of condition of the elements $z,\ z',\ldots$ by this one

$$\epsilon^{(i)} = p^{(i)}z + q^{(i)}z' + \cdots - \alpha^{(i)},$$

$\epsilon^{(i)}$ being the error of the observation. The elements $z,\ z',\ldots$ being determined by the most advantageous method, let us designate by $u,\ u',\ \ldots$ their errors; we will have, by naming $\epsilon'^{(i)}$ the remainder of the function

$$p^{(i)}z + q^{(i)}z' + \cdots - \alpha^{(i)}$$

when we have substituted for z, z', ... their values thus determined,

$$\epsilon^{(i)} = \epsilon'^{(i)} + p^{(i)}u + q^{(i)}u' + \cdots ,$$

that which gives

$$\mathrm{S}\epsilon^{(i)2} = \mathrm{S}\epsilon'^{(i)2} + 2\mathrm{S}\epsilon'^{(i)}(p^{(i)}u + q^{(i)}u' + \ldots) + \mathrm{S}(p^{(i)}u + q^{(i)}u' + \ldots)^2,$$

the integral sign S being extended to all the values of i, from $i = 0$ to $i = s - 1$. But, by the conditions of the most advantageous method, we have [502]

$$\mathrm{S}p^{(i)}\epsilon'^{(i)} = 0, \quad \mathrm{S}q^{(i)}\epsilon'^{(i)} = 0, \quad \ldots;$$

we have therefore

$$\mathrm{S}\epsilon^{(i)^2} = \mathrm{S}\epsilon'^{(i)^2} + \mathrm{S}(p^{(i)}u + q^{(i)}u' + \cdots)^2;$$

by comparing this value of $\mathrm{S}\epsilon^{(i)^2}$ to its preceding value $\frac{2k''}{k}a^2s + a^2r\sqrt{s}$, we will have

$$a^2r\sqrt{s} = \mathrm{S}\epsilon'^{(i)^2} - \frac{2k''}{k}a^2s + \mathrm{S}(p^{(i)}u + q^{(i)}u' + \cdots)^2.$$

Let us make

$$\mathrm{S}\epsilon'^{(i)^2} - \frac{2k''}{k}a^2s = t\sqrt{s},$$

$$u = \frac{\nu}{\sqrt{s}}, \quad u' = \frac{\nu'}{\sqrt{s}}, \quad u'' = \frac{\nu''}{\sqrt{s}}, \quad \ldots,$$

we will have

$$a^2r = t + \frac{\mathrm{S}(p^{(i)}\nu + q^{(i)}\nu' + \cdots)^2}{s\sqrt{s}};$$

the exponential $c^{-\frac{\beta'^2r^2}{4}}$ becomes thus

$$c^{-\frac{\beta'^2}{4a^2}\left[t+\frac{\mathrm{S}(p^{(i)}\nu+q^{(i)}\nu'+\cdots)^2}{s\sqrt{s}}\right]^2};$$

thus the probability of t is proportional to this exponential.

The analysis of § 21 of Book II leads to this general theorem, namely that the probability of the simultaneous existence of the quantities u, u', u'', ... is proportional to the exponential

$$c^{-\frac{k}{4k''a^2s}\mathrm{S}(p^{(i)}\nu+q^{(i)}\nu'+\cdots)^2};$$

the probability of the simultaneous existence of t, ν, ν', ν'', ... is therefore proportional to

$$c^{-\frac{\beta'^2}{4a^4}\left[t+\frac{\mathrm{S}(p^{(i)}\nu+q^{(i)}\nu'+\cdots)^2}{s\sqrt{s}}\right]^2-\frac{k}{4k''a^2s}\mathrm{S}(p^{(i)}\nu+q^{(i)}\nu'+\cdots)^2}.$$

By substituting for $\frac{4k''a^2s}{k}$ its value $2\text{S}\epsilon'^{(i)} - 2t\sqrt{s}$, this exponential is reduced, by neglecting the terms of order $\frac{1}{s}$, in the following function: [503]

$$\left[1 - \frac{t\sqrt{s}}{2(\text{S}\epsilon'^{(i)2})^2}\text{S}(p^{(i)}\nu + q^{(i)}\nu' + \cdots)^2\right] c^{-\frac{\beta'^2}{4a^2}\left[t+\frac{\text{S}(p^{(i)}\nu+q^{(i)}\nu'+\cdots)^2}{s\sqrt{s}}\right]^2 - \frac{\text{S}(p^{(i)}\nu+q^{(i)}\nu'+\cdots)^2}{2\text{S}\epsilon'^{(i)2}}}$$

Now, in order to have the probability that the value of ν is comprehended within some given limits, it is necessary: 1° to multiply this function by $dt\, d\nu\, d\nu' \ldots$; 2° to take the integral of the product for all the possible values of t, ν', ν'', ... and, with respect to ν, to integrate only within the given limits; 3° to divide the whole by this same integral taken with respect to all the possible values of t, ν, ν', ... By regarding $\text{S}\epsilon'^{(i)2}$ as a datum from observation, t varies only at the rate of the unknown value of $\frac{2k''a^2s}{k}$, and this value is able to vary from zero to infinity; t is therefore able to vary from $\frac{\text{S}\epsilon'^{(i)2}}{\sqrt{s}}$ to negative infinity; and, as $\text{S}\epsilon'^{(i)2}$ is of the order of s, t is able to vary from negative infinity to a positive value of order $\sqrt{s}$. The preceding exponential becomes, at that limit of the integral taken with respect to t, of the form c^{-Q^2s}, and will be able to be supposed null, because of the magnitude of s. Thus we are able to take the integral relative to t, from $t = -\infty$ to $t = \infty$. Similarly the integrals relative to ν' ν'', ... are able to be taken within the same limits. If we make

$$t + \frac{\text{S}(p^{(i)}\nu + q^{(i)}\nu' + \cdots)^2}{s\sqrt{s}} = t',$$

the integral relative to t' will be able to be taken with respect to t' from $t' = -\infty$ to $t' = \infty$.

Thence it is easy to conclude that the probability that ν is comprehended within the given limits is proportional to the integral

$$\int d\nu d\nu' \ldots \left\{1 + \frac{\left[\text{S}(p^{(i)}\nu + q^{(i)}\nu' + \cdots)^2\right]^2}{(2\text{S}\epsilon'^{(i)2})^2 s}\right\} c^{-\frac{\text{S}(p^{(i)}\nu+q^{(i)}\nu'+\cdots)^2}{2\text{S}\epsilon^{(i)2}}},$$

the integral being taken from ν', ν'', ... equal to $-\infty$ to their positive infinite values and with respect to ν within the given limits, and being divided by the same integral extended to the positive and negative infinite values of ν, ν', ν'', ... [504]

The consideration of the difference which is able to exist between $\frac{2k''}{k}a^2s$ and $\text{S}\epsilon'^{(i)2}$ introduces therefore into the expression of the probability of which there is concern only one term of order $\frac{1}{s}$, an order that I myself am permitted to neglect in my Work. Thence, the preceding integral becomes

$$\int d\nu d\nu' \ldots c^{-\frac{\text{S}(p^{(i)}\nu+q^{(i)}\nu'+\ldots)^2}{2\text{S}\epsilon'^{(i)2}}}.$$

If we make

$$p_1^{(i)} = p^{(i)} - \frac{q^{(i)}\text{S}p^{(i)}q^{(i)}}{\text{S}q^{(i)2}},$$
$$r_1^{(i)} = r^{(i)} - \frac{q^{(i)}\text{S}r^{(i)}q^{(i)}}{\text{S}q^{(i)2}},$$
$$t_1^{(i)} = t^{(i)} - \frac{q^{(i)}\text{S}t^{(i)}q^{(i)}}{\text{S}q^{(i)2}},$$
$$\dots\dots\dots\dots\dots\dots\dots$$

the exponential

$$c^{-\frac{\text{S}(p^{(i)}\nu+q^{(i)}\nu'+r^{(i)}\nu''\dots)^2}{2\text{S}\epsilon'^{(i)2}}}$$

will be able to be set under this form

$$c^{-\frac{\text{S}(p_1^{(i)}\nu+r_1^{(i)}\nu''\dots)^2}{2\text{S}\epsilon'^{(i)2}}-\frac{\text{S}q^{(i)2}}{2\text{S}\epsilon'^{(i)2}}\left(\nu'+\frac{\nu\text{S}p^{(i)}q^{(i)}+\nu''\text{S}r^{(i)}q^{(i)}+\dots}{\text{S}q^{(i)2}}\right)^2}.$$

By multiplying this quantity by $d\nu'$, and by integrating it from $\nu' = -\infty$ to $\nu' = \infty$, we will have a quantity proportional to

$$c^{-\frac{\text{S}(p_1^{(i)}\nu+r_1^{(i)}\nu''\dots)^2}{2\text{S}\epsilon'^{(i)2}}}$$

and in which the variable ν' has disappeared. By following the same process, we will make the variables ν'', ν''', ... vanish. We will arrive thus to an exponential of the form $c^{-\frac{\nu^2\text{S}p_{n-1}^{(i)2}}{2\text{S}\epsilon'^{(i)2}}}$, n being the number of elements. If we restore, instead of ν, its value [505] $u\sqrt{s}$, this exponential becomes

$$c^{-Pu^2},$$

by making

$$P = \frac{s\text{S}p^{(i)2}}{2\text{S}\epsilon'^{(i)2}}.$$

u being the error of the value of z, P is that which I name *weight* of this value. The probability that this error is comprehended within some given limits is therefore

$$\frac{\int du\sqrt{P}c^{P-u^2}}{\sqrt{\pi}},$$

the integral being taken within these limits, and π being the circumference of which the diameter is unity. But it is simpler to apply the process of which we have just made use to the final equations which determine the elements, in order to reduce them to one alone, that which gives an easy method to resolve these equations.

§ 2. Let us take the general equation of condition, and, for more simplicity, let us limit it to the six elements z, z', z'', z''', z^{iv}, z^{v}; it becomes then

(1) $$\epsilon^{(i)} = p^{(i)}z + q^{(i)}z' + r^{(i)}z'' + t^{(i)}z''' + \gamma^{(i)}z^{\text{iv}} + \lambda^{(i)}z^{\text{v}} - \alpha^{(i)}.$$

By multiplying it by $\lambda^{(i)}$ and reuniting all the products together, we will have

$$\mathrm{S}\lambda^{(i)}\epsilon^{(i)} = z\mathrm{S}\lambda^{(i)}p^{(i)} + z'\mathrm{S}\lambda^{(i)}q^{(i)} + \cdots - \mathrm{S}\lambda^{(i)}\alpha^{(i)},$$

the integral sign S extending to all the values of i, from $i = 0$ to $i = s - 1$, s being the number of observations employed. By the conditions of the most advantageous method, we have $\mathrm{S}\lambda^{(i)}\epsilon^{(i)} = 0$; the preceding equation will give therefore

$$\begin{aligned} z^{\mathrm{v}} = -z^{\mathrm{iv}}\frac{\mathrm{S}\lambda^{(i)}\gamma^{(i)}}{\mathrm{S}\lambda^{(i)2}} - z'''\frac{\mathrm{S}\lambda^{(i)}t^{(i)}}{\mathrm{S}\lambda^{(i)2}} - z''\frac{\mathrm{S}\lambda^{(i)}r^{(i)}}{\mathrm{S}\lambda^{(i)2}} \\ -z'\frac{\mathrm{S}\lambda^{(i)}q^{(i)}}{\mathrm{S}\lambda^{(i)2}} - z\frac{\mathrm{S}\lambda^{(i)}p^{(i)}}{\mathrm{S}\lambda^{(i)2}} + \frac{\mathrm{S}\lambda^{(i)}\alpha^{(i)}}{\mathrm{S}\lambda^{(i)2}} \end{aligned}$$

If we substitute this value into equation (1) and if we make [506]

$$\begin{aligned} \gamma_1^{(i)} &= \gamma^{(i)} - \lambda^{(i)}\frac{\mathrm{S}\lambda^{(i)}\gamma^{(i)}}{\mathrm{S}\lambda^{(i)2}}, \\ t_1^{(i)} &= t^{(i)} - \lambda^{(i)}\frac{\mathrm{S}\lambda^{(i)}t^{(i)}}{\mathrm{S}\lambda^{(i)2}}, \\ r_1^{(i)} &= r^{(i)} - \lambda^{(i)}\frac{\mathrm{S}\lambda^{(i)}r^{(i)}}{\mathrm{S}\lambda^{(i)2}}, \\ q_1^{(i)} &= q^{(i)} - \lambda^{(i)}\frac{\mathrm{S}\lambda^{(i)}q^{(i)}}{\mathrm{S}\lambda^{(i)2}}, \\ p_1^{(i)} &= p^{(i)} - \lambda^{(i)}\frac{\mathrm{S}\lambda^{(i)}p^{(i)}}{\mathrm{S}\lambda^{(i)2}}, \\ \alpha_1^{(i)} &= \alpha^{(i)} - \lambda^{(i)}\frac{\mathrm{S}\lambda^{(i)}\alpha^{(i)}}{\mathrm{S}\lambda^{(i)2}}, \end{aligned}$$

we will have

(2) $$\epsilon^{(i)} = p_1^{(i)}z + q_1^{(i)}z' + r_1^{(i)}z'' + t_1^{(i)}z''' + \gamma_1^{(i)}z^{\mathrm{iv}} - \alpha_1^{(i)};$$

by this means, the element z^{v} has disappeared from the equations of condition that equation (2) represents. By multiplying this equation by $\gamma_1^{(i)}$ and reuniting all the products together, by observing next that we have

$$\mathrm{S}\gamma_1^{(i)}\epsilon^{(i)} = 0$$

by virtue of the equations

$$0 = \mathrm{S}\lambda^{(i)}\epsilon^{(i)}, \quad 0 = \mathrm{S}\gamma^{(i)}\epsilon^{(i)}$$

that the conditions of the most advantageous method give, we will have

$$0 = z\mathrm{S}\gamma_1^{(i)}p_1^{(i)} + z'\mathrm{S}\gamma_1^{(i)}q_1^{(i)} + z''\mathrm{S}\gamma_1^{(i)}r_1^{(i)} + z'''\mathrm{S}\gamma_1^{(i)}t_1^{(i)} + z^{\mathrm{iv}}\mathrm{S}\gamma_1^{(i)2} - \mathrm{S}\gamma_1^{(i)}\alpha_1^{(i)};$$

whence we deduce

$$z^{\mathrm{iv}} = -z'''\frac{\mathrm{S}\gamma_1^{(i)}t_1^{(i)}}{\mathrm{S}\gamma_1^{(i)2}} - z''\frac{\mathrm{S}\gamma_1^{(i)}r_1^{(i)}}{\mathrm{S}\gamma_1^{(i)2}} - z'\frac{\mathrm{S}\gamma_1^{(i)}q_1^{(i)}}{\mathrm{S}\gamma_1^{(i)2}} - z\frac{\mathrm{S}\gamma_1^{(i)}p_1^{(i)}}{\mathrm{S}\gamma_1^{(i)2}} + \frac{\mathrm{S}\gamma_1^{(i)}\alpha_1^{(i)}}{\mathrm{S}\gamma_1^{(i)2}}.$$

If we substitute this value into equation (2) and if we make [507]

$$t_2^{(i)} = t_1^{(i)} - \gamma_1^{(i)}\frac{\mathrm{S}\gamma_1^{(i)}t^{(i)}}{\mathrm{S}\gamma_1^{(i)2}},$$

$$r_2^{(i)} = r_1^{(i)} - \gamma_1^{(i)}\frac{\mathrm{S}\gamma_1^{(i)}r^{(i)}}{\mathrm{S}\gamma_1^{(i)2}},$$

$$q_2^{(i)} = q_1^{(i)} - \gamma_1^{(i)}\frac{\mathrm{S}\gamma_1^{(i)}q_1^{(i)}}{\mathrm{S}\gamma_1^{(i)2}},$$

$$p_2^{(i)} = p_1^{(i)} - \gamma_1^{(i)}\frac{\mathrm{S}\gamma_1^{(i)}p_1^{(i)}}{\mathrm{S}\gamma_1^{(i)2}},$$

$$\alpha_2^{(i)} = \alpha_1^{(i)} - \gamma_1^{(i)}\frac{\mathrm{S}\gamma_1^{(i)}\alpha_1^{(i)}}{\mathrm{S}\gamma_1^{(i)2}},$$

we will have

(3) $$\epsilon^{(i)} = p_2^{(i)}z + q_2^{(i)}z' + r_2^{(i)}z'' + t_2^{(i)}z''' - \alpha_2^{(i)}.$$

By continuing thus, we will arrive to an equation of the form

(4) $$\epsilon^{(i)} = p_5^{(i)}z - \alpha_5^{(i)}.$$

There results from § 20 of Book II that, if the value of z is determined by this equation and if u is the error of this value, the probability of this error is

$$\sqrt{\frac{s\mathrm{S}p_5^{(i)2}}{2\mathrm{S}\epsilon'^{(i)2}\pi}}c^{-\frac{s\mathrm{S}p_5^{(i)2}}{2\mathrm{S}\epsilon'^{(i)2}}u^2},$$

$\mathrm{S}\epsilon'^{(i)2}$ being the sum of the squares of the remainders of the equations of condition, when we have substituted there the elements determined by the most advantageous method. The weight P of this error is therefore equal to $\frac{s\mathrm{S}p_5^{(i)2}}{2\mathrm{S}\epsilon'^{(i)2}}$.

The concern now is to determine $\mathrm{S}p_5^{(i)2}$. For this, we will multiply respectively each of the equations of condition represented by equation (1) first by the coefficient of the first element, and we will take the sum of these products; next by the coefficient of the second element, and we will take the sum of these products, and thus of the rest. We will have, by observing that by the conditions of the most advantageous method [508]
$\mathrm{S}p^{(i)}\epsilon^{(i)} = 0, \mathrm{S}q^{(i)}\epsilon^{(i)} = 0, \ldots$, the six equations following:

(A) $$\left\{\begin{array}{l} \overline{p\alpha} = p^{(2)}z + \overline{pq}z' + \overline{pr}z'' + \overline{pt}z''' + \overline{p\gamma}z^{\text{iv}} + \overline{p\lambda}z^{\text{v}}, \\ \overline{q\alpha} = \overline{pq}z + q^{(2)}z' + \overline{qr}z'' + \overline{qt}z''' + \overline{q\gamma}z^{\text{iv}} + \overline{q\lambda}z^{\text{v}}, \\ \overline{r\alpha} = \overline{rp}z + \overline{rq}z' + r^{(2)}z'' + \overline{rt}z''' + \overline{r\gamma}z^{\text{iv}} + \overline{r\lambda}z^{\text{v}}, \\ \overline{t\alpha} = \overline{tp}z + \overline{tq}z' + \overline{tr}z'' + t^{(2)}z''' + \overline{t\gamma}z^{\text{iv}} + \overline{t\lambda}z^{\text{v}}, \\ \overline{\gamma\alpha} = \overline{\gamma p}z + \overline{\gamma q}z' + \overline{\gamma r}z'' + \overline{\gamma t}z''' + \gamma^{(2)}z^{\text{iv}} + \overline{\gamma\lambda}z^{\text{v}}, \\ \overline{\lambda\alpha} = \overline{\lambda p}z + \overline{\lambda q}z' + \overline{\lambda r}z'' + \overline{\lambda t}z''' + \overline{\lambda\gamma}z^{\text{iv}} + \lambda^{(2)}z^{\text{v}}, \end{array}\right.$$

whence we must observe that we suppose

$$p^{(2)} = \mathrm{S}p^{(i)2}, \quad \overline{pq} = \mathrm{S}p^{(i)}q^{(i)}, \quad q^{(2)} = \mathrm{S}q^{(i)2}, \quad \overline{qr} = \mathrm{S}q^{(i)}r^{(i)}, \quad \ldots$$

If we multiply similarly the equations of condition represented by equation (2) respectively by the coefficients of z and if we add these products, next by the coefficients of z' by adding again these products, and thus in succession, we will have the following system of equations, by observing that $\mathrm{S}p_1^{(i)}\epsilon^{(i)} = 0, \mathrm{S}q_1^{(i)}\epsilon^{(i)} = 0, \ldots$, by the conditions of the most advantageous method.

$$\text{(B)} \qquad \left\{ \begin{aligned}
\overline{p_1\alpha_1} &= p_1^{(2)}z + \overline{p_1q_1}z' + \overline{p_1r_1}z'' + \overline{p_1t_1}z''' + \overline{p_1\gamma_1}z^{\text{iv}}, \\
\overline{q_1\alpha_1} &= \overline{p_1q_1}z + q_1^{(2)}z' + \overline{q_1r_1}z'' + \overline{q_1t_1}z''' + \overline{q_1\gamma_1}z^{\text{iv}}, \\
\overline{r_1\alpha_1} &= \overline{p_1r_1}z + \overline{q_1r_1}z' + r_1^{(2)}z'' + \overline{r_1t_1}z''' + \overline{r_1\gamma_1}z^{\text{iv}}, \\
\overline{t_1\alpha_1} &= \overline{p_1t_1}z + \overline{q_1t_1}z' + \overline{r_1t_1}z'' + t_1^{(2)}z''' + \overline{t_1\gamma_1}z^{\text{iv}}, \\
\overline{\gamma_1\alpha_1} &= \overline{p_1\gamma_1}z + \overline{q_1\gamma_1}z' + \overline{r_1\gamma_1}z'' + \overline{t_1\gamma_1}z''' + \gamma_1^{(2)}z^{\text{iv}},
\end{aligned} \right.$$

whence we must observe that

$$\overline{p_1q_1} = \mathrm{S}p_1^{(i)}q_1^{(i)}, \quad p_1^{(2)} = \mathrm{S}p_1^{(i)2}, \quad \ldots$$

By substituting, instead of $p_1^{(i)}$, $q_1^{(i)}$, ..., their preceding values, we have

$$\overline{p_1q_1} = \mathrm{S}p_1^{(i)}q_1^{(i)} - \frac{\mathrm{S}\lambda^{(i)}p^{(i)}\mathrm{S}\lambda^{(i)}q^{(i)}}{\mathrm{S}\lambda^{(i)2}}$$

or

$$\overline{p_1q_1} = \overline{pq} - \frac{\overline{\lambda p}\,\overline{\lambda q}}{\lambda^{(2)}};$$

we have similarly [509]

$$p_1^{(2)} = p^{(2)} - \frac{\overline{\lambda p}^2}{\lambda^{(2)}},$$

$$q_1^{(2)} = q^{(2)} - \frac{\overline{\lambda q}^2}{\lambda^{(2)}},$$

$$\overline{p_1r_1} = \overline{pr} - \frac{\overline{\lambda p}\,\overline{\lambda r}}{\lambda^{(2)}},$$

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

$$\overline{p_1\alpha_1} = \overline{p\alpha} - \frac{\overline{\lambda p}\,\overline{\lambda\alpha}}{\lambda^{(2)}},$$

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

Thus the coefficients of the system of equations (B) are deduced easily from the coefficients of the system of equations (A).

The equations of condition represented by equation (3) will give similarly the following system of equations

$$\text{(C)} \qquad \left\{ \begin{aligned}
\overline{p_2\alpha_2} &= p_2^{(2)}\,z + \overline{p_2q_2}\,z' + \overline{p_2r_2}\,z'' + \overline{p_2t_2}\,z''', \\
\overline{q_2\alpha_2} &= \overline{p_2q_2}\,z + q_2^{(2)}\,z' + \overline{q_2r_2}\,z'' + \overline{q_2t_2}\,z''', \\
\overline{r_2\alpha_2} &= \overline{p_2r_2}\,z + \overline{q_2r_2}\,z' + r_2^{(2)}\,z'' + \overline{r_2t_2}\,z''', \\
\overline{t_2\alpha_2} &= \overline{p_2t_2}\,z + \overline{q_2t_2}\,z' + \overline{r_2t_2}\,z'' + t_2^{(2)}\,z''',
\end{aligned} \right.$$

and we have

$$p_2^{(2)} = p_1^{(2)} - \frac{\overline{\gamma_1 p_1}^2}{\gamma_1^{(2)}},$$

$$\overline{p_2 q_2} = \overline{p_1 q_1} - \frac{\overline{\gamma_1 p_1}\,\overline{q_1 \gamma_1}}{\gamma_1^{(2)}},$$

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

$$\overline{p_2 \alpha_2} = \overline{p_1 \alpha_1} - \frac{\overline{\gamma_1 p_1}\,\overline{\gamma_1 \alpha_1}}{\gamma_1^{(2)}},$$

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

We will have similarly the system of equations

(D) $$\begin{cases} \overline{p_3\alpha_3} = p_3^{(2)}\, z + \overline{p_3 q_3}\, z' + \overline{p_3 r_3}\, z'', \\ \overline{q_3\alpha_3} = \overline{p_3 q_3}\, z + q_3^{(2)}\, z' + \overline{q_3 r_3}\, z'', \\ \overline{r_3\alpha_3} = \overline{p_3 r_3}\, z + \overline{q_3 r_3}\, z' + r_3^{(2)}\, z'', \end{cases}$$

by making [510]

$$p_3^{(2)} = p_2^{(2)} - \frac{\overline{p_2 t_2}^2}{t_2^{(2)}},$$

$$\overline{p_3 q_3} = \overline{p_2 q_2} - \frac{\overline{p_2 t_2}\,\overline{q_{21} t_2}}{t_2^{(2)}},$$

$$\overline{p_3 \alpha_3} = \overline{p_2 \alpha_2} - \frac{\overline{t_2 p_2}\,\overline{t_2 \alpha_2}}{t_2^{(2)}},$$

$$\cdots\cdots\cdots\cdots\cdots ;$$

we will have further

(E) $$\begin{cases} \overline{p_4\alpha_4} = p_4^{(2)} z + \overline{p_4 q_4} z', \\ \overline{q_4\alpha_4} = \overline{p_4 q_4} z + q_4^{(2)} z', \end{cases}$$

by making

$$p_4^{(2)} = p_3^{(2)} - \frac{\overline{p_3 r_3}^2}{r_3^{(2)}},$$

$$\overline{p_4 q_4} = \overline{p_3 q_3} - \frac{\overline{p_3 r_3}\,\overline{q_3 r_3}}{r_3^{(2)}},$$

$$\overline{p_4 \alpha_4} = \overline{p_3 \alpha_3} - \frac{\overline{p_3 r_3}\,\overline{\alpha_3 r_3}}{r_3^{(2)}},$$

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

Finally we will have

(F) $$\overline{p_5 \alpha_5} = p_5^{(2)} z,$$

by making

$$p_5^{(2)} = p_4^{(2)} - \frac{\overline{p_4q_4}^2}{q_4^{(2)}}, \quad \overline{p_5\alpha_5} = \overline{p_4\alpha_4} - \frac{\overline{p_4q_4}\,\overline{q_4\alpha_4}}{q_4^{(2)}};$$

$p_5^{(2)}$ is the value of $\mathrm{S}p_5^{(i)2}$, and the weight P will be

$$\frac{sp_5^{(2)}}{2\mathrm{S}\epsilon'^{(i)2}}.$$

We see by the sequence of the values of $p^{(2)}$, $p_1^{(2)}$, $p_2^{(2)}, \dots$ that they diminish without ceasing, and that thus, for the same number of observations, the weight P diminishes when the number of elements increase.

If we consider the sequence of equations which determine $\overline{p_5\alpha_5}$, we see that this function, developed according to the coefficients of the system of equations (A), is of the form [511]

$$\overline{p\alpha} + M\overline{q\alpha} + N\overline{r\alpha} + \cdots,$$

the coefficient of $\overline{p\alpha}$ being unity. It follows thence that if we resolve equations (A), by leaving $\overline{p\alpha}$, $\overline{q\alpha}$, $\overline{r\alpha}$, ... as indeterminates, $\frac{1}{p_5^{(2)}}$ will be, by virtue of equation (F), the coefficient of $\overline{p\alpha}$ in the expression of z. Similarly, $\frac{1}{q_5^{(2)}}$ will be the coefficient of $\overline{q\alpha}$ in the expression of z'; $\frac{1}{r_5^{(2)}}$ will be the coefficient of $\overline{r\alpha}$ in the expression of z''; and thus of the rest; that which gives a simple means to obtain $p_5^{(2)}$, $q_5^{(2)}$, ...; but it is simpler yet to determine them thus.

First equation (F) gives the value of $p_5^{(2)}$ and of z. If in the system of equations (E) we eliminate z instead of z', we will have a single equation in z', of the form

$$\overline{q_5\alpha_5} = q_5^{(2)}z';$$

by making

$$q_5^{(2)} = q_4^{(2)} - \frac{\overline{p_4q_4}^2}{p_4^{(2)}}, \quad \overline{q_5\alpha_5} = \overline{q_4\alpha_4} - \frac{\overline{p_4q_4}\,\overline{p_4\alpha_4}}{p_4^{(2)}}.$$

If in the system of equations (D) we eliminate z instead of z'', in order to conserve at the end of the calculation only z'', we will have $r_5^{(2)}$ by changing in the sequence of equations which, departing from this system, determine $p_5^{(2)}$, the letter p into the letter r, and reciprocally. We will have thus

$$r_4^{(2)} = r_3^{(2)} - \frac{\overline{p_3r_3}^2}{p_3^{(2)}},$$

$$\overline{r_4q_4} = \overline{r_3q_3} - \frac{\overline{p_3q_3}\,\overline{p_3r_3}}{p_3^{(2)}},$$

$$q_4^{(2)} = q_3^{(2)} - \frac{\overline{p_3q_3}^2}{p_3^{(2)}},$$

$$r_5^{(2)} = r_4^{(2)} - \frac{\overline{p_4q_4}^2}{q_4^{(2)}},$$

....

In order to have $t_5^{(2)}$, we will depart from the system of equations (C), by changing, in the sequence of the values $p_3^{(2)}, \overline{p_3 q_3}, \ldots, r_3^{(2)}, \overline{q_3 r_3}, \ldots$, the letter p into the letter t, and reciprocally. [512]

We will have similarly the value of $\gamma_5^{(2)}$, by departing from the system of equations (B) and changing in the sequence of values of $p_2^{(2)}, p_3^{(2)}, \ldots$, the letter p into the letter γ, and reciprocally.

Finally, we will have the value of $\lambda_5^{(2)}$ by changing, in the sequence of values of $p_1^{(2)}, p_2^{(2)}, \ldots,$ the letter p into the letter λ, and reciprocally.

§ 3. The error of which the value of z is susceptible being u, its probability is, as we have seen,

$$\frac{\sqrt{P} c^{-Pu^2}}{\sqrt{\pi}}.$$

By multiplying it by $u\,du$ and taking the integral from u null to u infinity, we will have

$$\frac{1}{2\sqrt{\pi}\sqrt{P}}$$

for the mean error to fear more respecting the value of z. This expression affected with the sign $-$ will be the mean error to fear to less respecting this value. I have given in § 21 of Book II the analytic expression of these mean errors, whatever be the number of elements. We will have therefore, by comparing it to the preceding, the value of P, and it is easy to recognize the identity of these expressions. We find thus, in the case of a single element

$$P = \frac{sp^{(2)}}{2\text{S}\epsilon'^{(i)2}}.$$

If we make generally, for any number of elements whatsoever,

$$P = \frac{s}{2\text{S}\epsilon'^{(i)2}} \frac{A}{B},$$

we find, for two elements,

$$A = p^{(2)} q^{(2)} - \overline{pq}^2,$$
$$B = q^{(2)}.$$

By applying these results to the equations (E), we will have the value of P relative to the element z. [513]

We find, for three elements,

$$A = p^{(2)} q^{(2)} r^{(2)} - p^{(2)} \overline{qr}^2 - q^{(2)} \overline{pr}^2 - r^{(2)} \overline{pq}^2 + \overline{pq}\,\overline{pr}\,\overline{qr},$$
$$B = q^{(2)} r^{(2)} - \overline{qr}^2.$$

These results applied to the equations (D) will give the value of P relative to the element z.

By continuing thus, we will have, whatever be the number of elements, the weight relative to the first element z. By changing in its expression p into q and q into p, we will have the weight relative to the second element z'. By changing, in the expression of the weight of the first element, p into r and r into p, we will have the weight relative to the third element z'', and thus in succession. But, when the number of elements surpasses three, it is much simpler to make use of the method of the previous section.

We will observe here that the mean error to fear respecting each element being, by §§ 20 and 21 of Book II, smaller in the system of factors which constitute the most advantageous method than in every other system, the value of P is the greatest possible. Thus, for one same error of an element in this method, the probability is smaller than in every other method, that which assures its superiority.

§ 4. All my analysis rests on the hypothesis that the facility of the errors is the same for the positive errors and for the negative errors; that which renders null the integral of the product of the error by its probability and by its differential, the integral being taken in all the extent of the limits of the errors, and the origin of the errors being in the middle of the interval which separates these limits. But, if the law of facility is different for the positive errors and for the negative errors, then the preceding integral [514] becomes null only in the case where this origin is at the point of the abscissa through where passes the ordinate of the center of gravity of the curve, of which the ordinates represent the law of facility of the errors represented themselves by the abscissas. For every other point, the mean error of the observation is this integral divided by the interval of the limits; and, if we have a great number of observations, the mean of the errors of these observations will be, by that which we have seen in Book II, equal very nearly to this quotient. By making therefore so that the sum of the errors is null, we will be able to suppose null the integral of which we just spoke, and then all my analysis subsists and becomes independent of the hypothesis of an equal facility of positive errors and of negative errors. We are able always to obtain this advantage by adding to the equations of condition an indeterminate element of which the coefficient is unity. It is that which takes place of itself in the equations of condition relative to the movement of the planets in longitude; because the correction of the epoch has unity there for coefficient. But, the addition of an element weakening, as we have said, the probability of the errors of the other elements, a probability which, for the same number of observations, diminishes when the number of elements which are supported on them is greater, it is necessary to recur to this addition only when we are able to fear that a constant cause favors the errors of one sign rather than those of a contrary sign. Besides, we will be assured of it easily, by making the sum of the positive remainders and that of the negative remainders of the equations of condition, when we will have substituted the values of the elements determined by the most advantageous method, without the addition of which we just spoke and by seeing if the excess of one of these sums over the other indicates a constant cause.

In order to leave no doubt on this object, I am going to apply the calculus. There results from § 22 of Book II that the probability that the sum of the errors of the observations equals

$$\frac{ak'}{k}s + ar\sqrt{s}$$

is proportional to the exponential [515

$$c^{-\frac{k^2r^2}{2(kk''-k'^2)}}.$$

This sum is $\mathrm{S}\epsilon^{(i)}$, and, by § 1, we have

$$\mathrm{S}\epsilon^{(i)} = \mathrm{S}\epsilon'^{(i)} + \mathrm{S}(p^{(i)}u + q^{(i)}u' + \cdots).$$

By the nature of the final equations, we have $\mathrm{S}\epsilon'^{(i)} = 0$; we have therefore

$$\frac{ak'}{k}s + ar\sqrt{s} = \mathrm{S}(p^{(i)}u + q^{(i)}u' + \cdots).$$

If we make, as in this section, $u = \frac{\nu}{\sqrt{s}}$, $u' = \frac{\nu'}{\sqrt{s}}$, ..., we will have thus

$$r = -\frac{k'}{k}\sqrt{s} + \frac{1}{as}\mathrm{S}(p^{(i)}\nu + q^{(i)}\nu' + \cdots).$$

Thus $\frac{k'}{k}$ is of order $\frac{1}{\sqrt{s}}$, and its square is of order $\frac{1}{s}$; we are able therefore to neglect it, having regard to $\frac{k''}{k}$. The probability of the simultaneous existence of r, ν, ν', ... is thus proportional to the exponential

$$c^{-\frac{k}{2k''}r^2 - \frac{\mathrm{S}(p^{(i)}\nu + q^{(i)}\nu' + \cdots)^2}{2\mathrm{S}\epsilon'^{(i)2}}}.$$

By multiplying it by dr, $d\nu'$, ..., and integrating it with respect to r, ν', ν'', ..., from negative infinity to positive infinity, we will have a quantity proportional to the probability of ν. By multiplying therefore this quantity by $d\nu$ and by taking the integral within some given limits, by dividing next by this same integral taken from $\nu = -\infty$ to $\nu = +\infty$, we will have the probability that the value of ν is contained within these limits. We see thus that the consideration of the values that k' is able to have and on which depends the difference of probability of the positive and negative errors has no sensible influence on the results of the general method exposed here above.

§ 5. Let us apply now this method to an example. For this, I have profited from [516]
the immense work that Bouvard has just finished on the movements of Jupiter and of Saturn, from which he has constructed very precise Tables. He has made use of all the oppositions observed by Bradley and by the astronomers who have followed him: he has discussed them anew and with the greatest care, that which has given to him 126 equations of condition for the movement of Jupiter in longitude and 129 equations for the movement of Saturn. In these last equations, Bouvard has made the mass of Uranus enter as indeterminate. Here are the final equations that he has concluded by the most

advantageous method:

$$
\begin{aligned}
7212'', 600 =& 795938z - 12729398z' \\
&+ 6788, 2z'' - 1959, 0z''' + 696, 13z^{\text{iv}} + 2602z^{\text{v}}, \\
-738297'', 800 =& - 12729398z + 424865729z' \\
&- 153106, 5z'' - 39749, 1z''' - 5459z^{\text{iv}} + 5722z^{\text{v}}, \\
237'', 782 =& 6788, 2z - 153106, 5z' \\
&+ 71, 8720z'' - 3, 2252z''' + 1, 2484z^{\text{iv}} + 1, 3371z^{\text{v}}, \\
-40'', 335 =& - 1959, 0z - 39749, 1z' \\
&- 3, 2252z'' + 57, 1911z''' + 3, 6213\text{v}^{\text{iv}} + 1, 1128z^{\text{v}}, \\
-343'', 455 =& 696, 13z - 5459z' \\
&+ 1, 2484z'' + 3, 6213z''' + 21, 543z^{\text{iv}} + 46, 310z^{\text{v}}, \\
-1002'', 900 =& 2602z + 5722z' \\
&+ 1, 3371z'' + 1, 1128z''' + 46, 310z^{\text{iv}} + 129z^{\text{v}}.
\end{aligned}
$$

In these equations, the mass of Uranus is supposed $\frac{1+z}{19504}$; the mass of Jupiter is supposed $\frac{1+z'}{1067,09}$; z'' is the product of the equation of the center by the correction of the perihelion employed first by Bouvard; z''' is the correction of the equation of the center; z^{iv} is the secular correction of the mean movement; z^{v} is the correction of the epoch of the longitude at the beginning of 1750. The second of the decimal degree is taken for unity.

By means of the preceding equations contained in the system (A), I have concluded the following, contained in the system (B): [517]

$$
\begin{aligned}
27441'', 68 =& 743454z - 12844814z' \\
&+ 6761, 23z'' - 1981, 45z''' - 237, 97z^{\text{iv}}, \\
-693812'', 58 =& - 12844814z + 424611920z' \\
&- 153165, 81z'' - 39798, 46z''' - 7513, 15z^{\text{iv}}, \\
248'', 1772 =& 6761, 23z - 153165, 81z' \\
&+ 71, 8581z'' - 3, 2367z''' + 0, 7684z^{\text{iv}}, \\
-31'', 6836 =& - 1981, 45z - 39798, 46z' \\
&- 3, 2367z'' + 57, 1815z''' + 3, 2218z^{\text{iv}}, \\
16'', 5783 =& - 237, 97z - 7513, 15z' \\
&+ 0, 7684z'' + 3, 2218z''' + 4, 9181z^{\text{iv}}.
\end{aligned}
$$

From these equations, I have deduced the following four, contained in the system

(C),

$$\begin{aligned}
28243'',85 =& 731939,5z - 1328350z' + 6798,41z'' - 1825,56z''',\\
-668486'',70 =& -13208350z + 413134432z' - 1519920z'' - 34876,7z''',\\
245'',5870 =& 6798,41z - 151992,0z' + 71,7381z'' - 3,7401z''',\\
-42'',5434 =& -1825,56z - 34876,7z' - 3,7401z'' + 55,0710z''';
\end{aligned}$$

these last equations give the following, contained in the system (D),

$$\begin{aligned}
26833'',55 =& 671414,7z - 14364541z' + 6674,43z'',\\
-695430'',0 =& -14364541z + 391046861z' - 154360,6z'',\\
242'',6977 =& 6674,43z - 154360,6z' + 71,4841z''.
\end{aligned}$$

Finally I have concluded thence the following two equations, contained in the system (E):

$$4172'',95 = 48442z + 48020z', \qquad -171455'',2 = 48020z + 57725227z'.$$

I stop myself at this system, because it is easy to conclude from it the values of the weight P relative to the two elements z and z' that I desired particularly to know. The [518]
formulas of § 3 give, for z,

$$P = \frac{s}{2\mathrm{S}\epsilon'^{(i)2}}\left[48442 - \frac{(48020)^2}{57725227}\right]$$

and, for z',

$$P = \frac{s}{2\mathrm{S}\epsilon'^{(i)2}}\left[57725227 - \frac{(48020)^2}{48442}\right].$$

The number s of the observations is here 129 and Bouvard has found

$$\mathrm{S}\epsilon'^{(i)2} = 31096;$$

we have therefore, for z,

$$\log P = 2,0013595$$

and, for z',

$$\log P = 5,0778624.$$

The preceding equations give

$$\begin{aligned}
z' =& -0,00305,\\
z =& 0,08916.
\end{aligned}$$

The mass of Jupiter is $\frac{1}{1067,09}(1 + z')$. By substituting for z' its preceding value, this mass becomes $\frac{1}{1070,35}$. The mass of the Sun is taken for unity. The probability that the error of z' is comprehended within the limits $\pm U$ is, by § 1,

$$\frac{\sqrt{P}}{\sqrt{\pi}}\int du\, c^{-Pu^2},$$

the integral being taken from $u = -U$ to $u = U$. We find thus the probability that the mass of Jupiter is comprehended within the limits

$$\frac{1}{1070,35} \pm \frac{1}{100}\,\frac{1}{1067,09},$$

equal to $\frac{1000000}{1000001}$; so that there are odds one million very nearly against one that the value $\frac{1}{1070,35}$ is not in error of a hundredth of its value; or, that which reverts to quite nearly the same, that after a century of new observations, added to the previous, and [519] discussed in the same manner, the new result will not differ from the previous by a hundredth of its value.

Newton had found, by the observations of Pound, on the elongations of the satellites of Jupiter, the mass of this planet equal to the 1067th part of that of the Sun, that which differs very little from the result of Bouvard.

The mass of Uranus is $\frac{1+z}{19504}$. By substituting for z its previous value, this mass becomes $\frac{1}{17907}$. The probability that this value is comprehended within the limits

$$\frac{1}{17907} \pm \frac{1}{4}\,\frac{1}{19504},$$

is equal to $\frac{2508}{2509}$, and the probability that this mass is comprehended within the limits

$$\frac{1}{17907} \pm \frac{1}{5}\,\frac{1}{19504}$$

is equal to $\frac{215,6}{216,6}$.

The perturbations that Uranus produces in the movement of Saturn being of little importance, we must not yet expect from the observations of this movement a great precision in the value of its mass. But, after a century of new observations, added to the previous and discussed in the same manner, the value of P will increase in a manner to give this mass with a great probability that its value will be contained within some narrow limits; that which will be much preferable to the use than the elongations of the satellites of Uranus, because of the difficulty to observe these elongations.

Bouvard, by applying the previous method to the 126 equations of condition which the observations of Jupiter have given to him and by supposing the mass of Saturn equal to $\frac{1+z}{3534,08}$, has found

$$z = 0,00620$$

and

$$\log P = 4,8856829.$$

These values give the mass of Saturn equal to $\frac{1}{3512,3}$, and the probability that this mass [520] is comprehended within the limits

$$\frac{1}{3512,3} \pm \frac{1}{100}\,\frac{1}{3534,08}$$

is equal to $\frac{11327}{11328}$.

Newton had found, by the observations of Pound on the greatest elongation of the fourth satellite of Saturn, the mass of this planet equal to $\frac{1}{3012}$, that which surpasses by a sixth the preceding result. There are odds of millions of billions against one that the one of Newton is in error, and we will not at all be surprised if we consider the difficulty to observe the greatest elongations of the satellites of Saturn. The facility to observe those of the satellites of Jupiter has rendered, as we have seen, the value much more exact than Newton has concluded from the observations of Pound.

On the probability of judgments.
1816

I have compared, in § 50 of Book II, the judgment of a tribunal which pronounces between two contradictory opinions to the result of the testimonies of many witnesses of the extraction of a ticket from an urn which contains only two tickets. There is however between these two cases this difference, namely, that the probability of the testimony is independent of the nature of the thing attested, because we suppose that the witness has not been able to be deceived on this thing; instead an object in litigation is able to be surrounded by such obscurities, that the judges, in their supposing all the good faith desirable, are able to be however of contrary opinions. The nature of the affair which is subject to them must therefore influence on their judgment. I will make this consideration enter into the following investigations, by applying it to the judgments in criminal matter.

In order to condemn an accused, without doubt the strongest proof of his offense is necessary to the judges. But a moral proof is never but a probability, and experience has only too well made known the errors of which the criminal judgments, even those [521] which appear to be most just, are yet susceptible. The possibility to repair these errors is the most solid argument of the philosophers who have wished to proscribe the pain of death. We should therefore abstain ourselves from judging, if it was necessary we await mathematical evidence. But, when the proofs have a force such that the product of the error to fear by its feeble probability is inferior to the danger which would result from the impunity of the crime, judgment is commanded by the interest of society. This judgment is reduced, if I do not deceive myself, to the solution of the following question: Has the proof of the offense of the accused the high degree of probability necessary in order that the citizens have less to fear the errors of the tribunals, if he is innocent and condemned, than his new attempts and those of the unfortunate persons who the example of his impunity would embolden, if he was culpable and absolved? The solution of this question depends on many elements very difficult to know. Such is the imminence of danger which would menace society if the accused criminal remained unpunished. Sometimes, this danger is so great that the magistrate sees himself obliged to renounce the prudent forms established for the certainty of innocence. But that which renders nearly always the question of which there is concern insoluble is the impossibility to estimate exactly the probability of the offense, and to fix that which is necessary for the condemnation of the accused. Each judge, in this regard, is forced to bring himself back to his proper feeling. He forms his opinion by comparing the diverse witnesses and the circumstances of which the offense is accompanied to the results of his reflections and of his experience; and, under this relation, a long habit of interrogating and judging the accused gives much advantage in order to know the truth in the midst of often contradictory indices.

The preceding question depends further on the magnitude of the punishment applied to the offense; because we require naturally, in order to pronounce death, proofs much stronger than to inflict a detention of some months. This is a reason to proportion the punishment to the offense, a grave punishment applied to a light offense must inevitably render absolved many a guilty person. The product of the probability of the offense by its gravity being the measure of the danger that absolution of the accused is [522] able to make society experience, we would be able to think that the punishment must depend on this probability. This is that which we do indirectly in the tribunals where we retain during some times the accused against whom are raised some very strong proofs, but insufficient to condemn. In the view to acquire new understanding, we deliver him not at all immediately into the midst of his fellow citizens, who would review it not without lively alarms. But the arbitrariness of this measure and the abuse that we are able to make of it has caused to reject it in the country where we attach a very great price to individual liberty.

Now, what is the probability that the decision of a tribunal which is able to condemn only by a given majority will be just, that is to say, conformed to the true solution of the question posed above? This important problem well resolved will give the means to compare the diverse tribunals among themselves. The majority of a single vote in a numerous tribunal indicates that the affair of which there is concern is nearly doubtful; the condemnation of the accused would be therefore then contrary to the principles of humanity, protectors of innocence. The unanimity of the judges would give a very great probability of a just decision; but, by being obliged, too many guilty persons would be absolved. It is necessary therefore either to limit the number of judges, if we wish that they be unanimous, or to increase the majority necessary to condemn, when the tribunal becomes more numerous. I will test by applying the calculus to this object, persuaded that the applications of this kind, when they are well conducted and based on some data that good sense suggests to us, are always preferable to the most specious reasonings.

The probability that the opinion of each judge is just enters as principal element in this calculation. This probability is evidently relative to each affair. If, in a tribunal of one thousand and one judges, five hundred one are of one opinion, and five hundred are of a contrary opinion, it is clear that the probability of the opinion of each judge surpasses quite little $\frac{1}{2}$; because, by supposing it sensibly greater, a single vote of difference would be an unlikely event. But, if the judges are unanimous, this indicates [523] in the proofs that degree of force which carries away the conviction. The probability of the opinion of each judge is therefore then very near to unity or of certitude; not unless some passions or some common prejudgments mislead all the judges. Beyond these cases, the ratio of the votes for or against the accused must alone determine this probability. I suppose thus that it is able to vary from $\frac{1}{2}$ to unity, but that it is not able to be below $\frac{1}{2}$. If this were not, the decision of the tribunal would be insignificant as the lot: it has value only as much as the opinion of the judge has more tendency to the truth than to the error. It is next by the ratio of the numbers of votes favorable or contrary to the accused that I determine the probability of this opinion.

These data suffice in order to have the general expression of the probability that the decision of the tribunal judging in a given majority is just. In our special tribunals composed of eight judges, five votes are necessary for the condemnation of an accused: the probability of the error to fear respecting the justness of the decision surpasses then $\frac{1}{4}$. If the tribunal were reduced to six members which would be able to condemn only with the plurality of four votes, the probability of the error to fear would be then below $\frac{1}{4}$; there would be therefore for the accused an advantage to this reduction of the tribunal. In both cases, the majority required is the same and equal to two. Thus, this majority remaining constant, the probability of the error increases with the number of judges. This is general, whatever be the majority required, provided that it remains the same. By taking therefore for rule the arithmetic relation, the accused is found in a position less and less advantageous in measure as the tribunal becomes more numerous. This relation is followed in the Chamber of the peers of England. One requires for the condemnation a majority of twelve votes, whatever be the number of judges. If we have belief that, the votes opposed are destroying reciprocally, the twelve remaining votes represent the unanimity of a jury of twelve members, required in the same country for the condemnation of an accused, we have been in a great error. Good sense shows that there is a difference between the tribunal of two hundred twelve judges, of whom one hundred twelve condemn the accused, while one hundred absolve him, and that of a tri- [524] bunal of twelve judges unanimous for condemnation. In the first case, the one hundred votes favorable to the accused permit thinking that the proofs are far from attaining the degree of force which draw the conviction. In the second case, the unanimity of the judges carry belief that they have attained this degree. But simple good sense does not suffice to estimate the extreme difference of the probability of error in these two cases. It is necessary then to recur to the calculus, and we find very nearly $\frac{1}{5}$ for the probability of the error in the first case, and only $\frac{1}{8192}$ for this probability in the second case, a probability which is not $\frac{1}{1000}$ of the first. This is a confirmation of the principle that the arithmetic relation is unfavorable to the accused when the number of judges increases. To the contrary, if we take for rule the geometric relation, the probability of the error of the decision diminishes when the number of the judges is increased. For example, in the tribunals which would be able to condemn only in the plurality of the two thirds of votes, the probability of the error to fear is nearly $\frac{1}{4}$ if the number of judges is six: it is below $\frac{1}{7}$ if this number is raised to twelve. Thus we must be regulated neither on the arithmetic relation, nor on the geometric relation, if we wish that the probability of error is never above nor below a determined fraction.

But to what fraction must we be fixed? It is here that the arbitrary commences, and the tribunals offer in this regard great varieties. In the special tribunals, where five votes out of eight suffice for the condemnation of the accused, the probability of the error to fear respecting the goodness of the judgment is $\frac{65}{256}$ or below $\frac{1}{4}$. The magnitude of this fraction is frightening; but that which must reassure a little is the consideration that, most often, the judge who absolves an accused regards him not as innocent. He pronounces only that it is not attained by some sufficient proofs in order that he be condemned. We are especially reassured by the pity that nature has put into the heart of man, and which disposes the mind to see with difficulty a guilty person in the accused submitted to his judgment. This sentiment, more quick in those who have not at all the habit of criminal judgments, outweighs the inconvenience attached to the [525] inexperience of juries. In a jury of twelve members, if the plurality required for the condemnation is of eight votes out of twelve, the probability of error to fear is $\frac{1093}{8192}$ or a little less than $\frac{1}{8}$: it is nearly $\frac{1}{22}$ if this plurality is of nine votes. In the cases of unanimity, the probability of error to fear is $\frac{1}{8192}$, that is more than one thousand times less than in our juries.

The solution of the problem that we just considered does not suffice to fix the convenient majority, in a tribunal of any number of judges whatsoever. It is necessary, for this, to know the probability of the offense below which an accused is not able to be condemned, without that the citizens having to dread more the errors of the tribunals, than the attacks which would be born from the impunity of a guilty person absolved. It is necessary next to determine the probability of the offense resulting from the decision of the tribunal and to fix the majority in a manner that these probabilities are equals. But it is impossible to obtain them. The first is, as we have said, relative to the position in which society is found, a variable position, very difficult to define well and always too complicated in order to be submitted to the calculus. The second depends on a thing entirely unknown, the law of probability of the opinion of each judge in the estimation that he makes of the probability of the offense. Seeing our ignorance of these two elements of the calculus, what is more reasonable than to depart from the solution of the single problem that we may resolve in this manner, the one of the probability of the error of the decision of a tribunal? This probability appears to me too high in our tribunals, and I think that in this regard it is acceptable to approach to the English jury where it is only $\frac{1}{8192}$. In fixing it at the fraction $\frac{1}{1024}$ and in determining the majority necessary to attain it, we place the accused in the position where he would be vis-à-vis of a jury of nine members, of which we would require unanimity; that which appears to me to guarantee sufficiently the innocent ones from the errors of the tribunals, and society from the pains that impunity of the guilty persons would produce. It must be extremely rare then that an accused is condemned with a probability less than that which is necessary to his condemnation; because the majority who condemn him [526] declare that the probability of his offense is at least equal to this necessary probability: the minority who absolve him declare that the first of these probabilities appears to it inferior to the second; but it is natural to believe that this inferiority is not very considerable. It must rarely happen that the mean probability which results from the totality of the judgments of the members of the tribunal is inferior to the probability required for the condemnation of the accused, if we reduce, by a convenient majority, the probability of the error to fear respecting the justice of the decision, to the fraction $\frac{1}{1024}$. The analysis furnishes, in order to have this majority, some formulas which I will expose here and that it is easy to reduce into a Table dependent on the number of the judges. But a parallel Table will appear too arbitrary to the common men who will prefer always one or the other of the arithmetic and geometric relations which they are able to imagine easily.

§ 1. A judge must not, in order to condemn an accused, expect the mathematical evidence that it is impossible to attain in moral things. But, when the probability of the offense is such that the citizens had more to dread the attempts which would be able to be born of his impunity than the errors of the tribunals, the interest of society requires the condemnation of the accused. I name a this degree of probability, and I suppose that the judge who condemns an accused pronounces thence that the probability of his offense is at least a. I name x the probability of this opinion of the judge, a probability that I will suppose equal or superior to $\frac{1}{2}$, and varying by some infinitely small degrees, equal to x and equally probable *a priori*. I suppose further that the tribunal is composed of $p+q$ judges, of whom p condemn the accused and q absolve him. The probability that the opinion of the tribunal is just will be proportional to $x^p(1-x)^q$, and the probability that it is not will be proportional to $(1-x)^p x^q$; the probability of the goodness of the judgment will be therefore, by § 1 of Book II,

$$(a)\qquad \frac{x^p(1-x)^q}{x^p(1-x)^q+(1-x)^p x^q}.$$

It is necessary to multiply this quantity by the probability of the value of x, taken from the observed event. This event is that the tribunal is itself divided into two parts of which the one, composed of p judges, condemn the accused, and of which the other, formed of q judges, absolve him. The probability of x is therefore the function $x^p(1-x)^q+(1-x)^p x^q$ divided by the sum of all the similar functions relative to all the values of x, from $x=\frac{1}{2}$ to $x=1$; it is consequently [527]

$$\frac{\left[x^p(1-x)^q+(1-x)^p x^q\right]dx}{\int x^p dx(1-x)^q}.$$

the integral of the denominator being taken from $x=0$ to $x=1$. By multiplying this function by the function (a), we will have

$$\frac{x^p(1-x)^q dx}{\int x^p dx(1-x)^q}$$

for the probability of the goodness of the judgment relative to x. The same probability relative to all the values of x is therefore

$$(b)\qquad \frac{\int x^p dx(1-x)^q}{\int x^p dx(1-x)^q},$$

the integral of the numerator being taken from $x=\frac{1}{2}$ to $x=1$, and that of the denominator being taken from $x=0$ to $x=1$. It follows thence that the probability of the error to fear respecting the goodness of the judgment is further expressed by formula (b), provided that we take the integral of the numerator from $x=0$ to $x=\frac{1}{2}$. We find thus this last probability equal to

(c)

$$\frac{1}{2^{p+q+1}}\left\{\begin{aligned}&1+\frac{p+q+1}{1}+\frac{(p+q+1)(p+q)}{1.2}+\frac{(p+q+1)((p+q)(p+q-1)}{1.2.3}\\&+\cdots+\frac{(p+q+1)((p+q)(p+q-1)\cdots(p+2)}{1.2.3\ldots q}\end{aligned}\right\}$$

If we require unanimity, q is null, and this expression becomes $\frac{1}{2^{p+1}}$.

§ 2. Let us determine presently the probability of the error to fear respecting the justice of the decision of the tribunal, when p and q are large numbers; that which renders the formula (c) very difficult to evaluate in numbers. It is necessary to distinguish here two cases, one in which $p-q$ is considerable, the other in which $p-q$ is rather [528] small. In the first case, we will make use of formula (o) of § 28 of Book II which gives, for the probability of error,

$$(e) \qquad \frac{(p+q)^{p+q+\frac{3}{2}}}{2^{p+q+\frac{3}{2}}p^{p+\frac{1}{2}}q^{q+\frac{1}{2}}(p-q)\sqrt{\pi}}\left\{1-\frac{p+q}{(p-q)^2}-\frac{[(p+q)^2-13pq]}{12pq(p+q)}\right\},$$

π being the circumference of which the diameter is unity.

In the second case, where $p-q$ is a small number relative to p, we will find easily, by the analysis of § 19 of Book II, the probability of error to fear equal to

$$(f) \qquad \frac{\int dt\, c^{-t^2}}{\sqrt{\pi}},$$

the integral being taken from

$$t^2 = \frac{(p-q)^2(p+q)}{8pq}$$

to infinity.

In order to give an example of each of these formulas, we suppose a tribunal formed of 144 judges, and that $\frac{5}{6}$ is necessary for condemnation of the accused. Then we have

$$p = 90, \quad q = 54,$$

and formula (e) gives $\frac{1}{773}$ for the probability of the error to fear respecting the goodness of the decision of the tribunal. In the case of unanimity of a jury of eight members, the probability of the error to fear is $\frac{1}{512}$; the accused is therefore then in a more favorable position than vis-à-vis a similar jury.

Let us suppose the tribunal formed of 212 judges and that a majority of twelve votes suffices for condemnation. In this case

$$p+q = 212, \quad p-q = 12,$$

and formula (f) gives $\frac{1}{4{,}889}$ for the probability of the error to fear.

On a disposition of the Code of criminal instruction.
15 November 1816 [529]

Article 351 of the Code of criminal instruction is thus conceived:

> "If nevertheless the accused is declared guilty only by a simple majority, the judges will deliberate among them on the same point: and if the opinion of the minority of the jurors is adopted by the majority of the judges, of such sort that by reuniting the number of votes, this number exceeds the one of the majority of the jurors and of the minority of the judges, the opinion favorable to the accused will prevail."

After this article, seven jurors declaring the accused guilty and five declaring him not guilty, the accused is condemned when three alone of the five judges of the Assize Court are reunited to the minority of the jurors. This appears to shock at the same time the rules of common sense and the principals of humanity, protectors of innocence. The Assize Court intervene then with justice, because the offense of the accused is not sufficiently established by a simple majority of the jury, that which the Calculus of Probabilities render indubitable. But, when the opinion of the Assize Court annuls the one of the majority of the jurors, far from confirming it, when the difference of the two votes, which gave to this majority only an insufficient preponderance, is reduced to a single vote, by the addition of the judges of whom the state and the wise must inspire confidence[1], is it not unjust to condemn the accused?

I propose therefore to reform thus the article cited: [530]

> "If nevertheless the accused is declared culpable only by a simple majority, the judges will deliberate among them on the same point; and if the opinion of the minority of jurors is adopted by the majority of the judges, this opinion will prevail."

If this reform appeared just, the indispensable duty to abrogate promptly all that which is able to compromise the innocence does not permit to await, in order to convert by law, the general revision of the criminal Code, a revision which demands much reflection and time. It is in order to fulfill this duty as much as it is possible to me that I publish this writing.

15 November 1816

[1]The difference of one vote gives to the majority a preponderance so much less as the number of judges is more considerable: simple good sense shows it without the help of the calculus. In the present question, the preponderance of the majority of the jurors diminishes therefore not only by the reduction of the two votes to one, but further by the increase of the number of voters, which is raised from twelve to seventeen. Generally, a constant difference between the majority and the minority below of which the accused is not able to be condemned is so much less favorable to him as the number of judges is greater: on the contrary, the constant ratio of the votes of the majority to those of the minority become to him more favorable, in measure as the number of judges increases. The ratio $\frac{5}{8}$, adopted by the Chamber of Peers of France, is very favorable to the accused before a tribunal so numerous. (We are able to see, on this object, the Supplement to my *Théorie analytique des Probabilités*, and the third edition of my *Essai philosophique sur les Probabilités*.)