GAUSS’S WORK
(1803-1826)
on the
Theory of Least Squares

An English Translation by Hale F. Trotter
PREFACE

This report\(^1\) consists of an English translation of the work of Gauss on the method of least squares. It has been prepared from a French translation by J. Bertrand which had been authorized by Gauss and was published in 1855, shortly after his death.

The connected exposition of the theory is contained in a long paper (in two parts and a later supplement) presented to the Royal Society of Göttingen in 1821, 1823 and 1826. Bertrand also included the relevant discussion in *Theoria Motus Corporum Coelestium*, and several discussions of applications in the reduction of astronomical data and in surveying. These are listed in the Table of Contents, with page references to Gauss’ collected works.

The present report makes no pretense of being a polished translation, but I believe it to be substantially accurate. In a number of places I have used modern terminology which does not correspond to a literal translation of the original. For example, “error medius metuendus” (“erreur moyen a craindre” in Bertrand’s translation) has been translated as “standard deviation” or “standard error to be expected” according to the text, since the literal translation as “average error to be feared” would actually be misleading.

This report owes a great deal to the typists, Mrs. Sandra Beirig and Miss Joan Fisher, and to H.J. Arnold, who carried out the tedious job of proof-reading the final copy.

5 August 1957

Hale F. Trotter

\(^{1}\)Trotter’s text is not copyrighted. This version has been generated from a hectographed pamphlet, bound pages which were reproduced from a spirit duplicator. There are several places where printing was very poor and not legible. The type setting of the mathematics was ambiguous (to be kind). In addition, there appeared to be several omissions of characters.

According to Farebrother, Trotter produced his translation for the U.S. Army.

In this reproduction, the original text has been replicated as precisely as possible, correcting only the obvious errors by making reference to Bertrand. Careful reading or comparison with Gauss’ original work will undoubtedly uncover more.
TABLE OF CONTENTS

(References to *Werke* are to Gauss’s collected works\(^2\), edited by the Royal Scientific Society of Göttingen)

<table>
<thead>
<tr>
<th>Work</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theory of the Combination of Observations which leads to the Smallest Errors (1821–1826), <em>(Theoria Combinationis Observationum Erroribus Minimis Obnoxiae, Werke, vol. 4, p. 1–93)</em></td>
<td>4</td>
</tr>
<tr>
<td>First Part (1821), (Pars Prior)</td>
<td></td>
</tr>
<tr>
<td>Second Part (1823), (Pars Posterior)</td>
<td>27</td>
</tr>
<tr>
<td>Supplement (1826), (Supplementum)</td>
<td>53</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Extract (paragraphs 13–14) from Discourse on the Elements of the Orbit of Pallas (1803–1809), <em>(Disquisitiones de Elementis Ellipticis Pallidis, Werke, vol. 6, p. 20–24)</em></td>
<td>100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Application of Calculus of Probabilities to a Problem in Practical Geometry (1823), <em>(Anwendung der Wahrscheinlichkeitsrechnung auf eine Aufgabe der praktischen Geometrie, Werke, vol. 9, p. 231–237)</em></td>
<td>114</td>
</tr>
</tbody>
</table>


\(^2\)Carl Friedrich Gauss Werke: Herausgegeben von der Königlichen Gesellschaft der Wissenschaften, Göttingen(vol. 4, 1880; vol. 6, 1874; vol. 7, 1906).
However much care is taken with observations of the magnitude of physical quantities, they are necessarily subject to more or less considerable errors. These errors, in the majority of cases, are not simple, but arise simultaneously from several distinct sources which it is convenient to distinguish into two classes.

Certain causes of errors depend, for each observation, on circumstances which are variable and independent of the result which one obtains: the errors arising from such sources are called irregular or random, and like the circumstances which produce them, their value is not susceptible of calculation. Such are the errors which arise from the imperfection of our senses and all those which are due to irregular exterior causes, such as, for example, the vibrations of the air which make vision less clear; some of the errors due to the inevitable imperfection of the best instruments belong to the same category. We may mention, for example, the roughness of the inside of the level, the lack of absolute rigidity, etc.

On the other hand, there exist causes which in all observations of the same nature produce an identical error, or depend on circumstances essentially connected with the result of observation. We shall call the errors of this category constant or regular.

It is evident that this distinction is relative up to a certain point and depends on how broad a sense one wishes to attach to the idea of observations of the same nature. For instance, if one repeats indefinitely the measurement of a single angle, the errors arising from an imperfect division of the circular scale will belong to the class of constant errors. If, on the other hand, one measures successively several different angles, the errors due to the imperfection of the division will be regarded as random as long as one has not formed the table of errors pertaining to each division.

We exclude consideration of regular errors from our discussion. It is up to the observer to seek out carefully the causes which may produce a constant error, in order to avoid them if possible, or at least to estimate their effect, so as to make a correction for them on each observation, which will then give the same result as if the constant cause had not existed. The situation is quite different with irregular errors: the latter, by their nature, cannot be calculated, and it is certainly necessary to tolerate them in observations. One can, however, by a suitable combination of the results, reduce their influence as much as possible. It is to this important question that the following discussion is devoted.
3.

The errors which, in observations of the same type, arise from a single determined cause are bounded between certain limits which one could no doubt assign if the nature of this cause itself were perfectly known. In the majority of cases, all the errors included between these extreme limits ought to be regarded as possible. A deeper knowledge of each cause would tell us whether all these errors have equal or unequal probability, and, in the second case, what the relative probability of each error is. The same remark applies to the total error which arises from the combination of several simple errors. This error also will be bounded between two limits, one of which will be the sum of the upper limits corresponding to the simple errors, and the other the sum of the lower limits. All the errors included between these limits will be possible, and each one can result in an infinity of ways from suitable values attributed to the partial errors. One can see nevertheless, setting aside the purely analytic difficulties, that it is possible to estimate the probability of each result, if one supposed known the probabilities associated with each of the simple causes.

Certain causes, however, produce errors which cannot vary according to a continuous law, but which, on the other hand, take on a finite number of values: we may mention as an example the errors which arise from the imperfect division of scales on instruments (at least if one wishes to class these among random errors), for the number of divisions in any given instrument is essentially finite. It is nevertheless clear that unless all the causes which combine to produce the total error are assumed to be of this type, their sum will be subject to the law of continuity, or at least, will vary continuously over several distinct intervals, if it happens that when all the possible values of the discontinuous errors are arranged in order of magnitude, the difference between two consecutive terms of the series is greater than the difference between the extreme limits of the errors subject to the law of continuity. In practice, such a case will almost never arise; it would involve excessively large inaccuracies in the construction of the instrument.

4.

Let us denote by $\phi(x)$ the relative probability of an error $x$. One should understand by that, because of the continuity of the errors, that $\phi(x)dx$ is the probability that the error is contained between the limits $x$ and $x + dx$. It is not in general possible to assign a priori the form of the function $\phi$, and one may even state that this function will never be known in practice. One can nevertheless establish several general characteristics which it must necessarily have: $\phi(x)$ is clearly a discontinuous function; it vanishes for all the values of $x$ outside of the extreme errors. For each value included between these extreme limits, the function is positive (except in the case mentioned at the end of the preceding paragraph); in the majority of cases, errors which are equal and of opposite signs will be equally likely, and one will have

$$\phi(x) = \phi(-x).$$

Finally, since small errors are more easily committed than large ones, $\phi(x)$ will in general have a maximum when $x = 0$ and will decrease continuously as $x$ increases.
The integral
\[ \int_{a}^{b} \phi(x) \, dx \]
expresses the probability that the error, still unknown, falls between the limits \( a \) and \( b \). From this one concludes that the value of this integral taken between the extreme limits of the possible errors will always be equal to one. And since \( \phi(x) \) is 0 for the values of \( x \) not included between these limits, one can say, in all cases, that
\[ \int_{-\infty}^{\infty} \phi(x) \, dx = 1 \]

Let us consider the integral
\[ \int_{-\infty}^{\infty} x \phi(x) \, dx \]
and represent its value by \( k \). If the causes of error are such that there is no reason for two equal errors of opposite signs to have unequal probabilities, one will have
\[ \phi(x) = \phi(-x), \]
and consequently
\[ k = 0. \]

We conclude from this that if \( k \) does not vanish and has, for instance, a positive value, there must necessarily exist a cause of error which produces only positive errors, or, at least, produces them more readily than it does negative errors. This quantity \( k \), which is the average of all the possible errors, or the average value of \( x \), can be denoted conveniently by the name of constant part of the error. Finally, one easily proves that the constant part of the total error is the sum of the constant parts of the simple errors which make it up.

If the quantity \( k \) is supposed known, and one subtracts it from the result of each observation, denoting the error of the observation corrected in this way by \( x' \), and the corresponding probability by \( \phi'(x') \), one will have
\[ x' = x - k, \quad \phi'(x') = \phi(x) \]
and consequently,
\[ \int_{-\infty}^{\infty} x' \phi'(x') \, dx' = \int_{-\infty}^{\infty} x \phi(x) \, dx - k \int_{-\infty}^{\infty} \phi(x) \, dx = k - k = 0; \]
so that the errors of the corrected observations have no constant part. Which, indeed, seems evident a priori.
From the value of the integral
\[ \int_{-\infty}^{\infty} x \phi(x) \, dx, \]
that is to say the average value of \( x \), we learn the existence or non-existence of a constant error, as well as the value of this error; similarly, the integral
\[ \int_{-\infty}^{\infty} x^2 \phi(x) \, dx, \]
that is to say the average value of \( x^2 \), seems very suitable for defining and measuring, in a general way, the uncertainty of a system of observations; so that between two systems of observations of unequal precision, one should regard as preferable that which gives a smaller value to the integral
\[ \int_{-\infty}^{\infty} x^2 \phi(x) \, dx. \]

If one objects that this convention is arbitrary and does not appear necessary, we readily agree. The question which concerns us here has something vague about it from its very nature, and cannot be made really precise except by some principle which is arbitrary to a certain degree. The determination of a magnitude by observation can be compared, with some appropriateness, to a game in which there is a loss to fear and no gain to hope for; each error committed being likened to a loss which one suffers, the relative undesirability of such a game should be expressed by the probable loss, that is to say by the sum of the products of the various possible losses by their respective probabilities. But what loss should one associate with a given error? This is something which is not clear in itself; the evaluation depends in part on our choice. It is clear to begin with that the loss should not be considered as proportional to the error committed; for under this hypothesis, since a positive error represents loss, a negative error would be considered as a gain; the magnitude of the loss ought, on the contrary, to be evaluated by a function of the error whose value is always positive. Among the infinite number of functions satisfying this condition, it seems natural to choose the simplest, which is, without doubt, the square of the error, and in this way we are led to the principle proposed above.

Laplace has considered the question in an analogous manner, but he adopted the absolute value of the error itself as measure of the loss. This hypothesis, unless we deceive ourselves, is no less arbitrary than ours; should one in fact consider an error as more or less regrettable than an error of half the size repeated twice, and should one, in consequence, give it an importance double or more than double that of the second? The question is not clear and it is one on which mathematical arguments have no bearing; each must resolve it to his own liking. One cannot deny, however, that Laplace’s hypothesis departs from the law of continuity and is consequently less suited to analytic study; ours, on the contrary, recommends itself by the generality and simplicity of its consequences.
Continuing with the preceding notation, let us put
\[
\int_{-\infty}^{\infty} \phi(x) x^2 dx = m^2;
\]
we shall call \( m \) the standard error to be expected, or simply the standard error, of the observations under consideration. We do not restrict this terminology to the immediate result of the observations, but rather extend it to every magnitude which can be deduced from them in any way whatever. One should be careful not to confuse this standard error with the arithmetic average of the errors, which was discussed in article 5.

If we compare several systems of observations or several quantities resulting from observations to which one does not grant the same precision, we shall regard their relative weight as inversely proportional to \( m^2 \) and their precision as inversely proportional to \( m \). In order to be able to represent the weights by numbers, one must take as unit the weight of some arbitrarily chosen system of observations.

If the errors of the observations have a constant part, by subtracting it from each result obtained, the standard error is diminished, and the weight and precision are increased. Keeping the notation of article 5, and denoting by \( m' \) the standard error of the corrected observations, one has in fact
\[
m'^2 = \int_{-\infty}^{\infty} x'^2 \phi'(x') \, dx' = \int_{-\infty}^{\infty} (x - k)^2 \phi(x) \, dx
\]
\[
= \int_{-\infty}^{\infty} x^2 \phi(x) \, dx - 2k \int_{-\infty}^{\infty} x \phi(x) \, dx + k^2 \int_{-\infty}^{\infty} \phi(x) \, dx
\]
\[
= m^2 - 2k^2 + k^2 = m^2 - k^2.
\]
If, instead of subtracting from each observation the number \( k \), one subtracts another number \( l \), the square of the standard error becomes
\[
m^2 - 2kl + l^2 = m'^2 + (l - k)^2.
\]

Letting \( \lambda \) be a given coefficient and \( \mu \) the value of the integral
\[
\int_{-\lambda m}^{\lambda m} \phi(x) \, dx,
\]
\( \mu \) will be the probability that the error of an observation is less than \( \lambda m \) in absolute value; \( 1 - \mu \) will be, on the other hand, the probability that this error is greater than \( \lambda m \). If \( \rho \) is the value of \( \lambda m \) which makes \( \mu = \frac{1}{2} \), the probabilities that the error be greater than or less than \( \rho \) are equal; \( \rho \) may thus be called the probable error. The
relation between $\rho$ and $\mu$ depends on the nature of the function $\phi$, which is unknown in the majority of cases. It is interesting to study this relation in certain particular cases.

I. If the extreme limits of possible errors are $+a$ and $-a$, and if between these limits all the errors are equally probable, the function $\phi(x)$ will be constant between these same limits, and consequently equal to $\frac{1}{2a}$. Consequently one will have

$$m = a\sqrt{\frac{1}{3}}, \quad \mu = \lambda\sqrt{\frac{1}{3}},$$

as long as $\lambda$ is less than or equal to $\sqrt{3}$; finally,

$$\rho = m\sqrt{\frac{3}{4}} = 0.8660254m,$$

and the probability that the error does not exceed the standard error is

$$\sqrt{\frac{1}{3}} = 0.5773503.$$

II. If $-a$ and $+a$ are again the limits of the possible errors, and if one supposes furthermore that the probability of these errors decreases as one goes out from the error 0 as do the terms of an arithmetic progression, one will have

$$\phi(x) = \frac{a - x}{a^2}$$

for the values of $x$ between 0 and $+a$, and

$$\phi(x) = \frac{a + x}{a^2}$$

for the values between 0 and $-a$: from this one deduces

$$m = a\sqrt{\frac{1}{6}}, \quad \mu = \lambda\sqrt{\frac{2}{3} - \frac{1}{6}\lambda^2}$$

as long as $\lambda$ is included between 0 and $\sqrt{6}$;

$$\lambda = \sqrt{6} - \sqrt{6 - 6\mu},$$

as long as $\mu$ is included between 0 and 1; and, finally,

$$\rho = m(\sqrt{6} - \sqrt{3}) = 0.7174389m.$$

In this case, the probability that the error is less than or equal to the standard error will be

$$\sqrt{\frac{2}{3} - \frac{1}{6}} = 0.6498299.$$
III. If we suppose the function $\phi(x)$ proportional to $e^{-x^2/h^2}$ [which in reality is only approximately true\(^3\)], it will be equal to

$$\phi(x) = \frac{e^{-x^2/h^2}}{h\sqrt{\pi}}$$

and one concludes from this

$$m = h\sqrt{\frac{1}{2}}.$$  

(See *Disquisitiones generales circa seriem infinitam*, article 28)

If one denotes by $\Theta(z)$ the value of the integral

$$\frac{2}{\sqrt{\pi}} \int_0^z e^{-z^2} \, dz,$$

one will have

$$\mu = \Theta \left( \lambda \sqrt{\frac{1}{2}} \right).$$

The following table gives several values of this quantity:

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6744897</td>
<td>0.5</td>
</tr>
<tr>
<td>0.8416213</td>
<td>0.6</td>
</tr>
<tr>
<td>1.0000000</td>
<td>0.6826895</td>
</tr>
<tr>
<td>1.0364334</td>
<td>0.7</td>
</tr>
<tr>
<td>1.2815517</td>
<td>0.8</td>
</tr>
<tr>
<td>1.6448537</td>
<td>0.9</td>
</tr>
<tr>
<td>2.5758293</td>
<td>0.99</td>
</tr>
<tr>
<td>3.2918301</td>
<td>0.999</td>
</tr>
<tr>
<td>3.8905940</td>
<td>0.9999</td>
</tr>
<tr>
<td>$\infty$</td>
<td>1.</td>
</tr>
</tbody>
</table>

Although the relation which connects $\lambda$ with $\mu$ depends on the nature of the function $\phi$, one can still establish several general results, which apply in all cases in which this function is non-increasing with the absolute value of $x$; one has then the following theorems:

- When $\mu$ is less than $\frac{2}{3}$, $\lambda$ does not exceed $\mu\sqrt{3}$;
- When $\mu$ exceeds $\frac{2}{3}$, $\lambda$ does not exceed $\frac{2}{3\sqrt{2/\pi}}$;
- When $\mu$ equals $\frac{2}{3}$, the two bounds coincide and $\lambda$ cannot be greater than $\sqrt{\frac{2}{3}}$.

---

\(^3\)To understand this remark, it is necessary to refer to a chapter of *Theoria Motus Corporum Coelestium* in which Gauss shows that this law of probability is the most reasonable which one can adopt. We reproduce this chapter in which the illustrious author made known for the first time the method of least squares at the end of the volume, page 127. —J.B.
To demonstrate this remarkable theorem, let us denote by \( y \) the value of the integral
\[
\int_{-x}^{+x} \phi(z) \, dz;
\]
then \( y \) will be the probability that an error is included between \(-x\) and \(+x\).

Let us put
\[
x = \psi(y), \quad d\psi(y) = \psi'(y) \, dy, \quad d\psi'(y) = \psi''(y) \, dy,
\]
then we have
\[
\psi(0) = 0 \quad \text{and} \quad \psi'(y) = \frac{1}{\phi(x) + \phi(-x)},
\]
and we conclude from this, remembering the hypotheses which have been made, that for \( y \) between 0 and 1, \( \psi'(y) \) is always increasing, or at least is non-decreasing, and that consequently \( \psi''(y) \) is always positive, or at least is non-negative. Now we have
\[
d.y \psi'(y) = \psi'(y) \, dy + y \psi''(y) \, dy,
\]
and consequently,
\[
y \psi'(y) - \psi(y) = \int_0^y y \psi''(y) \, dy;
\]
\( y \psi'(y) - \psi(y) \) therefore always has a positive value, or at least its value is never negative. It follows from this that \( 1 - \frac{\psi(y)}{y \psi'(y)} \) will always be positive and less than unity.

Let \( f \) be the value of this difference for \( y = \mu \); since \( \psi(\mu) = \lambda m \), we have
\[
f = 1 - \frac{\lambda m}{\mu \psi'(\mu)},
\]
from which we conclude
\[
\psi'(\mu) = \frac{\lambda m}{(1 - f) \mu}.
\]

With this established, let us consider this function
\[
\frac{\lambda m}{(1 - f) \mu} (y - \mu f),
\]
which we shall denote by \( F(y) \), and let us put
\[
d.F(y) = F'(y) \, dy;
\]
we then have, clearly,
\[
F(\mu) = \lambda m = \psi(\mu)
\]
\[
F'(\mu) = \frac{\lambda m}{(1 - f) \mu} = \psi'(\mu)
\]
Now since \( \psi'(y) \) increases continually (or at least does not decrease) as \( y \) increases, and since on the other hand, \( F'(y) \) is constant, the difference
\[
\psi'(y) - F'(y) = \frac{d[y \psi(y) - F(y)]}{dy}
\]
will be positive for \( y \) greater than \( \mu \), and negative for \( y \) less than \( \mu \). From this we conclude that the difference \( \psi(y) - F(y) \) is always positive, and consequently, \( \psi(y) \) will certainly be greater than \( F(y) \) in absolute value, as long as the function \( F(y) \) is positive, that is to say for \( y \) between \( \mu f \) and 1. The value of the integral
\[
\int_{\mu f}^{1} [F(y)]^2 \, dy
\]
will thus be less than the value of the integral
\[
\int_{\mu f}^{1} [\psi(y)]^2 \, dy,
\]
and \textit{a fortiori} less than
\[
\int_{0}^{1} [\psi(y)]^2 \, dy,
\]
that is to say less than \( m^2 \). Now the first of these integrals has the value
\[
\frac{\lambda^2 m^2 (1 - \mu f)^3}{3 \mu^2 (1 - f)^2};
\]
thus
\[
\lambda^2 < \frac{3 \mu^2 (1 - f)^2}{(1 - \mu f)^3},
\]
where \( f \) denotes, as one must remember, a number between 0 and 1.

If we consider \( f \) as a variable, the fraction
\[
\frac{3 \mu^2 (1 - f)^2}{(1 - \mu f)^3}
\]
will have as differential
\[
-\frac{3 \mu^2 (1 - f)}{(1 - \mu f)^4} (2 - 3 \mu + \mu f) \, df;
\]
this fraction will therefore be continually decreasing as \( f \) increases from 0 to 1, provided that one also has \( \mu < \frac{2}{3} \); its maximum value will correspond to \( f = 0 \) and will be equal to \( 3 \mu^2 \), so that in this case the coefficient \( \lambda \) will certainly be less than, or at least will not be greater than \( \mu \sqrt{3} \). Q.E.D.

On the other hand, when \( \mu \) is greater than \( \frac{2}{3} \), the value of the function will be a maximum when
\[
2 - 3 \mu + \mu f = 0,
\]
that is to say when
\[
f = 3 - \frac{2}{\mu},
\]
and this maximum value will be
\[
\frac{4}{9(1 - \mu)};
\]
consequently, in this case, the coefficient \( \lambda \) is no greater than \( \frac{2}{3\sqrt{1-\mu}} \), as we stated above.

Let us, for example, take \( \mu = \frac{1}{2} \); then \( \lambda \) cannot exceed \( \sqrt{\frac{3}{4}} \), that is to say that the probable error cannot exceed \( 0.8660254\,m \), to which it is equal in the first case examined (section 9); we deduce readily from our theorem that \( \mu \) is not less than \( \lambda \sqrt{\frac{3}{4}} \), as long as \( \lambda \) is less than \( \sqrt{\frac{3}{4}} \), and on the other hand, is not less than \( 1 - \frac{4}{9\lambda^2} \) when \( \lambda \) is greater than \( \sqrt{\frac{4}{3}} \).

11.

Since the integral

\[
\int_{-\infty}^{\infty} x^4 \phi(x) \, dx,
\]

arises in several problems which we have to deal with, it will be useful to evaluate it for several particular cases.

Let us put

\[
\int_{-\infty}^{\infty} x^4 \phi(x) \, dx = n^4.
\]

I. When

\[
\phi(x) = \frac{1}{2a},
\]

for the values of \( x \) between \(-a\) and \(+a\), one has

\[
n^4 = \frac{a^4}{5} = \frac{9}{5} m^4.
\]

II. When

\[
\phi(x) = \frac{a \mp x}{a^2}
\]

(case II, section 9), \( x \) being still bounded between \(-a\) and \(+a\), we have

\[
n^4 = \frac{1}{15} a^4 = \frac{12}{5} m^4.
\]

III. In the third case, when

\[
\phi(x) = \frac{e^{-x^2/h^2}}{h\sqrt{\pi}}
\]

we find, according to the results obtained in the memoir cited above,

\[
n^4 = \frac{3}{4} h^4 = 3 m^4.
\]

One can furthermore show that under the hypotheses of the preceding section, the ratio \( n^4/m^4 \) is never less than \( 9/5 \).
12.

Let us denote by \(x, x', x'', \ldots\) the errors committed in observations of the same type, and let us suppose that these errors are independent of each other. As before, let \(\phi(x)\) denote the relative probability of the error \(x\); let us consider a quantity \(y\), which is a rational function of the variables \(x, x', x'', \ldots\).

The multiple integral

\[
\int \phi(x)\phi(x')\phi(x'') \cdots dx dx' dx'' \cdots,
\]

extended over all the values of the variables \(x, x', x'', \ldots\), for which the value of \(y\) falls between the given limits 0 and \(\eta\), represents the probability that the value of \(y\) lies between 0 and \(\eta\). Now this integral is evidently a function of \(\eta\). Let us denote its differential by \(\psi(\eta)d\eta\), so that the integral considered is equal to

\[
\int_0^\eta \psi(\eta) d\eta
\]

Then \(\psi(\eta)\) represents the relative probability of an arbitrary value of \(y\). Since \(x\) may be regarded as a function of the variables \(y, x', x'', \ldots\), which we shall denote as \(f(y, x', x'', \ldots)\), the integral (1) will take the form

\[
\int \phi(f(y, x', x'', \ldots)) \frac{df(y, x', x'', \ldots)}{dy} \phi(x')\phi(x'') \cdots dy dx' dx'' \cdots,
\]

where \(y\) is to vary between 0 and \(\eta\), and the other variables take on all the values for which \(f(y, x', x'', \ldots)\) is real.

One has therefore

\[
\psi(y) = \int \phi(f(y, x', x'', \ldots)) \frac{df(y, x', x'', \ldots)}{dy} \phi(x')\phi(x'') \cdots dx' dx'' \cdots;
\]

where the integration, in which \(y\) is to be considered as a constant, extends over all the values of the variables \(x', x'', \ldots\), for which \(f(y, x', x'', \ldots)\) takes a real value.

13.

The preceding integration requires, it is true, the knowledge of the function \(\phi\), which is unknown in the majority of cases. Even if this function were known, the calculation would, in general, surpass the resources of analysis. Hence it is impossible to obtain the probability for each value of \(y\); but it is different if one wishes only the average value of \(y\), which is given by the integral

\[
\int y\psi(y) dy,
\]

extended over all possible values of \(y\).

If, from the nature of the function, or because of bounds imposed on \(x, x', x'', \ldots\), \(y\) does not take on all values, one may suppose that \(\psi(y)\) vanishes for all those values which \(y\) does not take on, and one can then extend the integration from \(-\infty\) to \(+\infty\).
The integral
\[ \int y\psi(y) \, dy, \]
taken between the fixed limits \( \eta \) and \( \eta' \), is equal to
\[ \int y\phi(x)\phi(x')\phi(x'') \cdots dx \, dx' \, dx'' \cdots, \]
taken for \( y \) between \( \eta \) and \( \eta' \) and extended over all the values of the variables \( x', x'' \), etc., for which \( f \) is real. Consequently this integral is equal to the integral
\[ \int y\phi(x)\phi(x')\phi(x'') \cdots dx \, dx' \, dx'' \cdots, \]
in which \( y \) is expressed as a function of \( x, x', x'' \), etc., the integration extending over all values of the variables for which \( y \) is included between \( \eta \) and \( \eta' \). Hence the integral
\[ \int_{-\infty}^{\infty} y\phi(y) \, dy \]
can be put into the form
\[ \int y\phi(x)\phi(x')\phi(x'') \cdots dx \, dx' \, dx'' \cdots, \]
the integration being extended over all real values of \( x, x', x'' \), that is to say from \( x = -\infty \) to \( x = +\infty \), \( x' = -\infty \) to \( x = +\infty \), etc.

14.

If the function \( y \) reduces to a sum of terms of the form
\[ Ax^\alpha x'^\beta x''^\gamma \cdots, \]
the value of the integral
\[ \int y\psi(y) \, dy, \]
extended over all values of \( y \), that is to say the average value of \( y \), will be equal to a sum of terms of the form
\[ A \int_{-\infty}^{\infty} x^\alpha \phi(x) \, dx \int_{-\infty}^{\infty} x'^\beta \phi(x') \, dx' \int_{-\infty}^{\infty} x''^\gamma \phi(x'') \, dx'' \cdots. \]

In other words, the average value of \( y \) is equal to a sum of terms obtained from those which make up \( y \), by replacing \( x^\alpha, x'^\beta, x''^\gamma \), etc. in the expression for \( y \) by their average values. The proof of this important theorem could easily be deduced from other considerations.
15.

Let us apply the preceding theorem to the case in which one has

$$y = \frac{x^2 + x'^2 + x''^2 + \cdots}{\sigma}$$

where \(\sigma\) denotes the number of terms in the numerator.

One finds at once that the average value of \(y\) is equal to \(m^2\) (where the letter \(m\) still has the same meaning as in section 5). The actual value of \(y\) may be greater than or less than its average value, just as the true value of \(x^2\) may in each case be greater than or less than \(m^2\); but the probability that the random value of \(y\) does not differ appreciably from \(m^2\) approaches certainty as \(\sigma\) becomes larger. To show this more clearly, since it is impossible to find this probability exactly, we shall find the standard error to be expected in setting \(y = m^2\). According to previous results (section 6), this error will be the square root of the average of the function

$$\left(\frac{x^2 + x'^2 + x''^2 + \cdots}{\sigma} - m^2\right)^2.$$

To find this, it suffices to note that the average value of a term such as \(\frac{x^4}{\sigma^2}\) is equal to \(\frac{n^4}{\sigma^2}\) (where \(n\) has the same meaning as in section 11), and that the average value of a term such as \(\frac{2x^2x'^2}{\sigma^2}\) is equal to \(\frac{2m^4}{\sigma^2}\); consequently, the average value of this function will be

$$\frac{n^4 - m^4}{\sigma}.$$

From this we conclude that if the number of random errors is sufficiently large, the value of \(m\) will be given, with high probability, by the formula

$$m = \sqrt{\frac{x^2 + x'^2 + x''^2 + \cdots}{\sigma}}$$

and the standard error to be expected in the determination of \(m^2\) will be equal to

$$\sqrt{\frac{n^4 - m^4}{\sigma}}.$$

Since this last formula contains the quantity \(n\), if one wishes only to get an idea of the degree of precision of this determination, it suffices to adopt some particular hypothesis for the function \(\phi\).

For example, if we take the third hypothesis of sections 9 and 11, this error will be equal to \(m^2 \sqrt{\frac{2}{n^2}}\). If one prefers, one may obtain an approximate value of \(n^4\) by means of the errors themselves, by using the formula

$$\frac{x^4 + x'^4 + x''^4 + \cdots}{\sigma}.$$
One may state in general that to obtain double precision in this determination will require four times the number of observations, that is to say that the weight of the determination is proportional to the number \( \sigma \).

One sees in the same way that if the errors of the observations contain a constant part, one can deduce from their arithmetic mean a value of the constant part, and this value will be the more accurate, the greater the number of observations. In this determination, the standard error to be expected is given by

\[
\sqrt{\frac{m^2 - k^2}{\sigma}}
\]

where \( k \) denotes the constant part and \( m \) the standard error of the uncorrected observations. It is given simply by \( \frac{m}{\sqrt{n}} \), if \( m \) represents the standard error of the observations corrected for the constant part (see section 8).

16.

In sections 12 to 15 we supposed that the errors \( x, x', x'' \), etc., belong to the same type of observation, so that the probability for each of these errors was represented by the same function. But it is clear that the general principles expounded in sections 12 to 14, can easily be applied to the more general case where the probabilities of the errors \( x, x', x'' \), etc. are represented by different functions

\[
\phi(x), \quad \phi'(x'), \quad \phi''(x''), \ldots,
\]

that is to say when these errors belong to observations which do not have the same degree of precision. Let us suppose that \( x \) denotes the error of an observation whose standard error to be expected is \( m; x', x'' \), etc., the errors of other observations whose standard errors to be expected are respectively \( m', m'' \), etc.: then the average value of the sum

\[
x^2 + x'^2 + x''^2 + \ldots
\]

will be

\[
m^2 + m'^2 + m''^2 + \ldots.
\]

Now if one knows as well that the quantities \( m, m', m'' \), etc. are respectively proportional to the numbers \( 1, \mu', \mu'' \), etc., the average value of the expression

\[
\frac{x^2 + x'^2 + x''^2 + \ldots}{1 + \mu'^2 + \mu''^2 + \ldots}
\]

will be equal to \( m^2 \). But if we take for \( m^2 \) the value which this expression assumes when we substitute in it the values of \( x, x', x'' \), etc. which are obtained by chance, the standard error affecting this determination will be, according to the preceding section,

\[
\frac{\sqrt{n^2 + n'^2 + n''^2 + \ldots - m^2 - m'^2 - m''^2 - \ldots}}{1 + \mu'^2 + \mu''^2 + \ldots}
\]

where \( n', n'' \), etc. have the same meaning with respect to the second and third observations as \( n \) has with respect to the first. If one supposes the numbers \( n, n', n'' \), etc. to be proportional to \( m, m', m'' \), etc., this standard error will be equal to

\[
\frac{\sqrt{(n^2 - m^2)(1 + \mu'^2 + \mu''^2 + \ldots)}}{1 + \mu'^2 + \mu''^2 + \ldots}
\]
but this way of determining an approximate value of $m$ is not the most advantageous. Let us consider the more general expression

$$y = x^2 + \alpha' x'^2 + \alpha'' x''^2 + \cdots$$

whose average value is also $m^2$, for any coefficients $\alpha', \alpha''$, etc. The standard error to be expected when one takes $y$ as a value of $m^2$, $y$ being calculated from the random errors $x, x', x''$, etc., will be, according to the preceding principles, given by the formula

$$\sqrt{\frac{n^4 - m^4}{1 + \alpha' \mu'^2 + \alpha'' \mu''^2 + \cdots}} + \alpha' \mu'^4 + \alpha'' \mu''^4 + \cdots$$

For this error to be as small as possible, it is necessary to put

$$\alpha' = \frac{n^4 - m^4}{n'^4 - m'^4} \mu'^2$$
$$\alpha'' = \frac{n^4 - m^4}{n''^4 - m''^4} \mu''^2$$

$$\vdots$$

These quantities cannot be evaluated unless one knows the ratios $\frac{n}{m}, \frac{n'}{m'}, \frac{n''}{m''}$, etc. Since one is ignorant of their exact value\(^4\), it seems safest to suppose them equal to each other (see section 11) and one will then have

$$\alpha' = \frac{1}{\mu'^2}, \quad \alpha'' = \frac{1}{\mu''^2}, \quad \ldots,$$

that is to say that the coefficients $\alpha', \alpha''$, etc. should be assumed equal to the relative weights of the various observations, taking as unit the weight of that to which the error $x$ corresponds. With this agreed on, let $\sigma$ denote, as above, the number of errors in question; the average value of the expression

$$\frac{x^2 + \alpha' x'^2 + \alpha'' x''^2 + \cdots}{\sigma}$$

will be equal to $m^2$, and when we take as the true value of $m^2$ the value of this expression determined by means of random errors $x, x', x''$, etc., the standard error to be expected will be

$$\frac{n^2 + \alpha'^2 n'^4 + \alpha''^2 n''^4 + \cdots - \sigma m^4}{\sigma}$$

\(^4\)One can scarcely conceive the possibility of exactly determining $\mu, \mu', \mu''$, etc., except in the particular case where, by nature of the function $\phi$, the errors $x, x', x''$, etc., which are proportional to $\mu, \mu', \mu''$, etc. are equally likely, that is to say the case where

$$\phi(x) = \mu' \phi'(\mu' x) = \mu'' \phi''(\mu'' x) \cdots .$$

---G.
Finally, if it is permissible to suppose that the quantities $n, n', n''$, etc. are proportional to $m, m', m''$, etc., this expression reduces to

$$\sqrt{n^4 - m^4}$$

and this result is identical with the one which we obtained in the case where all the observations were of the same type.

17.

When we obtain the value of a certain quantity from an observation which is not absolutely precise, and this quantity is connected analytically with an unknown magnitude, the result of this observation may furnish an erroneous value for the unknown, but there is nothing arbitrary about its determination which might give rise to a choice of methods.

But if several functions of the same unknowns are given by imperfect observations, each observation will furnish a value for the unknown, and one can also obtain values by combination of several observations. There are obviously infinitely many ways of arriving at an answer; the answer will be subject, in every case, to a possible error. The standard error to be expected may be greater or less, depending on the method of combination adopted.

The same thing will happen if several observed quantities depend on several unknowns simultaneously. According to whether the number of observations is equal to the number of unknowns, or smaller or greater than this number, the problem will be determinate, or indeterminate or overdetermined (at least in general), and, in this third case, the observations can be combined in an infinity of different ways to give the values of the unknowns. Among these combinations, the most advantageous should be chosen, that is to say, those which furnish values whose standard error to be expected is as small as possible. This problem is certainly the most important which the application of mathematics to natural philosophy presents.

In our *Theoria Motus Corporum Coelestium* we have shown how to find the most probable values of the unknowns when one knows the probability distribution of the errors of the observations, and since, in almost every case, this distribution remains hypothetical by its very nature, we applied this theory to the very plausible hypothesis that the probability of the error $x$ is proportional to $e^{-h^2x^2}$; from this hypothesis came the method which I followed, especially in astronomical calculations, and which is now employed by the majority of calculators under the name of *Method of least squares*.

Considering the question from another point of view, Laplace subsequently showed that this principle is preferable to all others, whatever be the probability distribution of the errors, provided that the number of observations is extremely large. But when this number is restricted, the question remains open; so that if one rejected our hypothetical law, the method of least squares would be preferable to others solely because it leads to the simplest calculations.

We hope, therefore, to satisfy mathematicians by demonstrating in this memoir that the method of least squares furnishes the most advantageous combinations of the observations, not merely approximately, but in an absolute sense, and this for an arbitrary
probability distribution for the errors and an arbitrary number of observations, provided that one takes as standard error, not the definition of Laplace, but that which we have given in sections 5 and 6.

It is necessary to give warning here that in the following discussion we shall be concerned only with random errors from which the constant part has been subtracted. It is up to the observer to avoid the causes of constant errors. We reserve the examination of the case in which the observations are affected by an unknown constant error for another occasion, and shall treat this question in another memoir.

18.

PROBLEM

Let \( U \) be a given function of the unknowns \( V, V', V'' \), etc.; we wish to know the standard error \( M \) to be expected in the determination of the value of \( U \) when, in place of the true values of \( V, V', V'' \), etc., one takes the values deduced from observations which are independent of each other, where \( m, m', m'' \), etc. are the standard errors corresponding to these various observations.

Solution. — Let us denote by \( e, e', e'' \), etc. the errors of the observed values \( V, V', V'' \), etc.; the error in the value of the function \( U \) arising from these can be expressed by the linear function

\[
\lambda e + \lambda' e' + \lambda'' e'' + \cdots = E
\]

where \( \lambda, \lambda', \lambda'' \), etc. represent the derivatives \( \frac{dU}{dV} \), etc. evaluated at the true values of \( V, V', V'' \), etc.

It is clear that this is the value of \( E \) if one assumes that the observations are sufficiently exact that the squares and products of the errors are negligible. It follows from this that the average value of \( E \) is zero, since we suppose that the errors of the observations have no constant part. Now the standard error \( M \) to be expected in the value of \( U \) will be the square root of the average value of \( E^2 \), that is to say that \( M^2 \) will be the average value of the sum

\[
\lambda^2 e^2 + \lambda'^2 e'^2 + \lambda''^2 e''^2 + \cdots + 2\lambda\lambda' e e' + 2\lambda\lambda'' e e'' + 2\lambda'\lambda'' e' e'' + \cdots;
\]

but the average value of \( \lambda^2 e^2 \) is \( \lambda^2 m^2 \), that of \( \lambda'^2 e'^2 \) is \( \lambda'^2 m'^2 \), etc., and finally the average values of the products \( 2\lambda\lambda' e e' \) are all zero; thus we shall have

\[
M = \sqrt{\lambda^2 m^2 + \lambda'^2 m'^2 + \lambda''^2 m''^2 + \cdots}
\]

It is suitable to add several comments on this solution.

I. Since we neglect all powers of the errors larger than the first, we can in our formula take for \( \lambda, \lambda', \lambda'' \), etc. the values of the derivatives \( \frac{du}{dv} \), etc., obtained from the observed values of \( V, V', V'' \), etc. Whenever \( U \) is a linear function, this substitution is rigorously exact.

II. If in place of the standard errors one prefers to introduce the weights of the observations, let \( p, p', p'' \), etc. be the respective weights relative to an arbitrary unit, and \( P \) the weight of the value of \( U \); one will have

\[
P = \frac{1}{\frac{\lambda^2}{p} + \frac{\lambda'^2}{p'} + \frac{\lambda''^2}{p''} + \cdots}
\]
III. Let \( T \) be another function of \( V, V', V'', \) etc., and put

\[
\frac{dT}{dV} = \kappa, \quad \frac{dT}{dV'} = \kappa', \quad \frac{dT}{dV''} = \kappa'', \ldots
\]

The error in \( T \) which arises from taking the results furnished by observation for \( V, V', V'', \) etc., will be

\[
\kappa e + \kappa' e' + \kappa'' e'' + \cdots = E',
\]

and the standard error to be expected in this determination will be

\[
\sqrt{\kappa^2 m^2 + \kappa'^2 m'^2 + \kappa''^2 m''^2 + \cdots}
\]

It is clear that the errors \( E \) and \( E' \) will not be independent of each other, and that the average value of the product \( EE' \) will not be zero like the average value of \( ee' \); it will be instead equal to

\[
\kappa \lambda m^2 + \kappa' \lambda' m'^2 + \kappa'' \lambda'' m''^2 + \cdots
\]

IV. The problem includes the case where the values of the quantities \( V, V', V'', \) etc. are not given immediately by observation, but are obtained from arbitrary combinations of direct observations. In order for this extension to be legitimate, it is necessary that these quantities be determined independently, that is to say that they should be furnished by different observations. If this condition of independence were not satisfied, the formula giving the value of \( M \) would no longer be correct. If, for example, one single observation had been used in the determination of both \( V \) and \( V' \), the errors \( e \) and \( e' \) would no longer be independent, and the average value of the product \( ee' \) would no longer be zero. In this case, if one knew the relation of \( V \) and \( V' \) to the results of the simple observations from which they were derived, one could calculate the average value of the product \( ee' \), as is indicated in remark III, and thus correct the formula for \( M \).

19.

Let \( V, V', V'' \), etc. be functions of the unknowns \( x, y, z, \) etc.; let \( \varpi \) be the number of these functions, \( \rho \) the number of unknowns; suppose that \( L, L', L'' \), etc. have been obtained from observations, directly or indirectly, as values of the functions, \( V, V', V'' \), etc., in such a way that these determinations are completely independent of each other. If \( \rho \) is greater than \( \varpi \), the problem of finding the unknowns is indeterminate. If \( \rho \) equals \( \varpi \), each of the unknowns \( x, y, z, \) etc. may be considered to be calculated as functions of \( V, V', V'', \) etc., so that the values of the former can be derived from the observed values of the latter, and the results of the preceding section permit us to calculate the relative precision of these various determinations. If \( \rho \) is smaller than \( \varpi \), each unknown \( x, y, z, \) etc. can be expressed as a function of \( V, V', V'', \) etc., in an infinity of ways, and these values will be, in general, different; they would coincide if (contrary to our hypothesis) the observations were perfectly exact. Moreover, it is clear that the various combinations furnish results whose precisions will be, in general, different.

However, if, in the second and third case, the quantities \( V, V', V'' \), etc. are such that \( \varpi - \rho + 1 \) or more of them can be considered as functions of the others, the problem
is over-determined relative to these latter functions and indeterminate relative to the
unknowns \( x, y, z, \) etc.; one could not then determine these latter unknowns even if the
values of the functions \( V, V', V'', \) etc. were known exactly: but we exclude this case
from our discussion.

If \( V, V', V'', \) etc. are not linear functions of the unknowns, one can put them in
this form by replacing the original unknowns by the difference between them and their
approximate values, which one supposes known, the standard errors to be expected in
the determinations

\[
V = L, \quad V' = L', \quad V'' = L'', \ldots
\]

being denoted respectively by \( m, m', m'', \) etc., and the weights of these determinations
by \( p, p', p'', \) etc., so that

\[
pm^2 = p'm'^2 = p''m''^2 = \ldots
\]

We suppose that the ratios of the standard errors as well as the weights are known,
one of them being taken arbitrarily. Then, if we put

\[
(V - L)\sqrt{p} = \nu, \quad (V' - L')\sqrt{p} = \nu', \ldots,
\]

the situation is just the same as if direct observations, of equal precision and with
standard error \( m\sqrt{p}, \) had given

\[
\nu = 0, \quad \nu' = 0, \quad \nu'' = 0, \ldots
\]

### 20. PROBLEM

Let us denote by \( v, v', v'', \) etc. the following linear functions of the variables \( x, y, z, \)
etc.,

\[
\begin{align*}
\nu &= ax + by + cz + \cdots + l, \\
\nu' &= a'x + b'y + c'z + \cdots + l', \\
\nu'' &= a''x + b''y + c''z + \cdots l'', \\
&\vdots
\end{align*}
\]

(1)

It is required to find among all systems of coefficients \( \kappa, \kappa', \kappa'', \) etc. for which

\[
\kappa \nu + \kappa' \nu' + \kappa'' \nu'' + \cdots = x - k,
\]

\( k \) being independent of \( x, y, z, \) etc., the one for which \( \kappa^2 + \kappa'^2 + \kappa''^2 + \cdots \) is a minimum.

**Solution.** — Let us put

\[
\begin{align*}
\alpha \nu + \alpha' \nu' + \alpha'' \nu'' + \cdots &= \xi, \\
\beta \nu + \beta' \nu' + \beta'' \nu'' + \cdots &= \eta, \\
\gamma \nu + \gamma' \nu' + \gamma'' \nu'' + \cdots &= \zeta, \\
&\vdots
\end{align*}
\]

(2)
ξ, η, ζ will be linear functions of \( x, y, z \), and one will have

\[
\begin{align*}
\xi &= x \sum a^2 + y \sum ab + z \sum ac + \cdots + \sum a, \\
\eta &= x \sum ab + y \sum b^2 + z \sum bc + \cdots + \sum b, \\
\zeta &= x \sum ac + y \sum bc + z \sum c^2 + \cdots + \sum c,
\end{align*}
\]

(3)

where

\[
\sum a^2 = a^2 + a'^2 + a''^2 + \cdots,
\]

and similarly for the other \( \sum \)'s.

The number of the quantities \( \xi, \eta, \zeta \), etc. is equal to \( \wp \), the number of the unknowns \( x, y, z \), etc.; one can, therefore, obtain, by elimination, an equation of the following form,

\[
x = A + (\alpha\alpha)\xi + (\alpha\beta)\eta + (\alpha\gamma)\zeta + \cdots,
\]

which will be satisfied identically when one substitutes for \( \xi, \eta, \zeta \), the values given in (3). Consequently, if one puts

\[
\begin{align*}
a(\alpha\alpha) + b(\alpha\beta) + c(\alpha\gamma) + \cdots &= \alpha, \\
a'(\alpha\alpha) + b'(\alpha\beta) + c'(\alpha\gamma) + \cdots &= \alpha', \\
a''(\alpha\alpha) + b''(\alpha\beta) + c''(\alpha\gamma) + \cdots &= \alpha'',
\end{align*}
\]

(4)

one will have identically

\[
\alpha\nu + \alpha'\nu' + \alpha''\nu'' + \cdots = x - A
\]

This equation shows that one must include the system

\[
\kappa = \alpha, \quad \kappa' = \alpha', \quad \kappa'' = \alpha'', \quad \ldots
\]

among the different systems of coefficients \( \kappa, \kappa', \kappa'', \) etc. One has moreover, for an arbitrary system,

\[
(\kappa - \alpha)\nu + (\kappa' - \alpha')\nu' + (\kappa'' - \alpha'')\nu'' + \cdots = A - k
\]

and this equation, being an identity, implies the following:

\[
(\kappa - \alpha)a + (\kappa' - \alpha')a' + (\kappa'' - \alpha'')a'' + \cdots = 0,
\]

\[
(\kappa - \alpha)b + (\kappa' - \alpha')b' + (\kappa'' - \alpha'')b'' + \cdots = 0,
\]

\[
(\kappa - \alpha)c + (\kappa' - \alpha')c' + (\kappa'' - \alpha'')c'' + \cdots = 0,
\]

\[\vdots\]

\[\text{---G.}\]

\[\text{---G.}\]
Let us multiply these equations by \((\alpha\alpha), (\alpha\beta), (\alpha\gamma)\), etc. and add them; we obtain, in virtue of the equations (4),

\[
(k - \alpha)\alpha + (k' - \alpha')\alpha' + (k'' - \alpha'')\alpha'' + \cdots = 0,
\]

that is to say

\[
k^2 + k'^2 + k''^2 + \cdots = \alpha^2 + \alpha'^2 + \alpha''^2 + \cdots + (k - \alpha)^2 + (k' - \alpha')^2 + \cdots;
\]

and consequently, the sum

\[
k^2 + k'^2 + k''^2 + \cdots
\]

will have a minimum value when one sets

\[
k = \alpha, \quad k' = \alpha', \quad k'' = \alpha'', \quad \ldots
\]

Moreover, this minimum value is obtained in the following way.

The equation (5) shows that one has

\[
\begin{align*}
\alpha\alpha + a\alpha' + a''\alpha'' + \cdots &= 1, \\
b\alpha + b\alpha' + b''\alpha'' + \cdots &= 0, \\
c\alpha + c\alpha' + c''\alpha'' + \cdots &= 0, \\
&\quad \vdots
\end{align*}
\]

Let us multiply these equations by \((\alpha\alpha), (\alpha\beta), (\alpha\gamma)\), etc., respectively, and add; taking the relations (4) into account, we find

\[
\alpha^2 + \alpha'^2 + \alpha''^2 + \cdots = (\alpha\alpha).
\]

When the observations have furnished the approximate equations

\[
\nu = 0, \quad \nu' = 0, \quad \nu'' = 0, \quad \ldots
\]

it is necessary, in order to determine the unknown \(x\), to choose a combination of the following form,

\[
k\nu + k'\nu' + k''\nu'' + \cdots = 0
\]

such that the unknown \(x\) has a coefficient equal to 1, and the other unknowns are eliminated.

The weight of this determination will be, from section 18,

\[
\frac{1}{k^2 + k'^2 + k''^2 + \cdots}
\]

According to the preceding section, one obtains the most suitable determination by taking

\[
k = \alpha, \quad k' = \alpha', \quad k'' = \alpha'', \quad \ldots
\]
then \( x \) will have the value \( A \). Clearly one obtains the same value without computing the multipliers \( \alpha, \alpha', \alpha'' \), etc. by carrying out the elimination process on the equations

\[
\xi = 0, \quad \eta = 0, \quad \zeta = 0, \quad \ldots;
\]

the weight of this determination will be

\[
\frac{1}{(\alpha \alpha)},
\]

and the standard error to be expected

\[
m\sqrt{p(\alpha \alpha)} = m'\sqrt{p'(\alpha \alpha)} = m''\sqrt{p''(\alpha \alpha)} = \ldots
\]

A similar procedure would lead to the most suitable values of the other unknowns \( y, z, \) etc., which will be those obtained by elimination from the equations

\[
\xi = 0, \quad \eta = 0, \quad \zeta = 0, \quad \ldots
\]

If we denote by \( \Omega \) the sum

\[
\nu^2 + \nu'^2 + \nu''^2 + \ldots,
\]

or, what comes to the same thing,

\[
p(V-L)^2 + p'(V'-L')^2 + p''(V''-L'')^2 + \ldots,
\]

we shall clearly have

\[
2\xi = \frac{d\Omega}{dx}, \quad 2\eta = \frac{d\Omega}{dy}, \quad 2\zeta = \frac{d\Omega}{dz}, \quad \ldots;
\]

consequently, the values of the unknowns obtained from the most suitable combination, and which we may call the \textit{most likely} values, are precisely those which give \( \Omega \) a minimum value. Now \( V-L \) represents the difference between the observed value and the calculated value; hence the most likely values of the unknowns are those which make the sum of the squares of the differences between the calculated and observed values of the quantities \( V, V', V'' \), etc., a minimum, these squares having been multiplied by the respective weights of the observations. I had established this principle some time ago by another method (\textit{Theoria Motus Corporum Coelestium}).

If one wishes to find the relative precision of each of these determinations one must obtain from the equations (3) the values of \( x, y, z \), etc. which will be presented in the following form:

\[
\begin{align*}
x &= A + (\alpha \alpha)\xi + (\alpha \beta)\eta + (\alpha \gamma)\zeta + \cdots, \\
y &= B + (\beta \alpha)\xi + (\beta \beta)\eta + (\beta \gamma)\zeta + \cdots, \\
z &= C + (\gamma \alpha)\xi + (\gamma \beta)\eta + (\gamma \gamma)\zeta + \cdots.
\end{align*}
\]

The most likely values of the unknowns \( x, y, z, \) etc. will be \( A, B, C \), etc. The weights of these determinations will be

\[
\frac{1}{(\alpha \alpha)}, \quad \frac{1}{(\beta \beta)}, \quad \frac{1}{(\gamma \gamma)}, \quad \ldots,
\]
and the standard errors to be expected

for \( x, m \sqrt{p(\alpha\alpha)} = m' \sqrt{p'(\alpha\alpha)}, \ldots \),

for \( y, m \sqrt{p(\beta\beta)} = m' \sqrt{p'(\beta\beta)}, \ldots \),

for \( z, m \sqrt{p(\gamma\gamma)} = m' \sqrt{p'(\gamma\gamma)}, \ldots \),

which agrees with the results obtained previously (Theoria Motus Corporum Coelestium).

22.

The case where there is only one unknown is the most frequent and is the simplest of all. One has then

\[ V = x, \quad V' = x, \quad V'' = x, \quad \ldots \; ; \]

it will be useful to say a few words about it.

One will have

\[ a = \sqrt{p}, \quad a' = \sqrt{p'}, \quad a'' = \sqrt{p''}, \quad \ldots \]

\[ l = -L \sqrt{p}, \quad l' = -L' \sqrt{p'}, \quad l'' = -L'' \sqrt{p''}, \quad \ldots \]

and, consequently,

\[ (\alpha\alpha) = \frac{1}{p + p' + p'' + \cdots} ; \]

\[ A = \frac{pL + p'L' + p''L'' + \cdots}{p + p' + p'' + \cdots} . \]

Thus, if, by several observations of unequal precision having the weights \( p, p', p'' \), etc., respectively, one has found, for one single quantity, a first value \( L \), a second \( L' \), a third \( L'' \), etc., the most likely value will be

\[ \frac{pL + p'L' + p''L'' + \cdots}{p + p' + p'' + \cdots} ; \]

and the weights of this determination will be

\[ p + p' + p'' + \cdots \]

If all the observations are of equal precision, the most likely value will be

\[ \frac{L + L' + L'' + \cdots}{\varpi} \]

that is to say the arithmetic mean of the observed values; taking as unit the weight of a single observation, the weight of the average will be \( \varpi \).
It still remains to expound some researches intended to extend and clarify the preceding theory.

First let us see whether the elimination which gives the variables $x, y, z$, etc., as functions of $\xi, \eta, \zeta$, etc., is always possible. Since the number of equations is equal to the number of unknowns, one knows that this elimination will be possible if $\xi, \eta, \zeta$, etc., are independent of each other; in the contrary case, it will be impossible.

Suppose for the moment that $\xi, \eta, \zeta$, etc. were not independent, but that

$$0 = F\xi + G\eta + H\zeta + \cdots + K$$

were an identical relation between these quantities; we would conclude from this that

$$F \sum a^2 + G \sum ab + H \sum ac + \cdots = 0,$$
$$F \sum ab + G \sum b^2 + H \sum bc + \cdots = 0,$$
$$F \sum ac + G \sum be + H \sum c^2 + \cdots = 0,$$
$$\vdots$$
$$F \sum al + G \sum bl + H \sum cl + \cdots = -K.$$

Putting

$$\left\{ \begin{align*}
aF + bG + cH + \cdots &= \Theta \\
a'F + b'G + c'H + \cdots &= \Theta' \\
a''F + b''G + c''H + \cdots &= \Theta'' \\
\vdots
\end{align*} \right. \quad (1)$$

we get

$$a\Theta + a'\Theta' + a''\Theta'' + \cdots = 0,$$
$$b\Theta + b'\Theta' + b''\Theta'' + \cdots = 0,$$
$$c\Theta + c'\Theta' + c''\Theta'' + \cdots = 0,$$
$$\vdots$$
$$l\Theta + l'\Theta' + l''\Theta'' + \cdots = -K.$$

Multiplying the equations (1) by $\Theta, \Theta', \Theta''$, etc. respectively, and adding, we get

$$\Theta^2 + \Theta'^2 + \Theta''^2 + \cdots = 0,$$
and this equation implies the following:

$$\Theta = 0, \quad \Theta' = 0, \quad \Theta'' = 0, \quad \ldots$$
Hence we conclude, in the first place, that \( K = 0 \). Secondly, the equations (1) show that the functions \( \nu, \nu', \nu'' \), etc. are such that their values do not change when the variables \( x, y, z \), etc. take increments proportional to \( F, G, H \), etc. It will clearly be the same for the functions \( V, V', V'' \), etc. Now this can occur only in the case in which it would be impossible to determine \( x, y, z \), etc. from the values of \( V, V', V'' \), etc., even if the latter were known exactly; but then the problem would be indeterminate by nature, and we shall exclude this case from our discussions.

24.

Let us denote by \( \beta, \beta', \beta'' \), etc., multipliers which play the same role relative to the unknown \( y \) as do the multipliers \( \alpha, \alpha', \alpha'' \), etc. relative to the unknown \( x \), that is to say such that

\[
\begin{align*}
 a(\beta \alpha) + b(\beta \beta) + c(\beta \gamma) + & \cdots = \beta, \\
 a'(\beta \alpha) + b'(\beta \beta) + c'(\beta \gamma) + & \cdots = \beta', \\
 a''(\beta \alpha) + b''(\beta \beta) + c''(\beta \gamma) + & \cdots = \beta'', \\
& \vdots
\end{align*}
\]

one has then identically

\[
\beta \nu + \beta' \nu' + \beta'' \nu'' + \cdots = y - B
\]

Let \( \gamma, \gamma', \gamma'' \), etc. be the analogous multipliers relative to the variable \( z \) so that one has

\[
\begin{align*}
 a(\gamma \alpha) + b(\gamma \beta) + c(\gamma \gamma) + & \cdots = \gamma, \\
 a'(\gamma \alpha) + b'(\gamma \beta) + c'(\gamma \gamma) + & \cdots = \gamma', \\
 a''(\gamma \alpha) + b''(\gamma \beta) + c''(\gamma \gamma) + & \cdots = \gamma'', \\
& \vdots
\end{align*}
\]

and, consequently,

\[
\gamma \nu + \gamma' \nu' + \gamma'' \nu'' + \cdots = z - C
\]

In the same way in which we found (section 20)

\[
\sum \alpha a = 1, \quad \sum \alpha b = 0, \quad \sum \alpha c = 0, \ldots, \sum \alpha l = -A,
\]

we obtain here

\[
\sum \beta a = 0, \quad \sum \beta b = 1, \quad \sum \beta c = 0, \ldots, \sum \beta l = -B, \\
\sum \gamma a = 0, \quad \sum \gamma b = 0, \quad \sum \gamma c = 1, \ldots, \sum \gamma l = -C;
\]

and so on.

One has also, as in section 20,

\[
\sum \alpha^2 = (\alpha \alpha), \quad \sum \beta^2 = (\beta \beta), \quad \sum \gamma^2 = (\gamma \gamma), \ldots
\]
Let us multiply the values of $\alpha, \alpha', \alpha''$, etc. (section 20), respectively by $\beta, \beta', \beta''$, etc., and add; we shall have

$$\alpha\beta + \alpha'\beta' + \alpha''\beta'' + \cdots = (\alpha\beta),$$

that is to say

$$\sum \alpha\beta = (\alpha\beta)$$

Multiplying $\beta, \beta', \beta''$, etc. by $\alpha, \alpha', \alpha''$, etc., respectively, and adding one finds

$$\alpha\beta + \alpha'\beta' + \alpha''\beta'' + \cdots = (\beta\alpha);$$

thus

$$(\alpha\beta) = (\beta\alpha).$$

One finds in the same way,

$$(\alpha\gamma) = (\gamma\alpha) = \sum \alpha\gamma, \quad (\beta\gamma) = (\gamma\beta) = \sum \beta\gamma, \quad \ldots$$

Let us denote by $\lambda, \lambda', \lambda''$, etc. the values which the functions $\nu, \nu', \nu''$, etc. take when one replaces $x, y, z$, etc. by their most likely values, $A, B, C$, etc., that is to say, let us put

$$aA + bB + cC + \cdots + l = \lambda,$$
$$a'A + b'B + c'C + \cdots + l' = \lambda',$$
$$a''A + b''B + c''C + \cdots + l'' = \lambda'',$$
$$\vdots$$

If we take

$$\lambda^2 + \lambda'^2 + \lambda''^2 + \cdots = M,$$

so that $M$ is the value of the function $\Omega$ corresponding to the most likely values of the variables, $M$ will be (section 20) the minimum value of $\Omega$. Consequently,

$$a\lambda + a'\lambda' + a''\lambda'' + \cdots$$

will be the value of $\xi$ when

$$x + A, \quad y = B, \quad z = C, \quad \ldots$$

This value is zero, because of the way in which $A, B, C$, etc. were obtained. One will thus have

$$\sum a\lambda = 0;$$

one obtains in the same way

$$\sum b\lambda = 0, \quad \sum c\lambda = 0, \quad \ldots,$$
and
\[ \sum \alpha \lambda = 0, \quad \sum \beta \lambda = 0, \quad \sum \gamma \lambda = 0, \ldots \]

Finally, multiplying the values of \( \lambda, \lambda', \lambda'', \) etc. by \( \lambda, \lambda', \lambda'', \) etc., and adding, we obtain
\[ l\lambda + l'\lambda' + l''\lambda'' + \cdots = \lambda^2 + \lambda'^2 + \lambda''^2 + \cdots, \]
that is to say
\[ \sum l\lambda = M. \]

Replacing \( x, y, z, \) in the equation
\[ \nu = ax + by + cz + \cdots + l, \]
by the expressions (7) [section 21], we obtain after easy reductions
\[
\begin{align*}
\nu &= \alpha \xi + \beta \eta + \gamma \zeta + \cdots + \lambda, \\
\nu' &= \alpha' \xi' + \beta' \eta' + \gamma' \zeta' + \cdots + \lambda', \\
\nu'' &= \alpha'' \xi'' + \beta'' \eta'' + \gamma'' \zeta'' + \cdots + \lambda'', \\
\vdots
\end{align*}
\]
Multiplying either these equations, or the equations (1) of section 20, by \( \lambda, \lambda', \lambda'', \) etc., respectively, and finally adding, one obtains the identity
\[ \lambda \nu + \lambda' \nu' + \lambda'' \nu'' + \cdots = M, \]

The function \( \Omega \) can be expressed in several forms which it is important to indicate.

Squaring the equations (1) [section 20], and adding them term by term, we find
\[
\begin{align*}
\Omega &= x^2 \sum a^2 + y^2 \sum b^2 + z^2 \sum c^2 + \cdots + 2xy \sum ab + 2xz \sum ac + \\
&\quad 2yz \sum bc + \cdots + 2x \sum al + 2y \sum bl + 2z \sum cl + \cdots + \sum l^2.
\end{align*}
\]
this is the first form.

Multiplying the same equations by \( \nu, \nu', \nu'', \) etc., respectively, and adding, one has
\[ \Omega = \xi x + \eta y + \zeta z + \cdots + l\nu + l'\nu' + l''\nu'' + \cdots; \]
replacing \( \nu, \nu', \nu'', \) etc. by the values indicated in the preceding section we find
\[ \Omega = \xi x + \eta y + \zeta z + \cdots - A\xi - B\eta - C\zeta - \cdots + M, \]
or
\[ \Omega = \xi(x - A) + \eta(y - B) + \zeta(z - C) + \cdots + M; \]
this is the second form.

Finally, replacing \( x - A, y - B, z - C, \) etc., in this second form, by the expressions

\[ (7) \] [section 21], we obtain the third form:

\[ \Omega = (\alpha\alpha)\xi^2 + (\beta\beta)\eta^2 + (\gamma\gamma)\zeta^2 + \cdots + 2(\alpha\beta)\xi\eta + 2(\alpha\gamma)\xi\zeta + 2(\beta\gamma)\eta\zeta + \cdots + M. \]

One can give a fourth form which is an immediate consequence of the third and of the formulas of the preceding sections

\[ \Omega = (\nu - \lambda)^2 + (\nu' - \lambda')^2 + (\nu'' - \lambda'')^2 + \cdots + M, \]

that is to say

\[ \Omega = M + \sum (\nu - \lambda)^2. \]

In this last form one sees clearly that \( M \) is the minimum value of \( \Omega \).

28.

Let \( e, e', e'' \), etc. be the errors committed in observations which have given

\[ V = L, \quad V' = L', \quad V'' = L'', \quad \ldots. \]

The true values of the functions, \( V, V', V'' \), etc. will then be

\[ L - e, \quad L' - e', \quad L'' - e'', \quad \ldots, \]

and the true values of \( \nu, \nu', \nu'' \), etc. will be

\[ -e\sqrt{p}, \quad -e'\sqrt{p'}, \quad -e''\sqrt{p''}, \quad \ldots, \]

respectively; consequently, the true value of \( x \) will be

\[ A - \alpha e\sqrt{p} - \alpha'e'\sqrt{p'} - \alpha''e''\sqrt{p''} - \cdots, \]

and the error committed in the most suitable determination of the unknown \( x \) will be, denoting it by \( E_x \),

\[ E_x = \alpha e\sqrt{p} + \alpha'e'\sqrt{p'} + \alpha''e''\sqrt{p''} + \cdots. \]

Similarly, the error committed in the most suitable determination of the value of \( y \) will be

\[ E_y = \beta e\sqrt{p} + \beta'e'\sqrt{p'} + \beta''e''\sqrt{p''} + \cdots. \]

The average value of the square \((E_x)^2\) will be

\[ m^2p(\alpha^2 + \alpha'^2 + \alpha''^2 + \cdots) = m^2p(\alpha\alpha). \]

Similarly the average value of \((E_y)^2\) will be

\[ m^2p(\beta\beta). \]
as we have already seen. One can also give the average value of the product $E_x E_y$, which will be

$$m^2 p(\alpha \beta + \alpha' \beta' + \alpha'' \beta'' + \cdots) = m^2 p(\alpha \beta).$$

One can state these results more briefly in the following way.

The average values of the squares $(E_x)^2$, $(E_y)^2$, etc. are equal to $\frac{m^2 p}{2}$ times the partial derivatives of the second order

$$\frac{d^2 \Omega}{d\xi^2}, \quad \frac{d^2 \Omega}{d\eta^2}, \quad \cdots,$$

and the average value of a product such as $E_x E_y$ is equal to $\frac{1}{2} m^2 p$ times $\frac{d^2 \Omega}{d\xi d\eta}$, considering $\Omega$ as the function of $\xi, \eta, \zeta$, etc.

29.

Let $t$ be a given linear function of the quantities $x, y, z$, etc., for example,

$$t = fx + gy + hz + \cdots + k;$$

the value of $t$ obtained from the most likely values of $x, y, z$, etc. will be

$$fA + gB + hC + \cdots + k;$$

we shall denote it by $K$. Denoting by $E_t$ the error committed in taking $K$ as the value of $t$, one has

$$E_t = f(E_x) + g(E_y) + h(E_z) + \cdots;$$

the average value of this error will obviously be zero, that is to say, the error will contain no constant part, but the average value of $(E_t)^2$, that is to say of the sum

$$f^2(E_x)^2 + 2fg(E_x)(E_y) + 2fh(E_x)(E_z) + \cdots +
+g^2(E_y)^2 + 2gh(E_y)(E_z) + \cdots +
+h^2(E_z)^2 + \cdots,$$

will be, according to the preceding section, equal to $m^2 p$ times the sum

$$f^2(\alpha \alpha) + 2fg(\alpha \beta) + 2fh(\alpha \gamma) + \cdots +
+g^2(\beta \beta) + 2gh(\beta \gamma) + \cdots +
+h^2(\gamma \gamma) + \cdots,$$

that is to say equal to the product of $m^2 p$ by the value of the function $\Omega - M$ when one sets

$$\xi = f, \quad \eta = g, \quad \zeta = h, \quad \cdots,$$

in it.

Let us denote by $\omega$ this value of $\Omega - M$; the standard error to be expected, when one takes $t = K$, will be $m\sqrt{p\omega}$, and the weight of this determination will be $\frac{1}{\omega}$. 

Since one has identically
\[ \Omega - M = (x - A)\xi + (y - B)\eta + (z - C)\zeta + \cdots, \]
\( \omega \) will be equal to the value of the expression
\[ (x - A)f + (y - B)g + (z - C)h + \cdots \]
[which gives \((t - K)\)], when one replaces \(x, y, z, \) etc. by the values corresponding to \(\xi = f, \eta = g, \zeta = h, \) etc.

Finally, noting that \(t\) expressed as a function of the quantities \(\xi, \eta, \zeta, \) etc. will have \(K\) as constant part, if one supposes
\[ t = F\xi + G\eta + H\zeta + \cdots + K, \]
one will have
\[ \omega + fF + gG + hH + \cdots. \]

30.

We have seen that the function \(\Omega\) takes on its absolute minimum \(M\) when one puts
\[ x = A, \quad y = B, \quad z = C, \quad \ldots, \]
or, what amounts to the same thing,
\[ \xi = 0, \quad \eta = 0, \quad \zeta = 0, \quad \ldots \]
If one assigns another value to one of the unknowns, for instance, if one puts
\[ x = A + \Delta, \]
the other unknowns still being considered as variables, \(\Omega\) can take on a relative minimum which can be located with the help of the equations
\[ x = A + \Delta, \quad \frac{d\Omega}{dy} = 0, \quad \frac{d\Omega}{dz} = 0, \quad \ldots, \]
and consequently
\[ \eta = 0, \quad \zeta = 0, \quad \ldots; \]
now since
\[ x = A + (\alpha\alpha)\xi + (\alpha\beta)\eta + (\alpha\gamma)\zeta + \cdots, \]
one concludes from this
\[ \xi = \frac{\Delta}{(\alpha\alpha)}. \]
In the same way one finds
\[ y = B + \frac{(\alpha\beta)}{(\alpha\alpha)}\Delta, \quad z = C + \frac{(\alpha\gamma)}{(\alpha\alpha)}\Delta, \quad \ldots \]
The relative minimum of $\Omega$ will be

$$(\alpha \alpha) \xi^2 + M = M + \frac{\Delta^2}{(\alpha \alpha)}$$

Conversely we conclude from this that if $\Omega$ is not to exceed $M + \mu^2$ the value of $x$ is necessarily included between the limits $A - \mu \sqrt{(\alpha \alpha)}$ and $A + \sqrt{(\alpha \alpha)}$. It is important to note that $\mu \sqrt{(\alpha \alpha)}$ becomes equal to the standard error to be expected in the most likely value of $x$, if one puts

$$\mu = m \sqrt{p};$$

that is to say if $\mu$ is the standard error of observations of unit weight.

More generally, let us try to find the smallest value of the function $\Omega$ which can correspond to a given value of $t$, where, as in the preceding section, $t$ denotes the linear expression

$$fx + gy + hz + \cdots + k,$$

whose most likely value is $K$; let us denote the given value of $t$ by $K + \kappa$. According to the theory of maxima the solution of the problem will be given by the equations

$$\frac{d\Omega}{dx} = \Theta \frac{dt}{dx},$$

$$\frac{d\Omega}{dy} = \Theta \frac{dt}{dy},$$

$$\frac{d\Omega}{dz} = \Theta \frac{dt}{dz},$$

$$\cdots,$$

or

$$\xi = \Theta f,$$

$$\eta = \Theta g,$$

$$\zeta = \Theta h,$$

$$\cdots,$$

where $\Theta$ denotes an undetermined multiplier.

If, as in the preceding section, we put

$$t = F\xi + G\eta + H\zeta + \cdots + K,$$

we shall have

$$K + \kappa = \Theta(f F + gG + hH + \cdots) + K;$$

this gives

$$\Theta = \frac{\kappa}{\omega},$$

where $\omega$ has the same significance as in the preceding section.
Since $\Omega - M$ is a homogeneous function of the second degree with respect to $\xi, \eta, \zeta$, etc., its value for
\[
\xi = \Theta f, \quad \eta = \Theta g, \quad \zeta = \Theta h,
\]
is obviously
\[
\Theta^2 \omega
\]
and consequently, the minimum value of $\Omega$ when
\[
t = K + \kappa,
\]
is
\[
M + \Theta^2 \omega = M + \frac{\kappa^2}{\omega}
\]
Conversely, if $\Omega$ is to stay below a given value $M + \mu^2$, the value of $t$ will necessarily be included between the limits $K - \mu \sqrt{\omega}, K + \mu \sqrt{\omega}$, and $\mu \sqrt{\omega}$ will be the standard error to be expected in the most likely value of $t$, if $\mu$ represents the standard error of observations of unit weight.

31.

When the number of unknowns $x, y, z$, etc., is fairly large, the determination of the numerical values of $A, B, C$, etc. by the ordinary elimination process is a considerable task. It is for this reason that we indicated in *Theoria Motus Corporum Coelestium*, and developed later in a *Memoir on the elements of the orbit of Pallas*, a method which simplifies this work as much as possible.

The function $\Omega$ should be brought to the following form:
\[
\frac{u^0}{A^0} + \frac{u^1}{B^1} + \frac{u''}{C''} + \frac{u'''}{D'''} + \cdots + M,
\]
where the denominators $A^0, B^1, C'', D'''$, etc. are constants; $u^0, u', u'', u''', \ldots$, etc. are linear functions of $x, y, z$, etc., such that the second $u'$ does not contain $x$, the third $u''$ contains neither $x$ nor $y$, the fourth contains neither $x, y$ nor $z$ and so on, so that the last, $u^{(\pi-1)}$, contains only the last of the unknowns $x, y, z$, etc.; finally the coefficients of $x, y, z$, etc. in $u^0, u', u'', u''', \ldots$, etc. are respectively equal to $A^0, B^1, C'', D'''$, etc. Then one puts
\[
u^0 = 0, \quad u' = 0, \quad u'' = 0, \quad u''' = 0, \quad \ldots,
\]
and one obtains very easily the values of $x, y, z$, etc. by solving these equations, beginning with the last. I do not think it necessary to repeat again the algorithm which leads to this transformation of the function $\Omega$.

However, the elimination which it is necessary to carry out to find the weights of these determinations requires calculations which are much longer still. We have shown, in *Theoria Motus Corporum Coelestium*, that the weight of the determination of the last unknown, which is the only one appearing in $u^{(\pi-1)}$, is equal to the last term of the sequence of denominators $A^0, B^1, C''$, etc. This derivation is easy; and

---

6One will find these calculations in a Note at the end of the volume. —J.B.
several calculators, wishing to avoid a tedious elimination process, have had the idea, for lack of another method, of repeating the indicated transformation considering each unknown in succession as the last. I hope then, that mathematicians will be grateful to me for indicating a new method for calculating the weights of the determinations, which seem to me to leave nothing to be desired on this point.

32.

Let us put

$$\begin{align*}
  u^0 &= A^0 x + B^0 y + C^0 z + \cdots + L^0, \\
  u' &= B'y + C''z + \cdots + L', \\
  u'' &= C''z + \cdots + L'', \\
  \vdots & \quad \vdots
\end{align*}$$

(1)

We then have the identity

$$\frac{1}{2} d\Omega = \xi dx + \eta dy + \zeta dz + \cdots$$

$$= u^0 \frac{du^0}{A^0} + u' \frac{du'}{B^0} + u'' \frac{du''}{C^0} + \cdots$$

$$= u^0 \left( dx + \frac{B^0}{A^0} dy + \frac{C^0}{A^0} dz + \cdots \right) + u' \left( dy + \frac{C'}{B'} dz + \cdots \right)$$

$$+ u'' (dz + \cdots ) + \cdots ;$$

from which we deduce

$$\begin{align*}
  \xi &= u^0, \\
  \eta &= \frac{B^0}{A^0} u^0 + u', \\
  \zeta &= \frac{C^0}{A^0} u^0 + \frac{C'}{B'} u' + u'', \\
  \vdots & \quad \vdots
\end{align*}$$

(2)

The values of $u^0, u', u''$, etc., obtained from these equations, will appear in the following form:

$$\begin{align*}
  u^0 &= \xi, \\
  u' &= A' \xi + \eta, \\
  u'' &= A'' \xi + B'' \eta + \zeta, \\
  \vdots & \quad \vdots
\end{align*}$$

(3)

Taking the complete differential of the equation

$$\Omega = \xi(x - A) + \eta(y - B) + \zeta(z - C) + \cdots + M,$$
we derive the equation
\[ \frac{1}{2} d\Omega = \xi dx + \eta dy + \zeta dz + \cdots , \]
which becomes
\[ \frac{1}{2} d\Omega = (x - A) d\xi + (y - B) d\eta + (z - C) d\zeta + \cdots . \]
This expression should be identical with that obtained from the equations (3), that is to say
\[ \frac{u^0}{A^0} d\xi + \frac{u'}{B'} (A' d\xi + d\eta) + \frac{u''}{C''} (A'' d\xi + B'' d\eta + d\zeta) + \cdots ; \]
hence one has
\[
\begin{align*}
    x &= \frac{u^0}{A^0} + A' \frac{u'}{B'} + A'' \frac{u''}{C''} + \cdots + A, \\
    y &= \frac{u'}{B'} + B'' \frac{u''}{C''} + \cdots + B, \\
    z &= \frac{u''}{C''} + \cdots + C, \\
    &\vdots \\
\end{align*}
\]
On substituting into these expressions, the values of \( u^0, u', u'', \) etc., taken from the equations (3), the elimination is effected. In order to determine the weights, we have
\[
\begin{align*}
    (\alpha\alpha) &= \frac{1}{A^0} + \frac{A'^2}{B^2} + \frac{A''^2}{C'^2} + \frac{A'''^2}{D'^2} + \cdots , \\
    (\beta\beta) &= \frac{1}{B^0} + \frac{B'^2}{C'^2} + \frac{B''^2}{D'^2} + \cdots , \\
    (\gamma\gamma) &= \frac{1}{C''^2} + \frac{C'''^2}{D''^2} + \cdots , \\
    &\vdots \\
\end{align*}
\]
The simplicity of these formulas leaves nothing to be desired. One may find equally simple formulas to express the other coefficients \((\alpha\beta), (\alpha\gamma), (\beta\gamma), \) etc.; but, since their use is less frequent, we shall not bother to set them out.

33.

The importance of the subject has induced us to reduce everything to a form suitable for calculation and to derive the explicit expressions for the coefficients \( A', A'', A''', B'', B'''', \) etc.

This calculation can be approached in two ways: the first consists in substituting back into the equations (2), the values of \( u^0, u', u'', \) etc. obtained from the system (3), which ought to give identities; and the second on the other hand, by expressing the
identities obtained from the system (2) when one substitutes in it the values of \( \xi, \eta, \zeta \), obtained from the system (3).

The first method leads to the following formulas:

\[
\begin{align*}
\frac{\mathcal{B}^0}{\mathcal{A}^0} + A' &= 0 \\
\frac{\mathcal{C}^0}{\mathcal{A}^0} + \frac{\mathcal{C}'}{\mathcal{B}'} A' + A'' &= 0 \\
\frac{\mathcal{D}^0}{\mathcal{A}^0} + \frac{\mathcal{D}'}{\mathcal{B}'} A' + \frac{\mathcal{D}''}{\mathcal{C}''} A'' + A''' &= 0 \\
&\vdots
\end{align*}
\]

From these formulas we derive \( A', A'', A''', \) etc. One then has

\[
\begin{align*}
\frac{\mathcal{C}'}{\mathcal{B}'} + B'' &= 0, \\
\frac{\mathcal{D}'}{\mathcal{B}'} + \frac{\mathcal{D}''}{\mathcal{C}''} B'' + B''' &= 0, \\
&\vdots
\end{align*}
\]

which give \( B'', B''', \) etc.; then

\[
\frac{\mathcal{D}''}{\mathcal{C}''} + C''' = 0, \\
&\vdots
\]

which gives \( C''', \) etc.; and so on.

The second method gives the following system:

\[
A^0 A' + B^0 = 0,
\]

from which one derives \( A' \);

\[
A^0 A'' + B^0 B'' + C^0 = 0, \\
B^0 B'' + C' = 0,
\]

from which one derives \( B'' \) and \( A'' \):

\[
A^0 A''' + B^0 B''' + C^0 C''' + D^0 = 0, \\
B^0 B''' + C' C''' + D' = 0, \\
C'' C''' + D'' = 0,
\]

from which one derives \( C''', B''', A''' \); and so on.

There is very little to choose between the two systems of formulas when one wants the weights of the determinations of all the unknowns \( x, y, z, \) etc.; but when one wants
to find only one of the quantities \((\alpha\alpha), (\beta\beta), (\gamma\gamma),\) etc., the first system is considerably better.

The combination of equations (1) and (4) leads to the same formulas as well, and furnishes besides a second method for obtaining the most likely values \(A, B,\) etc., which are

\[
A = -\frac{L^0}{B^0} - A'\frac{L'}{B'} - A''\frac{L''}{C''} - A'''\frac{L'''}{D'''} - \cdots,
\]

\[
B = -\frac{L'}{B'} - B''\frac{L''}{C''} - B'''\frac{L'''}{D'''} - \cdots,
\]

\[
C = -\frac{L''}{C''} - C'''\frac{L'''}{D'''} - \cdots,
\]

\[\vdots \]

The other calculation is identical with the ordinary calculation in which one supposes

\[
u^0 = 0, \quad u' = 0, \quad u'' = 0, \quad \ldots
\]

34.

The results obtained in section 32 are only particular cases of a more general theorem which can be stated in the following way:

**THEOREM:** If \(t\) represents the following linear function of the unknowns \(x, y, z,\) etc.,

\[t = fx + gy + hz + \cdots + k,
\]

whose expression as a function of the variables \(u^0, u', u'',\) etc. is

\[t = k^0 u^0 + k' u' + k'' u'' + \cdots + K,
\]

\(K\) will be the most likely value of \(t\) and the weight of this determination will be

\[
\frac{1}{A^0 k^0 + B' k'^2 + C'' k''^2 + \cdots}
\]

**PROOF:** The first part of the theorem is obvious, since the most likely value of \(t\) must correspond to the values

\[
u^0 = 0, \quad u' = 0, \quad u'' = 0, \quad \ldots
\]

To prove the second part, let us note that one has

\[
\frac{1}{2} d\Omega = \xi dx + \eta dy + \zeta dz + \cdots,
\]

\[dt = f dx + g dy + h dz + \cdots,
\]

and consequently, when

\[
\xi = f, \quad \eta = g, \quad \zeta = h, \quad \ldots
\]
one has

\[ d\Omega = 2dt \]

whatever the differentials \( dx, dy, dz, \) etc. It follows that if we continue to suppose

\[ \xi = f, \quad \eta = g, \quad \zeta = h, \quad \ldots, \]

we shall have

\[ \frac{u^0}{A^0} du^0 + \frac{u'}{B'} du' + \frac{u''}{C''} du'' + \cdots = k^0 du^0 + k' du' + k'' du'' + \cdots \]

Now one sees easily that if the differentials \( dx, dy, dz, \) etc. are independent of each other, it will be the same for \( du^0, du', du'', \) etc.; and consequently for

\[ \xi = f, \quad \eta = g, \quad \zeta = h, \quad \ldots, \]

we shall have

\[ u^0 = A^0 k^0, \quad u' = B' k', \quad u'' = C'' k'', \quad \ldots \]

Consequently, the value of \( \Omega \) corresponding to the same hypothesis will be

\[ A^0 k^{02} + B' k'^2 + C'' k''^2 + \cdots + M; \]

which, according to section 29, proves our theorem.

However, if one wishes to carry out the transformation of the function \( t \), without making use of the formulas (4) (section 32), one has at once the relations

\[ f = A^0 k^0, \quad g = B^0 k^0 + B' k', \quad h = C^0 k^0 + C' k' + C'' k'', \quad \ldots \]

which allow \( k^0, k', k'', \) etc. to be determined, and finally we have

\[ K = -L^0 k^0 - L' k' - L'' k'' - \cdots. \]

We shall discuss the following problem in detail, as much on account of its practical utility as of the simplicity of its solution.

To find the changes which the most likely values of the unknowns undergo when a new equation is adjoined, and to determine the weights of these new determinations.

Let us keep the notations used above. The original equations, reduced so as to have unit weight, will be

\[ \nu = 0, \quad \nu' = 0, \quad \nu'' = 0, \quad \ldots; \]

and one will have

\[ \Omega = \nu^2 + \nu'^2 + \nu''^2 + \cdots, \]
\(\xi, \eta, \zeta, \text{ etc., will be the partial derivatives}\)
\[
\frac{d\Omega}{2dx}, \frac{d\Omega}{2dy}, \frac{d\Omega}{2dz}, \ldots
\]
and finally one will have, by elimination,
\[
\begin{aligned}
x &= A + (\alpha\alpha)\xi + (\alpha\beta)\eta + (\alpha\gamma)\zeta + \cdots, \\
y &= B + (\alpha\beta)\xi + (\beta\beta)\eta + (\beta\gamma)\zeta + \cdots, \\
z &= C + (\alpha\gamma)\xi + (\beta\gamma)\eta + (\gamma\gamma)\zeta + \cdots, \\
\end{aligned}
\]
(1)

Now let us suppose that we have a new approximate equation
\[
n^* = 0,
\]
whose weight we shall suppose equal to unity. Let us find the changes in the most likely values \(A, B, C, \text{ etc.}\) in the coefficients \((\alpha\alpha), (\beta\beta), \text{ etc.}\).

Let us put
\[
\Omega + \nu^* = \Omega^*,
\]
\[
\frac{1}{2} \frac{d\Omega^*}{dx} = \xi^*, \quad \frac{1}{2} \frac{d\Omega^*}{dy} = \eta^*, \quad \frac{1}{2} \frac{d\Omega^*}{dz} = \zeta^*, \quad \ldots
\]
and let
\[
x = A^* + (\alpha\alpha^*)\xi^* + (\alpha\beta^*)\eta^* + (\alpha\gamma^*)\zeta^* + \cdots
\]
be the result of the elimination.

Finally, let
\[
n^* = fx + gy + hz + \cdots + k,
\]
which becomes, taking note of the equations (1),
\[
n^* = F\xi + G\eta + H\zeta + \cdots + K,
\]
and set
\[
Ff + Gg + Hh + \cdots = \omega;
\]
\(K\) is then obviously the most likely value of the function \(n^*\), as it is deduced from the original equations without taking into account the values \(O\) furnished by the new observations, and \(\frac{1}{\omega}\) is the weight of this determination.

Now we have
\[
\xi^* = \xi + f\nu^*, \quad \eta^* = \eta + g\nu^*, \quad \zeta^* = \zeta + h\nu^*, \quad \ldots
\]
and consequently
\[
F\xi^* + G\eta^* + H\zeta^* + \cdots + K = \nu^*(1 + Ff + Gg + Hh + \cdots);
\]
from which we deduce
\[ \nu^* = \frac{F\xi^* + G\eta^* + H\zeta^* + \cdots + K}{1 + \omega} \]

One has, besides,
\[
x = A + (\alpha\alpha)\xi^* + (\alpha\beta)\eta^* + (\alpha\gamma)\zeta^* + \cdots - \nu^*[f(\alpha\alpha) + g(\alpha\beta) + h(\alpha\gamma) + \cdots] = A + (\alpha\alpha)\xi^* + (\alpha\beta)\eta^* + \cdots - F\nu^* \]
\[= A + (\alpha\alpha)\xi^* + (\alpha\beta)\eta^* + \cdots - \frac{F}{1 + \omega} (F\xi^* + G\eta^* + H\zeta^* + \cdots + K). \]

We deduce from this
\[A^* = A - \frac{FK}{1 + \omega}\]
which will be the most likely value of \(x\), obtained from all the observations.

One has also
\[(\alpha\alpha^*) = (\alpha\alpha) - \frac{F^2}{1 + \omega};\]
and consequently,
\[
\frac{1}{(\alpha\alpha) - \frac{F^2}{1 + \omega}}
\]
will be the weight of this determination.

Similarly, one finds that the most likely value of \(y\) obtained from all the observations is
\[B^* = B - \frac{GK}{1 + \omega};\]
and the weight of this determination will be
\[
\frac{1}{(\beta\beta) - \frac{G^2}{1 + \omega}};
\]
and so on.

Thus the problem is solved.

Let us add a few remarks.

I. On substituting the new values \(A^*, B^*, C^*, \) etc. the function \(\nu^*\) takes on its most likely value
\[K - \frac{K}{1 + \omega} (Ff + Gg + Hh + \cdots) = \frac{K}{1 + \omega}, \]
and since one has identically
\[\nu^* = \frac{F}{1 + \omega} \xi^* + \frac{G}{1 + \omega} \eta^* + \frac{H}{1 + \omega} \zeta^* + \cdots + \frac{K}{1 + \omega}, \]
the weight of this determination will be (section 29)
\[
\frac{1 + \omega}{Ff + Gg + Hh + \cdots} = \frac{1}{\omega} + 1.
\]
These results could be deduced immediately from the rules set forth at the end of section 21. The original set of equations gave, in fact, the determination \( \nu^* = K \), with weight \( \frac{1}{\omega} \), and the new observation gave another determination \( \nu^* = 0 \), independent of the first, with unit weight; their combination gives the determination

\[
\nu^* = \frac{K}{1 + \omega},
\]

which has weight

\[
\frac{1}{\omega} + 1.
\]

II. One concludes from the foregoing that for

\[
x = A^*, \quad y = B^*, \quad z = C^*, \ldots,
\]

one should have

\[
\xi^* = 0, \quad \eta^* = 0, \quad \zeta^* = 0, \ldots,
\]

and consequently

\[
\xi = -\frac{fK}{1 + \omega}, \quad \eta = -\frac{gK}{1 + \omega}, \quad \zeta = -\frac{hK}{1 + \omega}, \ldots
\]

Since, however,

\[
\Omega = \xi(x - A) + \eta(y - B) + \zeta(z - C) + \cdots + M,
\]

\[
\Omega^* = \Omega + \nu^* \omega^2
\]

one must have, for these same values,

\[
\Omega = \frac{K^2}{(1 + \omega)^2} (Ff + Gg + Hh + \cdots) + M + \frac{\omega K^2}{(1 + \omega)^2}
\]

and

\[
\Omega^* = M + \frac{\omega K^2}{(1 + \omega)^2} + \frac{K^2}{(1 + \omega)^2} = M + \frac{K^2}{1 + \omega}.
\]

III. Comparing these results with those of section 30, we see here that the function \( \Omega \) has the smallest value which it can take when one imposes the condition

\[
\nu^* = \frac{K}{1 + \omega}.
\]

We shall give here only the solution of the following problem, which is very similar to the preceding; but we shall refrain from giving the proof which the reader will readily supply with the help of what has preceded.

To find the changes in the most likely values of the unknowns and the weights of the new determinations, when the weight of one of the original observations is changed.
Let us suppose that after having finished the calculation one notices that one has given to some observation, for instance, the first, which gave \( V = L \), too large or too small a weight, and that it would be more accurate to assign it the weight \( p^* \) in place of the weight \( p \); it is unnecessary to begin the calculation over again, and it is more convenient to obtain the necessary corrections with the help of the following formulas.

The most likely values of the unknowns will become

\[
\begin{align*}
x &= A - \frac{(p^* - p)\alpha \lambda}{p + (p^* - p)(a\alpha + b\beta + c\gamma + \cdots)}, \\
y &= B - \frac{(p^* - p)\beta \lambda}{p + (p^* - p)(a\alpha + b\beta + c\gamma + \cdots)}, \\
z &= C - \frac{(p^* - p)\gamma \lambda}{p + (p^* - p)(a\alpha + b\beta + c\gamma + \cdots)}, \\
\vdots
\end{align*}
\]

and the respective weights of these determinations will be obtained by taking the reciprocals of

\[
\begin{align*}
(\alpha \alpha) &= \frac{(p^* - p)\alpha^2}{p + (p^* - p)(a\alpha + b\beta + c\gamma + \cdots)}, \\
(\beta \beta) &= \frac{(p^* - p)\beta^2}{p + (p^* - p)(a\alpha + b\beta + c\gamma + \cdots)}, \\
(\gamma \gamma) &= \frac{(p^* - p)\gamma^2}{p + (p^* - p)(a\alpha + b\beta + c\gamma + \cdots)}, \\
\vdots
\end{align*}
\]

This solution is suitable for the case in which, after having finished the calculations it is necessary to reject one of the observations completely, since this amounts to making \( p^* = 0 \); similarly, \( p^* = \infty \) leads to the case in which the equation \( V = L \) which in the calculation had been considered approximate becomes rigorously exact.

If, after the calculation is finished, several new equations should be adjoined to the original, or if the weights attributed to several of them were in error, the calculations of the corrections would become very complicated and it would be better to begin all over again.

In sections 15 and 16 we gave an approximate method of determining the precision of a system of observations\(^7\) but this method supposes that the real errors one has actually encountered in a long sequence of observations are known exactly. Now this condition is satisfied only very rarely, not to say never.

\(^{7}\) Our previous research on the same subject (Zeitschrift für Astronomie und verwandte Wissenschaften, Vol. 1, page 185) were based on the hypotheses concerning the probability of errors to which we were led to in Theoria Motus Corporum Coelestium.

One will find this Memoir at the end of the volume. —J.B.
If the quantities whose approximate values are given by observation depend on one or several unknowns, following a given probability distribution, one can find by the method of least squares the most probable values of these unknowns; if one then calculates the corresponding values of the observed quantities, these latter may be considered as close to the true values: so that the differences between them and the observed values will represent the errors committed with greater and greater certainty as the number of observations increases. This is the procedure followed in practice by calculators who have attempted in complicated cases to evaluate the precision of the observations a posteriori. Although this method is sufficient in many cases, it is theoretically incorrect and may sometimes lead to serious errors; for this reason it is very important to treat the question with more care.

Let us keep the notations of section 19. The method under consideration consists in regarding \( A, B, C, \) etc. as the true values of the unknowns \( x, y, z, \) etc. and \( \lambda, \lambda', \lambda'', \) etc. as the true values of the functions \( \nu, \nu', \nu'', \) etc. If all the observations are of equal precision and their common weight \( p = p' = p'' = \cdots \) is taken as unity, these same quantities, taken with the opposite sign, on this assumption represent the errors of the observations, and consequently, according to section 15,

\[
x = \sqrt{\frac{\lambda^2 + \lambda'^2 + \lambda''^2 + \cdots}{\omega}} = \sqrt{\frac{M}{\omega}}
\]

will be the standard error of the observations. If the observations are not of the same precision, \(-\lambda, -\lambda', -\lambda'', \) etc. will represent the errors of the observations multiplied by the square roots of the respective weights, and the rules of section 16 will lead to the same formula

\[
\sqrt{\frac{M}{\omega}},
\]

which previously expressed the standard error of these observations when their weight is equal to unity.

But the exact calculation would clearly require replacing \( \lambda, \lambda', \lambda'', \) etc. by the values of \( \nu, \nu', \nu'', \) etc., derived from the true values of the unknowns \( x, y, z, \) etc. and the quantity \( M \) by the corresponding value \( \Omega. \) Although one cannot determine this last value, we are nevertheless certain that it is larger than \( M, \) which is its minimum. It will not attain this minimum except in the infinitesimally probable case in which the most likely values of the unknowns coincide with the true values. We can thus assert that in general the standard error calculated by the ordinary practice is smaller than the precise standard error, and that consequently an excessive precision is attributed to the observations.

Let us see what a rigorous theory gives.

38.

To begin with, it is necessary to find out how the quantity \( M \) depends on the actual errors of the observations. Let us denote these errors as in section 28, by \( e, e', e'', \) etc.
and let us put for simplicity
\[ e \sqrt{p} = \epsilon, \quad e' \sqrt{p'} = \epsilon', \quad e'' \sqrt{p''} = \epsilon'' \]
and
\[ m \sqrt{p} = m' \sqrt{p'} = m'' \sqrt{p''} = \cdots = \mu \]

Let
\[ A - x^0, \quad B - y^0, \quad C - z^0, \quad \ldots, \]
be the true values of the unknowns \( x, y, z, \ldots \), for which \( \xi, \eta, \zeta, \ldots \) are \( -\xi^0, -\eta^0, -\zeta^0, \ldots \), respectively. The corresponding values of \( \nu, \nu', \nu'', \ldots \) will obviously be \( -\epsilon, -\epsilon', -\epsilon'', \ldots \); so that one has
\[
\begin{align*}
\xi^0 &= a\epsilon + a'\epsilon' + a''\epsilon'' + \cdots, \\
\eta^0 &= b\epsilon + b'\epsilon' + b''\epsilon'' + \cdots, \\
\zeta^0 &= c\epsilon + c'\epsilon' + c''\epsilon'' + \cdots, \\
\ddots & \\
x^0 &= \alpha\epsilon + \alpha'\epsilon' + \alpha''\epsilon'' + \cdots, \\
y^0 &= \beta\epsilon + \beta'\epsilon' + \beta''\epsilon'' + \cdots, \\
z^0 &= \gamma\epsilon + \gamma'\epsilon' + \gamma''\epsilon'' + \cdots, \\
\ddots & \\
\end{align*}
\]
and finally,
\[ \Omega = \epsilon^2 + \epsilon'^2 + \epsilon''^2 + \cdots \]
will be the value of the function \( \Omega \) corresponding to the true values of the variables \( x, y, z, \ldots \).

Since
\[ \Omega = M + (x - A)\xi + (y - B)\eta + (z - C)\zeta + \cdots, \]
is an identity, one has also
\[ M = \Omega^0 - x^0\xi^0 - y^0\eta^0 - z^0\zeta^0 - \cdots. \]

From this it is obvious that \( M \) is a homogeneous function of the second degree of the errors \( \epsilon, \epsilon', \epsilon'', \ldots \); this function may become smaller or greater for various values of the errors. Since we do not know these values, it is wise to look closely at the function \( M \) and first of all to calculate its average value according to the principles of probability. We shall obtain this average value by replacing the squares \( \epsilon^2, \epsilon'^2, \ldots \) by \( m^2, m'^2, \ldots \) and omitting the terms \( \epsilon\epsilon', \epsilon\epsilon'', \ldots \), whose average value is zero; or, what comes to the same thing, by replacing each square \( \epsilon^2, \epsilon'^2, \epsilon''^2, \ldots \) by \( \mu^2 \), and neglecting \( \epsilon\epsilon', \epsilon\epsilon'', \ldots \). According to this the term \( \Omega^0 \) will give \( \omega\mu^2 \); the term \( -x^0\xi^0 \) will give
\[ -(a\alpha + a'\alpha' + a''\alpha'' + \cdots)\mu^2 = -\mu^2; \]
and each of the other parts will also give \(-\mu^2\), so that the total average value will be
\[
(\varpi - \rho)\mu^2,
\]
where \(\varpi\) denotes the number of observations and \(\rho\) the number of unknowns. The true value of \(M\) will be influenced by chance and can be greater or less than this average value, but the difference will become smaller and smaller as the number of observations increases; so that
\[
\sqrt{\frac{M}{\varpi - \rho}}
\]
can be considered as an approximate value of \(\mu\): consequently, the value of \(\mu\) furnished by the erroneous method of which we spoke in the preceding section should be increased in a ratio of \(\sqrt{\varpi - \rho}\) to \(\sqrt{\varpi}\).

39.

In order to see more clearly to what extent it is permissible to consider the value of \(M\) given by the observations as equal to the exact value it is necessary to find out the standard error to be expected when one puts
\[
\mu^2 = \frac{M}{\varpi - \rho}
\]
This standard error is the square root of the average value of the quantity
\[
\left(\frac{\Omega^0 - x^0\xi^0 - y^0\eta^0 - z^0\zeta^0 - \cdots - (\varpi - \rho)\mu^2}{\varpi - \rho}\right)^2
\]
which we shall write as follows
\[
\left(\frac{\Omega^0 - x^0\xi^0 - y^0\eta^0 - z^0\zeta^0 - \cdots}{\varpi - \rho}\right)^2 - \frac{2\mu^2}{\varpi - \rho} [\Omega^0 - x^0\xi^0 - y^0\eta^0 - z^0\zeta^0 - \cdots - (\varpi - \rho)\mu^2] - \mu^4;
\]
and since the average value of the second term is obviously zero, the question reduces to finding the average value of the function
\[
\psi = (\Omega^0 - x^0\xi^0 - y^0\eta^0 - z^0\zeta^0 - \cdots)^2
\]
Let us denote this average by \(N\); the standard error we seek will be
\[
\sqrt{\frac{N}{(\varpi - \rho)^2} - \mu^4}
\]
If one expands the function \(\psi\), one sees that it is a homogeneous function of the errors \(e, e', e'', \ldots\), or what comes to the same thing \(e, e', e'', \ldots\); thus one finds the average value by:

1. replacing the fourth powers \(e^4, e'^4, e''^4, \ldots\) by their average values;
2. replacing the products $e^2e'^2$, $e^2e''^2$, etc. by their average values, that is to say by $m^2m'^2$, $m^2m''^2$, etc.;

3. neglecting the products such as $e^3e'$, $e^2e'e''$, etc.

We shall suppose (section 16) that the average values of $e^4$, $e'^4$, $e''^4$, etc. are proportional to $m^4$, $m'^4$, $m''^4$, etc., so that their ratios are $\nu^4/\mu^4$, $\nu^4$ denoting the average value of the fourth powers of the errors for the observations of unit weight.

The preceding rules can be restated in this way:

Replace each fourth power $\epsilon^4$, $\epsilon'^4$, $\epsilon''^4$, etc. by $\nu^4$; each product $\epsilon^2\epsilon'^2$, $\epsilon^2\epsilon''^2$, etc. by $\mu^4$, and neglect all the terms such as $\epsilon^3\epsilon'$, $\epsilon^2\epsilon'\epsilon''$, $\epsilon\epsilon'\epsilon''\epsilon'''$.

Once these principles are understood, one sees easily that:

I. The average value of $\Omega^0\xi^0$ is $\varpi\nu^4 + (\varpi^2 - \varpi)\mu^4$

II. The average value of the product $\epsilon^2x^0\xi^0$ is

$$a\alpha\nu^4 + (a'\alpha' + a''\alpha'' + \cdots)\mu^4 = a\alpha(\nu^4 - \mu^4) + \mu^4,$$

since $a\alpha + a'\alpha' + a''\alpha'' + \cdots = 1$.

Similarly, the average value of $\epsilon'^2x^0\xi^0$ is

$$a'\alpha'(\nu^4 - \mu^4) + \mu^4;$$

the average value of $\epsilon''^2x^0\xi^0$ is

$$a''\alpha''(\nu^4 - \mu^4) + \mu^4;$$

and so on.

Thus the average value of the product

$$(\epsilon^2 + \epsilon'^2 + \epsilon''^2 + \cdots)x^0\xi^0 \text{ or } \Omega^0x^0\xi^0$$

will be $\nu^4 - \mu^4 + \varpi\mu^4$.

The products $\Omega^0y^0\eta^0$ or $\Omega^0z^0\zeta^0$, etc. will have the same average value; thus the product

$$\Omega^0(x^0\xi^0 + y^0\eta^0 + z^0\zeta^0 + \cdots)$$

will have

$$\rho\nu^4 + \rho(\varpi - 1)\mu^4$$

as average value.

III. In order to abbreviate the expansions which will follow, we shall adopt the following notation. We shall give to the sign $\Sigma$ a more extended sense than we have done up to now, and shall use it to denote the sum of the similar, though not identical, terms
which arise from all the permutations of the observations. Thus in this notation we shall have
\[ x^0 = \sum \alpha \epsilon \]
\[ x^{02} = \sum \alpha^2 \epsilon^2 + \sum \alpha \epsilon \epsilon'. \]
Calculating the average value of \( x^{02} \) term by term, we shall have first, as average value of the product \( \alpha^2 \epsilon^2 \xi^0 \),
\[ a^2 \alpha^2 \nu^4 + \alpha^2(a''^2 + a''\ldots) \mu^4 = a^2 \alpha^2(\nu^4 - \mu^4) + \nu^4 \mu^4 \sum a^2. \]
Similarly, the average value of the product \( \alpha'^2 \epsilon'^2 \xi^0 \)
\[ a'^2 \alpha'^2(\nu^4 - \mu^4) + \alpha'^2 \mu^4 \sum a^2; \]
and so on.
Consequently, the average value of the product \( \xi^{02} \sum \alpha^2 \epsilon^2 \) will be
\[ (\nu^4 - \mu^4) \sum a^2 \alpha^2 + \mu^4 \sum a^2. \sum \alpha^2 \]
Now the average value of \( \alpha \epsilon \epsilon' \xi^{02} \) is \( 2\alpha \epsilon' a \mu^4 \). The average value of \( \alpha \epsilon' \epsilon'' \xi^{02} \) is \( 2\alpha \epsilon' a'' \mu^4 \), and so on. From this one deduces easily that the average value of the product \( \xi^{02} \sum \alpha \epsilon \epsilon' \) is
\[ 2\mu^4 \sum a^2 \alpha^2 = \mu^4 \left[ (\sum a) \alpha^2 - \sum a^2 \alpha^2 \right] = \mu^4 \left( 1 - \sum a^2 \alpha^2 \right). \]
With this established we have as average value of the product \( x^{02} \xi^{02} \)
\[ (\nu^4 - 3\mu^4) \sum a^2 \alpha^2 + 2\mu^4 + \mu^4 \sum a^2. \sum \alpha^2 \]
IV. In an analogous way, the average value of the product \( x^0 y^0 \xi^0 \eta^0 \) is found to be
\[ \nu^4 \sum ab\alpha\beta + \mu^4 \sum aab' \beta' + \mu^4 \sum ab\alpha' \beta' + \mu^4 \sum a\beta b' \alpha' \]
Now one has
\[ \sum aab' \beta' = \sum a\alpha. \sum b\beta - \sum aab\beta \]
\[ \sum ab\alpha' \beta' = \sum ab. \sum \alpha\beta - \sum ab\alpha \beta \]
\[ \sum a\beta b' \alpha' = \sum a\beta. \sum b\alpha - \sum a\beta b\alpha \]
\[ \sum a\alpha = 1, \quad \sum b\beta = 1, \quad \sum a\beta = 0, \quad \sum b\alpha = 0; \]
thus this average value will be
\[ (\nu^4 - 3\mu^4) \sum ab\alpha\beta + \mu^4 \left( 1 + \sum ab. \sum \alpha\beta \right). \]
V. Similar calculation would give that the average value of $x^0 z^0 \xi^0 \zeta^0$ is
\[
(\nu^4 - 3\mu^4) \sum a\alpha \gamma + \mu^4 \left(1 + \sum ac \sum \alpha \gamma \right);
\]
and so on. Adding these one obtains the average value of the product
\[
x^0 \xi^0 (x^0 \xi^0 + y^0 \eta^0 + z^0 \zeta^0 + \cdots);
\]
which is
\[
(\nu^4 - 3\mu^4) \sum [a\alpha (a\alpha + b\beta + c\gamma + \cdots)] + \\
(\rho + 1)\mu^4 + \mu^4 \left(\sum a^2 \sum \alpha^2 + \sum ab \sum \alpha \beta + \sum ac \sum \alpha \gamma + \cdots\right) \\
= (\nu^4 - 3\mu^4) \sum [a\alpha (a\alpha + b\beta + c\gamma + \cdots)] + (\rho + 2)\mu^4
\]
VI. One finds in the same way
\[
(\nu^4 - 3\mu^4) \sum [b\beta (a\alpha + b\beta + c\gamma + \cdots)] + (\rho + 2)\mu^4
\]
as average value of the product
\[
y^0 \eta^0 (x^0 \xi^0 + y^0 \eta^0 + z^0 \zeta^0 + \cdots),
\]
and
\[
(\nu^4 - 3\mu^4) \sum [c\gamma (a\alpha + b\beta + c\gamma + \cdots)] + (\rho + 2)\mu^4
\]
as the average value of the product
\[
z^0 \zeta^0 (x^0 \xi^0 + y^0 \eta^0 + z^0 \zeta^0 + \cdots);
\]
and so on.
Thus adding, we have the average value of the square
\[
(x^0 \xi^0 + y^0 \eta^0 + z^0 \zeta^0 + \cdots)^2;
\]
it will be
\[
(\nu^4 - 3\mu^4) \sum [(a\alpha + b\beta + c\gamma + \cdots)^2] + (\rho^2 + 2\rho)\mu^4.
\]
VII. Finally we conclude from all these preliminary results that
\[
N = (\varpi - 2\rho)\nu^4 + (\varpi^2 - \varpi - 2\varpi \rho + 4\rho + 2\rho^2)\mu^4 + \\
(\nu^4 - 3\mu^4) \sum [(a\alpha + b\beta + c\gamma + \cdots)^2] \\
= (\varpi - \rho)(\nu^4 - \mu^4) + (\varpi - \rho)^2\mu^4 - \\
(\nu^4 - 3\mu^4) \left[\rho - \sum (a\alpha + b\beta + c\gamma + \cdots)^2\right]
\]
Thus the standard error to be expected when one takes

\[ \mu^2 = \frac{M}{\varpi - \rho} \]

will be

\[ \sqrt{\frac{\nu^4 - \mu^4}{\varpi - \rho} - \frac{\nu^4 - 3\mu^4}{(\varpi - \rho)^2} \left[ \rho - \sum (a\alpha + b\beta + c\gamma + \cdots)^2 \right]} \]

\[ 40. \]

The quantity \( \sum [(a\alpha + b\beta + c\gamma + \cdots)^2] \) which comes into the preceding expression cannot be reduced to a simpler form in general. One can however give two limits within which its value is necessarily included.

1. From the preceding relations one easily obtains

\[
(a\alpha + b\beta + c\gamma + \cdots)^2 + (a\alpha' + b\beta' + c\gamma' + \cdots)^2 + \\
(a\alpha'' + b\beta'' + c\gamma'' + \cdots)^2 + \cdots = a\alpha + b\beta + c\gamma + \cdots;
\]

and from this we conclude that

\[ a\alpha + b\beta + c\gamma + \cdots \]

is a positive quantity less than one, or at least no greater than one. The same is true of the quantity

\[ a\alpha' + b\beta' + c\gamma' + \cdots, \]

which is equal to the sum

\[
(a\alpha' + b\beta' + c\gamma' + \cdots)^2 + (a\alpha' + b\beta' + c\gamma' + \cdots)^2 + \\
(a\alpha'' + b\beta'' + c\gamma'' + \cdots)^2 + \cdots;
\]

similarly, \( a''\alpha'' + b''\beta'' + c''\gamma'' + \cdots \) will be less than one; and so on. Thus

\[ \sum [(a\alpha + b\beta + c\gamma + \cdots)^2] \]

is necessarily less than \( \varpi \).

2. One has

\[ \sum [(a\alpha + b\beta + c\gamma + \cdots)^2] = \rho, \]

since \( \sum a\alpha = 1, \sum b\beta = 1, \sum c\gamma = 1, \ldots \); from this one derives easily that \( \sum [(a\alpha + b\beta + c\gamma + \cdots)^2] \) is greater than, or at least is no less than \( \frac{\varpi^2}{\varpi} \).

Consequently the term

\[ \frac{\nu^4 - 3\mu^4}{(\varpi - \rho)^2} \left[ \rho - \sum (a\alpha + b\beta + c\gamma + \cdots)^2 \right] \]

necessarily lies between the limits

\[ -\frac{\nu^4 - 3\mu^4}{\varpi - \rho} \text{ and } \frac{\nu^4 - 3\mu^4 \rho}{\varpi - \rho} \]
and hence between the wider limits

\[-\frac{\nu^4 - 3\mu^4}{\varpi - \rho} \quad \text{and} \quad \frac{\nu^4 - 3\mu^4}{\varpi - \rho}\]

Thus the square of the standard error to be expected for the value

\[\mu^2 = \frac{M}{\varpi - \rho}\]

is included between the limits

\[\frac{2\nu^4 - 4\mu^4}{\varpi - \rho} \quad \text{and} \quad \frac{2\mu^4}{\varpi - \rho};\]

so that one can attain as high a degree of precision as one wishes, provided that the number of observations is sufficiently large.

It is very remarkable that under the hypothesis of section 9 (III) on which we had previously based the theory of least squares, the second term in the expression for the square of the standard error disappears completely (since one has \(\nu^4 - 3\mu^4 = 0\)); and since in order to find \(\mu\), the approximate value of the standard error of the observations, it is necessary in every case to treat the sum

\[\lambda^2 + \lambda'^2 + \lambda''^2 + \cdots = M\]

as if it were equal to the sum of the squares of \(\varpi - \rho\) random errors; it results from this, that under this hypothesis, the precision of this determination becomes equal to that which we found in section 15 for the determination deduced from \(\varpi - \rho\) true errors.
Supplement

to Theory of the Combination of Observations

which leads to the Smallest Errors

1826

In the preceding memoir, we supposed that the quantities to be determined by
means of imperfect observations depended on certain unknown parameters, in terms
of which they could be expressed as functions: the problem then consisted in deducing
from the observations as exactly as possible, the value of these parameters.

In the majority of cases, the question is presented in essentially this way; but some-
times the situation is slightly different, so that one might even wonder, at first glance,
whether the problem could be reduced to the preceding. It is not unusual, in fact, for the
quantities which are observed not to be explicitly expressed as functions of parameters
and for them not to appear reducible to such a form, except by difficult or ambiguous
operations. On the other hand, it often happens that the nature of the problem furnishes
certain conditions which the observed values should satisfy exactly.

However, on closer examination, one sees that this case does not differ essentially
from the preceding, and that it can be reduced to it. In fact, if one calls \( \varpi \) the number
of observed magnitudes, and \( \sigma \) the number of conditioning equations, nothing prevents us
from choosing \( \varpi - \sigma \) of the observed magnitudes and considering them as the only un-
knowns, the others being considered as functions of them, defined by the conditioning
equations. By this artifice, we reduce the problem to the case of the preceding memoir.

Nevertheless, although this procedure often leads to the result in a fairly convenient
manner, one cannot deny that it is somewhat unnatural, and consequently, it is desirable
to treat the problem under the other form, which, moreover, admits a very elegant solu-
tion. Furthermore this new solution leads to more rapid calculations than the preceding
whenever \( \sigma \) is less than \( \frac{1}{2} \varpi \), or, what comes to the same thing, whenever the number
of elements which we have denoted by \( \rho \) in the preceding memoir is greater than \( \frac{\varpi}{2} \); it
is in this case to be preferred even though it would be quite easy from the nature of the
problem to get rid of the conditioning equations without ambiguity.

2.

Let us denote by \( \nu, \nu', \nu'' \), etc. the \( \varpi \) quantities, whose values are furnished by the
observations. Let us suppose that a single unknown depends on them and that it is
expressed by a known function \( u(\nu, \nu', \nu'' \ldots) \). Let \( l, l', l'' \), etc. be the values of the
derivatives

\[
\frac{du}{d\nu}, \quad \frac{du}{d\nu'}, \quad \frac{du}{d\nu''}, \quad \ldots,
\]

taken at the true values of \( \nu, \nu', \nu'' \), etc., If one were to substitute the true values
of \( \nu, \nu', \nu'' \), etc. in the function \( u \), one would obtain the true value of \( u \); but if the
observations were effected by errors, \( e, e', e'' \), etc., a total error for \( u \) would result
which would be represented by

\[
le + l'e' + l''e'' + \cdots,
\]

provided that, as we shall always suppose (when \( u \) is not linear), one may neglect the
squares and products of the errors \( e, e', e'' \), etc.
Although the magnitude of the errors $e$, $e'$, $e''$, etc. is uncertain, the uncertainty attached to the value found for $u$ can generally be measured by the standard error to be expected in the determination adopted. According to the principles developed in the first memoir, this standard error is

$$\sqrt{l^2m^2 + l'^2m'^2 + l''^2m''^2 + \cdots},$$

where $m$, $m'$, $m''$, etc. are the standard errors of the various observations. If all the observations are affected with the same degree of uncertainty, this expression becomes

$$m\sqrt{l^2 + l'^2 + l''^2 + \cdots}$$

It is clear, moreover, that up to the degree of approximation which we are considering, one may replace $l, l', l''$, etc. by the values of the derivatives

$$\frac{du}{d\nu}, \frac{du}{d\nu'}, \frac{du}{d\nu''}, \ldots,$$

taken at the observed values of $\nu, \nu', \nu''$, etc.

3.

When the quantities $\nu, \nu', \nu''$, etc. are independent, the unknown can be determined in only one way, and the uncertainty attached to the result can be neither avoided nor diminished. The observations give a value of the unknown which, in this case, has nothing arbitrary about it. The situation is completely different when the quantities $\nu, \nu', \nu''$, etc. are connected by necessary relations, which we shall suppose expressed by $\sigma$ equations,

$$X = 0, \quad Y = 0, \quad Z = 0, \quad \ldots,$$

where $X, Y, Z, \ldots$ denote given functions of the variables $\nu, \nu', \nu''$, etc.; for in this case, one can substitute for the function $u$, any other function $U$ such that the difference $U - u$ vanishes identically in consequence of the equations

$$X = 0, \quad Y = 0, \quad Z = 0, \quad \ldots,$$

If the observations were absolutely exact, this substitution would in no way change the results; but because of the inevitable errors, a different result will correspond to each form chosen for $u$, and the error committed instead of being

$$le + l'e' + l''e'' + \cdots,$$

will become

$$Le + L'e' + L''e'' + \cdots,$$

where $L, L', L''$, etc. denote the derivatives

$$\frac{dU}{d\nu}, \frac{dU}{d\nu'}, \frac{dU}{d\nu''}, \ldots$$
Although it is impossible to determine the value of the various errors, we can nevertheless compare the standard errors to be expected in the use of the various combinations. The most advantageous combination will be the one which gives a minimum value for the standard error. This error being

$$\sqrt{L^2m^2 + L'^2m'^2 + L''m''^2 + \cdots}$$

we must seek to make the sum

$$L^2m^2 + L'^2m'^2 + L''m''^2 + \cdots$$

as small as possible.

The function $U$, infinite in number, by which one may replace $u$, will be distinguished from each other in our considerations, except by the values which they give for $L, L', L''$, etc.; thus we must first find the relations which hold between the systems of values which these coefficients can take. Let us denote by

$$a, a', a'', \ldots,$$
$$b, b', b'', \ldots,$$
$$c, c', c'', \ldots,$$
$$\vdots \quad \vdots \quad \vdots \quad \ddots$$

the values of the derivatives

$$\frac{dX}{d\nu}, \frac{dX}{d\nu'}, \frac{dX}{d\nu''}, \cdots,$$
$$\frac{dY}{d\nu}, \frac{dY}{d\nu'}, \frac{dY}{d\nu''}, \cdots,$$
$$\frac{dZ}{d\nu}, \frac{dZ}{d\nu'}, \frac{dZ}{d\nu''}, \cdots,$$
$$\vdots \quad \vdots \quad \vdots \quad \ddots$$

taken at the true values of $\nu, \nu', \nu''$, etc. It is clear that if one gives $\nu, \nu', \nu''$, etc., increments $d\nu, d\nu', d\nu''$, etc., which do not change $X, Y, Z$, etc., and consequently leave their value 0, these increments, which will satisfy the equations

$$0 = a'd\nu + a''d\nu'' + \cdots,$$
$$0 = b'd\nu' + b''d\nu'' + \cdots,$$
$$0 = c'd\nu' + c''d\nu'' + \cdots,$$
$$\vdots$$

will not change the value of $U - u$, and consequently give

$$0 = (l - L)d\nu + (l' - L')d\nu' + (l'' - L'')d\nu'' + \cdots$$
From this one easily concludes that $L$, $L'$, $L''$, etc. must have the form

\[
L = l + ax + by + cz + \cdots, \\
L' = l' + a'x + b'y + c'z + \cdots, \\
L'' = l'' + a''x + b''y + c''z + \cdots,
\]

where $x$, $y$, $z$, etc. denote certain fixed coefficients. Conversely, it is clear that for all values of $x$, $y$, $z$, etc., one can form a function $U$ for which the values $L$, $L'$, $L''$, etc. will be precisely those furnished by these equations, and, by our preceding remarks, this function can be substituted for $u$. The simplest form that one can give it is

\[
U = u + xX + yY + zZ + \cdots + u',
\]

where $u'$ denotes a function of $\nu$, $\nu'$, $\nu''$, etc., which vanishes identically when $X$, $Y$, $Z$, etc. are 0, and whose value in the actual case will be a maximum or minimum (since its derivatives must all vanish). But introducing this function makes no difference to the results.

5.

It is now easy to assign values to $x$, $y$, $z$, etc. which make the sum

\[
L^2 m^2 + L'^2 m'^2 + L''^2 m''^2 + \cdots,
\]

a minimum.

It is clear that in order to achieve this end, knowledge of the absolute standard errors $m$, $m'$, $m''$, etc. is not necessary; it is sufficient to know their ratios. Let us then introduce, in place of these quantities, the weights of the observations, $p$, $p'$, $p''$, etc., that is to say numbers reciprocally proportional to $m^2$, $m'^2$, $m''^2$, etc. The quantities $x$, $y$, $z$, etc. must be determined in such a way that the polynomial

\[
\frac{(ax + by + cz + \cdots + l)^2}{p} + \frac{(a'x + b'y + c'z + \cdots + l')^2}{p'} + \cdots
\]

takes on its minimum value. Suppose that $x^0$, $y^0$, $z^0$, etc. are the fixed values of $x$, $y$, $z$, etc., giving this minimum, and let us adopt the following notations:

\[
\frac{a^2}{p} + \frac{a'^2}{p'} + \frac{a''^2}{p''} + \cdots = (aa) \\
\frac{ab}{p} + \frac{a'b'}{p'} + \frac{a''b''}{p''} + \cdots = (ab) \\
\frac{ac}{p} + \frac{a'c'}{p'} + \frac{a''c''}{p''} + \cdots = (ac) \\
\cdots
\]
The condition for a minimum obviously requires that one have

\[
\begin{align*}
0 &= (aa)x^0 + (ab)y^0 + (ac)z^0 + \cdots + (al), \\
0 &= (ba)x^0 + (bb)y^0 + (bc)z^0 + \cdots + (bl), \\
0 &= (ca)x^0 + (cb)y^0 + (cc)z^0 + \cdots + (cl), \\
\vdots
\end{align*}
\]

(1)

After the values of \(x^0, y^0, z^0\), etc. have been determined from these equations, one will put

\[
\begin{align*}
ax^0 + by^0 + cz^0 + \cdots + l &= L, \\
a'x^0 + b'y^0 + c'z^0 + \cdots + l' &= L', \\
a''x^0 + b''y^0 + c''z^0 + \cdots + l'' &= L'', \\
\vdots
\end{align*}
\]

(2)

and the most suitable function for determining our unknown, the one with the smallest standard error, will be the one whose derivatives are equal to \(L, L', L''\), etc. for the values of the variables under consideration. The weight of the determination obtained in this way will be

\[
P = \frac{1}{\frac{L^2}{p} + \frac{L'^2}{p'} + \frac{L''^2}{p''} + \cdots}
\]

(3)

that is to say, that \(\frac{1}{P}\) will be precisely the value which the polynomial considered above takes when \(x, y, z\), etc. are given the values satisfying the equations (1).
6.

In the preceding section we showed how to determine the function \( U \) which gives the most suitable determination of the unknown \( u \). Let us now examine the value which results from it. Let us denote this value by \( K \); one obtains it by substituting the observed values of the quantities \( \nu, \nu', \nu'' \), etc. into \( U \). Let \( k \) be the value taken by \( u \) when one makes the same substitutions, and finally, let \( \kappa \) be the true value of this unknown, which one would obtain by substitution of the true values of \( \nu, \nu', \nu'' \), etc. either in \( u \) or in \( U \). One will have

\[
k = \kappa + le + l'e' + l''e'' + \cdots,
\]

\[
K = \kappa + Le + L'e' + L''e'' + \cdots,
\]

and consequently,

\[
K = k + (L - l)e + (L' - l')e' + (L'' - l'')e'' + \cdots.
\]

Substituting in this equation the values furnished by (2) in place of \( L - l, L' - l', L'' = l'' \), etc., and setting

\[
\begin{align*}
\text{(4)}
\begin{cases}
ae + a'e' + a''e'' + \cdots &= A, \\
b'e + b'e' + b''e'' + \cdots &= B, \\
c'e' + c''e'' + \cdots &= C,
\end{cases}
\end{align*}
\]

we obtain

\[
(5)
K = k + Ax^0 + By^0 + Cz^0 + \cdots
\]

It is not possible to calculate \( A, B, C, \) etc. by means of the formulas (4) since the errors \( e, e', e'' \), etc., which occur in the formulas, have unknown values, but one sees exactly that these quantities \( A, B, C, \) etc. are nothing but the values of \( X, Y, Z, \) etc. which correspond to the observed values of \( \nu, \nu', \nu'' \), etc.; then the system of equations (1), (3), (5), gives the complete solution for our problem. Indeed, it is clear that one can apply the remark made at the end of section 2 about the quantities \( l, l', l'' \), etc. to the calculation of \( a, a', a'', \ldots, b, b', b'', \) etc. that is to say we replace the true values of \( \nu, \nu', \nu'' \), etc. by the observed values.

7.

One may replace the formula (3), which represents the weight of the most probable determination, by several expressions which it is useful to indicate; to begin with, note that by adding the equations (2) after multiplying them by \( \frac{a}{p}, \frac{a'}{p'}, \frac{a''}{p''}, \) etc. one will have

\[
(aa)x^0 + (ab)y^0 + (ac)z^0 + \cdots = \frac{aL}{p} + \frac{a'L'}{p'} + \frac{a''L''}{p''} + \cdots
\]
The first term is zero; thus denoting the second member by \((aL)\), in accordance with the notation we adopted, one has

\[(aL) = 0\]

and similarly

\[\begin{align*}
(bL) &= 0 \\
(cL) &= 0 \\
&
\end{align*}\]

Finally, let us multiply the equations (2) by \(\frac{L}{p'}, \frac{L'}{p'}, \frac{L''}{p''}, \text{ etc.}\); we shall have on adding

\[
\frac{ll}{p} + \frac{l'l'}{p'} + \frac{l''l''}{p''} + \cdots = \frac{L^2}{p} + \frac{L'^2}{p'} + \frac{L''^2}{p''} + \cdots,
\]

and consequently, we obtain this second expression for the weight

\[P = \frac{1}{\frac{ll}{p} + \frac{l'l'}{p'} + \frac{l''l''}{p''} + \cdots}\]

If, finally, we add the same equations (2) after multiplying them by \(\frac{l}{p}, \frac{l'}{p'}, \frac{l''}{p''}, \text{ etc.}\), we obtain a third expression for the weight

\[P = \frac{1}{(al)x^0 + (bl)y^0 + (cl)z^0 + \cdots + (ll)}\]

where we have put

\[
\frac{ll}{p} + \frac{l'l'}{p'} + \frac{l''l''}{p''} + \cdots = (ll).
\]

From this one obtains easily the fourth expression for the weight

\[
\frac{1}{P} = (ll) - (aa)x^0 + (bb)y^0 - (cc)z^0 + \cdots
- 2(ab)x^0y^0 - 2(ac)x^0z^0 - 2(bc)y^0z^0 + \cdots
\]

8.

The general solution which we have just expounded applies principally to the case in which one has a single unknown to determine. On the other hand, when one wants to find the most likely values of several unknowns, depending on the same observations, or when one does not know which unknowns it is preferable to derive from the observations, it is convenient to proceed in a different way, which we shall now consider.

Let us consider \(x, y, z, \text{ etc.}\) as unknowns, and set

\[
\begin{align*}
(aa)x + (ab)y + (ac)z + \cdots &= \xi \\
(ab)x + (bb)y + (bc)z + \cdots &= \eta \\
(ac)x + (bc)y + (cc)z + \cdots &= \zeta \\
&
\end{align*}
\]

(6)
Suppose that one obtains from these, by elimination,

\[
\begin{align*}
(\alpha\alpha)\xi + (\alpha\beta)\eta + (\alpha\gamma)\zeta + \cdots &= x \\
(\beta\alpha)\xi + (\beta\beta)\eta + (\beta\gamma)\zeta + \cdots &= y \\
(\gamma\alpha)\xi + (\gamma\beta)\eta + (\gamma\gamma)\zeta + \cdots &= z \\
&\vdots
\end{align*}
\]  

(7)

Note that the symmetrically placed coefficients are necessarily equal, that is to say that

\[
\begin{align*}
(\beta\alpha) &= (\alpha\beta) \\
(\gamma\alpha) &= (\alpha\gamma) \\
(\gamma\beta) &= (\beta\gamma) \\
&\vdots
\end{align*}
\]

which is a result of the theory of solution of equations which we shall demonstrate later. We have

\[
\begin{align*}
x^0 &= -(\alpha\alpha)(al) - (\alpha\beta)(bl) - (\alpha\gamma)(cl) - \cdots, \\
y^0 &= -(\beta\alpha)(al) - (\beta\beta)(bl) - (\beta\gamma)(cl) - \cdots, \\
z^0 &= -(\gamma\alpha)(al) - (\gamma\beta)(bl) - (\gamma\gamma)(cl) - \cdots, \\
&\vdots
\end{align*}
\]

(8)

and setting

\[
\begin{align*}
(\alpha\alpha)A + (\alpha\beta)B + (\alpha\gamma)C + \cdots &= A \\
(\beta\alpha)A + (\beta\beta)B + (\beta\gamma)C + \cdots &= B \\
(\gamma\alpha)A + (\gamma\beta)B + (\gamma\gamma)C + \cdots &= C \\
&\vdots
\end{align*}
\]

(9)

we obtain

\[ K = k - A(al) - B(bl) - C(cl) - \cdots, \]

which, if we put

\[
\begin{align*}
aA + bB + cC + \cdots &= p\epsilon, \\
a' A + b'B + c'C + \cdots &= p'\epsilon', \\
a'' A + b''B + c''C + \cdots &= p''\epsilon'', \\
&\vdots
\end{align*}
\]

(10)

becomes

\[ K = k - \epsilon - l\epsilon' - l'\epsilon'' - \cdots \]

(11)
Comparing the equations (7) and (9) shows that the auxiliary quantities $A$, $B$, $C$, etc. are the values taken by the unknowns $x$, $y$, $z$, etc., when one assumes

$$
\xi = A, \quad \eta = B, \quad \zeta = C, \quad \ldots;
$$

from this one obviously concludes

$$
\begin{align*}
(aa)A + (ab)B + (ac)C + \cdots &= A \\
(ba)A + (bb)B + (bc)C + \cdots &= B \\
(ca)A + (cb)B + (cc)C + \cdots &= C \\
&\vdots
\end{align*}
$$

(12)

Adding the equations (10) after having multiplied them by $\frac{a}{p}$, $\frac{a'}{p'}$, $\frac{a''}{p''}$, etc. one obtains

$$
\begin{align*}
A &= a\epsilon + a'\epsilon' + a''\epsilon'' + \cdots, \\
and similarly
B &= b\epsilon + b'\epsilon' + b''\epsilon'' + \cdots, \\
C &= c\epsilon + c'\epsilon' + c''\epsilon'' + \cdots, \\
&\vdots
\end{align*}
$$

(13)

Since $A$ is, as we have said, the value of $X$ when the observed values are substituted for $\nu$, $\nu'$, $\nu''$, etc., one sees easily that if the corrections $-\epsilon$, $-\epsilon'$, $-\epsilon''$, etc. are applied to each of these quantities, the value of $X$ will become equal to 0, and that similarly $Y$, $Z$, etc. will vanish. The equation (11) shows also that $K$ is the value taken by $u$ as a consequence of the same substitutions.

Defining *compensation of observations* to be the application of the corrections $-\epsilon$, $-\epsilon'$, $-\epsilon''$, etc. to the directly observed magnitudes, we are led to a very important consequence:

The compensated observations, as we have shown, satisfy all the conditioning equations exactly, and give every function of the observed quantities the value which would result from the most suitable combinations of the unmodified observations; and although the conditioning equations are too few in number, for one to deduce from them the exact value of the errors, we have at least found, by the preceding, the *most likely* errors. It is under this name that the quantities $\epsilon$, $\epsilon'$, $\epsilon''$, etc. have hitherto been designated.

10.

When the number of observations is larger than that of the conditioning equations, besides the system of most likely corrections, one can find an infinite number which exactly satisfy the conditioning equations.
It is worth while to examine the relations interconnecting these various systems. Let \(-E, -E', -E'', \ldots\) be such a system of corrections distinct from the most likely system; we shall have

\[
\begin{align*}
aE + a'E' + a''E'' + \cdots &= \mathcal{A} \\
bE + b'E' + b''E'' + \cdots &= \mathcal{B} \\
cE + c'E' + c''E'' + \cdots &= \mathcal{C} \\
&\vdots
\end{align*}
\]

Multiplying these equations by \(A, B, C, \ldots\), and adding, and taking note of equations \((10)\), we obtain

\[
peE + p'\epsilon' E' + p''\epsilon'' E'' + \cdots = A\mathcal{A} + B\mathcal{B} + C\mathcal{C} + \cdots
\]

But the equations \((13)\) combined in the same way give

\[
pe^2 + p'\epsilon'^2 + p''\epsilon''^2 + \cdots = A\mathcal{A} + B\mathcal{B} + C\mathcal{C} + \cdots ;
\]

and by combining these two results one easily deduces

\[
pe^2 + p'\epsilon'^2 + p''\epsilon''^2 + \cdots = pe^2 + p'\epsilon'^2 + p''\epsilon''^2 + \cdots \\
+ p(E - \epsilon)^2 + p'(E' - \epsilon')^2 + p''(E'' - \epsilon'')^2 + \cdots
\]

and consequently, the sum

\[
pE^2 + p'\epsilon'^2 + p''\epsilon''^2 + \cdots
\]

is necessarily larger than

\[
pe^2 + p'\epsilon'^2 + p''\epsilon''^2 + \cdots ,
\]

and this may be stated in the following way.

**Theorem:** The sum of the squares of corrections which reconcile the observations with the conditioning equations, when multiplied respectively by the weights of the corresponding observations, give a sum which is a minimum for the most likely corrections.

One recognizes precisely the principles of least squares, from which, moreover, the equations \((12)\) and \((10)\) may easily be deduced. The minimum sum, which we shall henceforth denote by \(S\), is equal, according to equation \((14)\), to

\[
A\mathcal{A} + B\mathcal{B} + C\mathcal{C} + \cdots .
\]

11.

The determination of the most likely errors, which is independent of \(l, l', l'', \ldots\), gives the most convenient procedure, whatever the use one wishes finally to make of the observations. Besides, one sees without difficulty, that to achieve this end, it is not necessary to carry out the *indefinite* elimination, that is to say, to calculate \((\alpha\alpha), (\alpha\beta), \ldots\); it suffices to deduce from these equations \((12)\), by a *definite* elimination the auxiliary
quantities $A, B, C,$ etc. which we shall call, in the following, the *correlatives* of the conditioning equations

$$X = 0, \quad Y = 0, \quad Z = 0, \quad \ldots$$

As a final step, these quantities will be substituted in the equation (10).

This method leaves nothing to be desired when one wants only the most likely values of the quantities furnished by the observation. But the situation is different when one wants the weights of each of the values found as well. In this case, whichever of the preceding formulas one wishes to use, it is necessary to know $L, L', L'', \text{etc.}$ or, what comes to the same thing, $x^0, y^0, z^0, \text{etc.;}$ for this reason, it will be useful to study more closely the elimination which gives these quantities and to obtain a more convenient method for the determination of the weights.

12.

The relations between the quantities which we are considering are considerably simplified by the consideration of the second degree function

$$(aa)x^2 + 2(ab)xy + 2(ac)xz + \cdots + (bb)y^2 + 2(bc)yz + \cdots + (cc)z^2 + \cdots,$$

which we shall denote by $T.$

This function is clearly equal to

$$\frac{(ax + by + cz + \cdots)^2}{p} + \frac{(a'x + b'y + c'z + \cdots)^2}{p'} + \cdots$$

Furthermore, one clearly has

$$(16) \quad T = \xi x + \eta y + \zeta z + \cdots;$$

and finally, if one expresses $x,$ $y,$ $z,$ etc. in terms of $\xi, \eta, \zeta,$ etc., one will have

$$T = (aa)\xi^2 + 2(\alpha\beta)\xi\eta + 2(\alpha\gamma)\xi\zeta + \cdots$$
$$+ (\beta\beta)\eta^2 + 2(\beta\gamma)\eta\zeta + \cdots + (\gamma\gamma)\zeta^2 + \cdots$$

The theory developed above furnishes two systems of values for the quantities $x, y, z,$ etc., $\xi, \eta, \zeta,$ etc. The first is

$$x = x^0, \quad y = y^0, \quad z = z^0, \quad \ldots,$$
$$\xi = -(al), \quad \eta = -(bl), \quad \zeta = -(cl), \quad \ldots;$$

and to this system the value

$$T = (ll) - \frac{1}{P}$$

corresponds, as one sees by comparing equation (16) with the third form of the weight $P,$ or by direct consideration of the form (4).
The second system of values is

\[ x = A, \quad y = B, \quad z = C, \quad \ldots, \]
\[ \xi = A, \quad \eta = B, \quad \zeta = C, \quad \ldots; \]

and the corresponding value of \( T \) is \( T = S \), as is clear from the formulas (10) and (15), and again from the formulas (14) and (16).

13.

We should, first of all, subject the function \( T \) to a transformation similar to that which has been indicated in *Theoria Motus*, article 182 and, with further developments in the *Researches on Pallas*.

To this end, let us set

\[
(bb, 1) = (bb) - \frac{(ab)}{(aa)},
\]

\[
(bc, 1) = (bc) - \frac{(ab)(ac)}{(aa)},
\]

\[
(bd, 1) = (bd) - \frac{(ab)(ad)}{(aa)},
\]

\[
\vdots
\]

\[
(cc, 2) = (cc) - \frac{(ac)^2}{(aa)} - \frac{(bc, 1)^2}{(bb, 1)},
\]

\[
(cd, 2) = (cd) - \frac{(ac)(ad)}{(aa)} - \frac{(bc, 1)(bd, 1)}{(bb, 1)},
\]

\[
\vdots
\]

\[
(dd, 3) = (dd) - \frac{(ad)^2}{(aa)} - \frac{(bd, 1)^2}{(bb, 1)} - \frac{(cd, 2)^2}{(cc, 2)},
\]

\[
\vdots
\]

and, finally\(^8\)

\[
(bb, 1)y + (bc, 1)z + (bd, 1)w + \cdots = \eta',
\]

\[
(cc, 2)z + (cd, 2)w + \cdots = \zeta'',
\]

\[
(dd, 3)w + \cdots = \phi''',
\]

\[
\vdots
\]

One will have

\[
T = \frac{\xi^2}{(aa)} + \frac{\eta'^2}{(bb, 1)} + \frac{\zeta''^2}{(cc, 2)} + \frac{\phi'''^2}{(dd, 3)} + \cdots
\]

---

\(^8\)In the preceding calculations it was sufficient to give three letters in each series to make clear the law of formulation; it has seemed necessary to use a fourth here, to make the algorithm clearer. —G.
The quantities \( \eta', \zeta'', \phi''' \), etc. can be obtained from \( \xi, \eta, \zeta, \phi \), etc. by the following equations

\[
\eta' = \eta - \frac{(ab)}{(aa)} \xi,
\]
\[
\zeta'' = \zeta - \frac{(bc, 1)}{(bb, 1)} \eta',
\]
\[
\phi''' = \phi - \frac{(ad)}{(aa)} \xi - \frac{(bd, 1)}{(bb, 1)} \eta' - \frac{(cd, 2)}{(cc, 2)} \zeta'',
\]
\[ \vdots \]

and from these we shall easily obtain all the formulas needed for our ends. Thus, to determine the correlatives \( A, B, C \), etc. we put

\[
\begin{align*}
B' &= B - \frac{(ab)}{(aa)} A, \\
C'' &= C - \frac{(ac)}{(aa)} A - \frac{(bc, 1)}{(bb, 1)} B', \\
D''' &= D - \frac{(ad)}{(aa)} A - \frac{(bd, 1)}{(bb, 1)} B' - \frac{(cd, 2)}{(cc, 2)} C'', \\
\vdots
\end{align*}
\]

and then \( A, B, C, D \), etc. are obtained by the following formulas beginning with the last one:

\[
\begin{align*}
(aa) A + (ab) B + (ac) C + (ad) D + \cdots &= A, \\
(bb, 1) B + (bc, 1) C + (bd, 1) D + \cdots &= B', \\
(cc, 2) C + (cd, 2) D + \cdots &= C'', \\
(dd, 3) D + \cdots &= D''', \\
\vdots
\end{align*}
\]

For \( S \) we have the new formula

\[
S = \frac{A^2}{(aa)} + \frac{B'^2}{(bb, 1)} + \frac{C''^2}{(cc, 2)} + \frac{D'''^2}{(dd, 3)} + \cdots
\]

and finally, the weight \( P \) which is to be attributed to the most likely determination of the quantity \( u \), will be given by the formula

\[
\frac{1}{P} = (ll) - \frac{(al)^2}{(aa)} - \frac{(bl, 1)^2}{(bb, 1)} - \frac{(cl, 2)^2}{(cc, 2)} - \frac{(dl, 3)^2}{(dd, 3)} - \cdots;
\]
in this formula,

\begin{equation}
\begin{cases}
(bl, 1) = (bl) - \frac{(ab)(al)}{(aa)} \\
(cl, 2) = (cl) - \frac{(ac)(al)}{(aa)} - \frac{(bc)(bl)}{(bb)} \\
(dl, 3) = (dl) - \frac{(ad)(al)}{(aa)} - \frac{(bd)(bl)}{(bb)} - \frac{(cd)(cl)}{(cc)}
\end{cases}
\end{equation}

The formulas (17), . . . , (21), whose simplicity leaves nothing to be desired, furnish the complete solution of our problem.

14.

Having solved the problem we were aiming for, we shall tackle a few secondary questions which will further clarify this theory.

In the first place, we shall examine whether it can happen that the elimination which gives \( x, y, z, \) etc. as a function of \( \xi, \eta, \zeta, \) etc., could, in certain cases become impossible. This would clearly happen if the functions \( \xi, \eta, \zeta, \) etc. were not independent of each other. Let us suppose for the moment that it were so, so that one of them could be expressed as a function of the others so that the relation

\[ \alpha \xi + \beta \eta + \gamma \zeta + \cdots = 0, \]

held identically, \( \alpha, \beta, \gamma, \) etc., denoting fixed numbers.

Then one will have

\[ \alpha(aa) + \beta(ab) + \gamma(ac) + \cdots = 0, \]
\[ \alpha(ab) + \beta(bb) + \gamma(bc) + \cdots = 0, \]
\[ \alpha(ac) + \beta(bc) + \gamma(cc) + \cdots = 0, \]

\[ \vdots \]

and if we set

\[ \alpha a + \beta b + \gamma c + \cdots = p\Theta, \]
\[ \alpha a' + \beta b' + \gamma c' + \cdots = p'\Theta', \]
\[ \alpha a'' + \beta b'' + \gamma c'' + \cdots = p''\Theta'', \]

\[ \vdots \]

we deduce from these

\[ a\Theta + a'\Theta' + a''\Theta'' + \cdots = 0, \]
\[ b\Theta + b'\Theta' + b''\Theta'' + \cdots = 0, \]
\[ c\Theta + c'\Theta' + c''\Theta'' + \cdots = 0, \]

\[ \vdots \]
and consequently

\[ p\Theta^2 + p'\Theta'^2 + p''\Theta''^2 + \cdots = 0; \]

since \( p, p', p'' \), etc. are positive by their nature, this equation requires

\[ \Theta = 0, \quad \Theta' = 0, \quad \Theta'' = 0, \quad \ldots, \]

If we consider the total differentials \( dX, dY, dZ \), etc., corresponding to the values of \( \nu, \nu', \nu'' \), etc. given directly by the observations, these differentials

\[
\begin{align*}
& a d\nu + a' d\nu' + a'' d\nu'' + \cdots, \\
& b d\nu + b' d\nu' + b'' d\nu'' + \cdots, \\
& c d\nu + c' d\nu' + c'' d\nu'' + \cdots, \\
& \vdots
\end{align*}
\]

according to the preceding results will be inter-related with each other so that by multiplying them respectively by \( \alpha, \beta, \gamma \), etc. the sum of the products will be identically 0, so that among the equations

\[ X = 0, \quad Y = 0, \quad Z = 0, \quad \ldots, \]

there is at least one which we may consider as useless for it will be satisfied whenever the others are.

On examining the equation more closely, one sees that this conclusion is applicable only to values of the variables which are infinitesimally different from those given by the observations. There are indeed two cases to be distinguished: the first is that in which one of the equations \( X = 0, Y = 0, Z = 0, \ldots, \) is included among the others in a general and absolute way, and can consequently be suppressed; the second is that in which one of the functions \( X, Y, Z, \) etc., \( X \) for instance, takes on a maximum or minimum value or, more generally, a value such that the differential vanishes when the other equations remain satisfied, for the particular values of \( \nu, \nu', \nu'', \) etc. which the observations give.

But since we consider only variations in our variables whose squares are negligible, this second case (which comes up only very rarely in applications) can be assimilated to the first, and one of the conditioning equations can be suppressed as superfluous.

If the remaining equations are independent in the sense which we have just indicated, one may, according to the preceding, be sure that the elimination is possible. We shall, however, restrain ourselves from further discussion on this matter, which merits attention as a theoretical subtlety rather than as a question of practical utility.

15.

In the first memoir, section 37 and following, we showed how to determine, a posteriori, to a very good approximation, the weight of a determination. If the approximate
value of \( \varpi \) quantities are given by equally precise observations, and one compares them with the values which result for them from the most likely hypotheses which one can make on the \( \rho \) elements on which they depend, we have seen that the result of adding the squares of the differences obtained and dividing the sum by \( \varpi - \rho \) could be considered an approximate value of the square of the standard error inherent in this type of observation.

If the observations are unequally precise, the only modification which one must apply to the preceding precepts, is that one must multiply the squares of the differences by the respective weights of the corresponding observations, and the standard error obtained this way is relative to that of those observations whose weight is taken as unity.

In the present case, the sum of the squares which we speak of obviously coincides with the sum \( S \), and the differences \( \varpi - \rho \) with the number \( \sigma \) of conditioning equations. Consequently, for the standard error of observations of unit weight, we have the expression \( \sqrt{\frac{S}{\sigma}} \) and the determination will be the more worthy of confidence the larger \( \sigma \) is.

It is worth while to establish this result independently of the arguments of the first memoir; in doing so, we shall introduce several new notations. Let us suppose that to the values \( \xi = a, \, \eta = b, \, \zeta = c, \, \ldots, \) there correspond \( x = \alpha, \, y = \beta, \, z = \gamma, \, \ldots, \) so that one has
\[
\begin{align*}
\alpha &= a(\alpha \alpha) + b(\alpha \beta) + c(\alpha \gamma) + \cdots \\
\beta &= a(\alpha \beta) + b(\beta \beta) + c(\beta \gamma) + \cdots \\
\gamma &= a(\alpha \gamma) + b(\gamma \beta) + c(\gamma \gamma) + \cdots 
\end{align*}
\]
and that to the values \( \xi = a', \, \eta = b', \, \zeta = c', \, \ldots, \) there correspond \( x = \alpha', \, y = \beta', \, z = \gamma', \, \ldots; \) and finally, that to the values \( \xi = a'', \, \eta = b'', \, \zeta = c'', \, \ldots, \) there correspond \( x = \alpha'', \, y = \beta'', \, z = \gamma'', \, \ldots, \) and so on.

The combinations of equations (4) and (9) yields
\[
\begin{align*}
A &= \alpha e + \alpha' e' + \alpha'' e'' + \cdots, \\
B &= \beta e + \beta' e' + \beta'' e'' + \cdots, \\
C &= \gamma e + \gamma' e' + \gamma'' e'' + \cdots
\end{align*}
\]
and, since one has
\[ S = AA + BB + CC + \cdots, \]
one will have
\[
S = \left( \alpha e + \alpha' e' + \alpha'' e'' + \cdots \right) \left( \alpha e + \alpha' e' + \alpha'' e'' + \cdots \right) + \\
\left( \beta e + \beta' e' + \beta'' e'' + \cdots \right) \left( \beta e + \beta' e' + \beta'' e'' + \cdots \right) + \\
\left( \gamma e + \gamma' e' + \gamma'' e'' + \cdots \right) \left( \gamma e + \gamma' e' + \gamma'' e'' + \cdots \right) + \cdots.
\]

The series of observations which furnish the quantities \( \nu, \nu', \nu'', \)\ etc. affected by the random errors \( e, e', e'' \)\ etc., may be considered as an experiment which does not actually yield the magnitude of each error, but which, by means of the rules expounded above, permits the determination of the quantity \( S \), a known function of all the errors. In such an experiment, some of the errors will be larger, others smaller; but the larger the number of errors used, the greater the probability that \( S \) differs only slightly from its average value; the difficulty thus reduces to finding the average value of \( S \).

By the principles set forth in the first memoir which it is unnecessary to reproduce here, one finds for this average value
\[
(a\alpha + b\beta + c\gamma + \cdots)\mu^2 + (a'\alpha' + b'\beta' + c'\gamma' + \cdots)\mu'\mu'' = \cdots
\]
Denoting by \( \mu \) the standard error which corresponds to observations of unit weight, so that one has
\[
\mu^2 = pm^2 = p' m'^2 = p'' m''^2 = \cdots,
\]
the preceding expression can be written as follows:
\[
\left( \frac{a\alpha}{p} + \frac{a'\alpha'}{p'} + \frac{a''\alpha''}{p''} + \cdots \right) \mu^2 + \left( \frac{b\beta}{p} + \frac{b'\beta'}{p'} + \frac{b''\beta''}{p''} + \cdots \right) \mu'^2 + \\
\left( \frac{c\gamma}{p} + \frac{c'\gamma'}{p'} + \frac{c''\gamma''}{p''} + \cdots \right) \mu''^2
\]
But we have found
\[
\frac{a\alpha}{p} + \frac{a'\alpha'}{p'} + \frac{a''\alpha''}{p''} + \cdots = (aa)(\alpha\alpha) + (ab)(\alpha\beta) + (ac)(\alpha\gamma) + \cdots;
\]
and the right-hand side equals one, as is seen readily by comparing the equations (6) and (7).

One finds similarly
\[
\frac{b\beta}{p} + \frac{b'\beta'}{p'} + \frac{b''\beta''}{p''} + \cdots = 1,
\]
\[
\frac{c\gamma}{p} + \frac{c'\gamma'}{p'} + \frac{c''\gamma''}{p''} + \cdots = 1,
\]
and so on.
According to this, the average value of $S$ is $\sigma \mu^2$, and if one considers it permissible to consider the chance value of $S$ as equal to its average value, one concludes

$$\mu = \sqrt{\frac{S}{\sigma}}$$

17.

One may estimate the reliability of this determination by calculating the standard error to be expected, either in its own value, or in that of its square. The latter will be the square root of the average value of the expression

$$\left( \frac{S}{\sigma} - \mu^2 \right)^2$$

which can be obtained by arguments similar to those set forth in the first memoir (section 39 and following). We omit them for the sake of brevity, and merely indicate the result.

The standard error to be expected in the determination of the square $\mu^2$ is given by

$$\sqrt{\frac{2\mu^2}{\sigma} + \frac{\nu^4 - 3\mu^4}{\sigma^2} N},$$

where $\nu^4$ is the average value of the fourth powers of the errors taken to have unit weight, and $N$ is the sum

$$(a\alpha + b\beta + c\gamma + \cdots)^2 + (a'\alpha' + b'\beta' + c'\gamma' + \cdots)^2 + (a''\alpha'' + b''\beta'' + c''\gamma'' + \cdots)^2 + \cdots.$$ 

This sum cannot in general be simplified; but, by a procedure analogous to that used in section 40 of the first memoir, one can show that its value lies between $\varpi$ and $\frac{\sigma^2}{\varpi}$. Under the hypothesis on which we had originally established the method of least squares, the term containing this sum vanishes because

$$\nu^4 = 3\mu^4,$$

and the precision which ought to be attributed to the determination

$$\mu = \sqrt{\frac{S}{\sigma}}$$

is consequently the same as if one had operated on $\sigma$ observations affected by precisely known errors, in accordance with the precepts of sections 15 and 16 of the first memoir.

18.

In order to adjust the observations, there are, as we have said, two operations to be carried out; firstly, it is necessary to determine the correlatives of the conditioning equations, that is to say, $A, B, C, \text{etc.}$, which satisfy the equations (12); secondly, to
substitute these quantities in the equation (10). The adjustment so obtained may be called \textit{perfect} and \textit{complete} in opposition to \textit{imperfect} and \textit{incomplete} adjustment. We shall apply this last term to adjustments which arise from the same equations (10), when one substitutes therein values of $A$, $B$, $C$, which do not satisfy the equations (12), that is to say, which satisfy only some (or none) of them. We shall not concern ourselves here with such a system of corrections, and we shall not even grant them the name of adjustment.

When the equations (10) are satisfied, the systems (12) and (13) become equivalent, and the difference which we speak of can then be stated as follows: The completely adjusted observations will satisfy the conditioning equations

$$X = 0, \quad Y = 0, \quad Z = 0, \quad \ldots$$

the incompletely adjusted observations will satisfy only some of these equations, and perhaps none; the adjustment which results in satisfying all the equations is necessarily complete.

19.

As a result of the very definition of adjustments, two systems of adjustments taken together will give a third, and one sees that it matters little whether the rules given for obtaining a perfect adjustment are applied to the original observations or to observations already imperfectly adjusted.

Let $-\Theta$, $-\Theta'$, $-\Theta''$, etc. be a system of incomplete adjustments, resulting from the formulas

\begin{equation}
\begin{aligned}
\Theta p &= A^0a + B^0b + C^0c + \cdots, \\
\Theta' p' &= A^0a' + B^0b' + C^0c' + \cdots, \\
\Theta'' p'' &= A^0a'' + B^0b'' + C^0c'' + \cdots, \\
\end{aligned}
\end{equation}

(I)

The observations thus modified do not satisfy all the conditioning equations. Let $A^*$, $B^*$, $C^*$, etc. be the values of $X$, $Y$, $Z$, etc., when the values obtained in this way are substituted for $\nu$, $\nu'$, $\nu''$, etc. One must now find the values $A^*$, $B^*$, $C^*$, etc., which satisfy the equations

\begin{equation}
\begin{aligned}
A^* &= A^*(aa) + B^*(ab) + C^*(ac) + \cdots, \\
B^* &= A^*(ab) + B^*(bb) + C^*(bc) + \cdots, \\
C^* &= A^*(ac) + B^*(bc) + C^*(cc) + \cdots, \\
\end{aligned}
\end{equation}

(II)

and when this is done, the complete adjustment of the modified observations will be accomplished by the new modifications $-\kappa$, $-\kappa'$, $-\kappa''$, etc., where $\kappa$, $\kappa'$, $\kappa''$, etc. are
derived from the formulas

\[
\begin{align*}
\kappa p &= A^*a + B^*b + C^*c + \cdots, \\
\kappa' p' &= A^*a' + B^*b' + C^*c' + \cdots, \\
\kappa'' p'' &= A^*a'' + B^*b'' + C^*c'' + \cdots, \\
\end{align*}
\]

Let us see how these corrections agree with the complete adjustment of the original observations. First it is clear that we have

\[
\begin{align*}
A^* &= A - a\Theta - a'\Theta' - a''\Theta'' - \cdots, \\
B^* &= B - b\Theta - b'\Theta' - b''\Theta'' - \cdots, \\
C^* &= C - c\Theta - c'\Theta' - c''\Theta'' - \cdots, \\
\end{align*}
\]

Giving \(\Theta, \Theta', \Theta'',\) etc., in these equations, the values furnished by the system (I), and for \(A^*, B^*, C^*,\) etc. those given by the system (II), we obtain

\[
\begin{align*}
A &= (A^0 + A^*)(aa) + (B^0 + B^*)(ab) + \cdots, \\
B &= (A^0 + A^*)(ab) + (B^0 + B^*)(bb) + \cdots, \\
C &= (A^0 + A^*)(ac) + (B^0 + B^*)(bc) + \cdots, \\
\end{align*}
\]

and it follows from this that the correlatives of the conditioning equations (12) are

\[
A = A^0 + A^*, \quad B = B^0 + B^*, \quad C = C^0 + C^*, \quad \ldots,
\]

and then the equations (10), (I) and (III) show that we have

\[
\epsilon = \Theta + \kappa, \quad \epsilon' = \Theta' + \kappa', \quad \epsilon'' = \Theta'' + \kappa'', \quad \ldots
\]

consequently, the complete adjustment has the same value for each unknown whether one calculates it directly, or whether one obtains it indirectly by starting from an incomplete adjustment.

20.

When there are a great many conditioning equations, the determination of the correlative quantities \(A, B, C,\) etc. may require such long calculations that the patience of the computer may be unequal to the task; in this case it can be advantageous to obtain a complete adjustment by the help of a series of approximations based on the theorem of the preceding article. To do this one separates the conditioning equations into two or more groups, and first tries to find an adjustment which will satisfy the equations of the first group. Subsequently, one takes the values modified by this first calculation, and
corrects them further, taking account only of the equations of the second group. This second calculation will give results which, in general, will no longer satisfy the equations of the first group, and it will be necessary, if one has formed only two groups, to return then to the first and satisfy them by means of still further corrections. The observations will then be subjected to a fourth adjustment, in which one considers only the conditions of the second group; and operating thus alternately on the one and the other group of equations, one will make corrections which will necessarily become smaller and smaller. If the groups have been suitably chosen, one will arrive quite rapidly at values which will not be changed by further corrections.

When one forms more than two groups, one must proceed in the same way, the various groups being used successively until the last, after which one returns to the first to use them again in the same order. We do no more than indicate this procedure, the success of which depends greatly on the ability of the computer.

21.

We still have to give the demonstration of the lemma used in section 8. For greater clarity, we shall adopt more suitable notations to illuminate the proof.

Let \( x^0, x', x'', \) etc. be unknowns; suppose that by solving the equations

\[
\begin{align*}
n^{00} x^0 + n^{01} x' + n^{02} x'' + \cdots &= X^0, \\
n^{10} x^0 + n^{11} x' + n^{12} x'' + \cdots &= X', \\
n^{20} x^0 + n^{21} x' + n^{22} x'' + \cdots &= X'', \\
&\vdots
\end{align*}
\]

the following are obtained

\[
\begin{align*}
N^{00} x^0 + N^{01} x' + N^{02} x'' + \cdots &= x^0, \\
N^{10} x^0 + N^{11} x' + N^{12} x'' + \cdots &= x', \\
N^{20} x^0 + N^{21} x' + N^{22} x'' + \cdots &= x'', \\
&\vdots
\end{align*}
\]

By substituting in the first two equations of the second system, the values of \( X^0, X', X'', \) etc., given by the first system, we shall obtain two identities:

\[
\begin{align*}
x^0 &= N^{00} \left( n^{00} x^0 + n^{01} x' + n^{02} x'' + \cdots \right) \\
&\quad + N^{01} \left( n^{10} x^0 + n^{11} x' + n^{12} x'' + \cdots \right) \\
&\quad + N^{02} \left( n^{20} x^0 + n^{21} x' + n^{22} x'' + \cdots \right) \\
&\quad + \cdots \\
&\vdots
\end{align*}
\]

\[
\begin{align*}
x' &= N^{10} \left( n^{00} x^0 + n^{01} x' + n^{02} x'' + \cdots \right) \\
&\quad + N^{11} \left( n^{10} x^0 + n^{11} x' + n^{12} x'' + \cdots \right) \\
&\quad + N^{12} \left( n^{20} x^0 + n^{21} x' + n^{22} x'' + \cdots \right) \\
&\quad + \cdots
\end{align*}
\]
Since these equations are identities, one may substitute any quantities one wishes in place of \(x^0, x', x'',\) etc. Let us put in the first
\[
x^0 = N^{10}, \quad x' = N^{11}, \quad x'' = N^{12}, \ldots,
\]
and in the second
\[
x^0 = N^{00}, \quad x' = N^{01}, \quad x'' = N^{02}, \ldots
\]
By subtracting these two equations term by term, we obtain
\[
N^{10} - N^{01} = (N^{00}N^{11} - N^{10}N^{01})(n^{01} - n^{10})
+ (N^{00}N^{12} - N^{10}N^{02})(n^{02} - n^{20})
+ (N^{00}N^{13} - N^{10}N^{03})(n^{03} - n^{30}) + \cdots
+ (N^{01}N^{12} - N^{11}N^{02})(n^{12} - n^{21})
+ (N^{01}N^{13} - N^{11}N^{03})(n^{13} - n^{31}) + \cdots
+ (N^{02}N^{13} - N^{12}N^{03})(n^{23} - n^{32}) + \cdots
\]
This can be more briefly written in the following way:
\[
N^{10} - N^{01} = \sum (N^{0\alpha}N^{1\beta} - N^{1\alpha}N^{0\beta})(n^{\alpha\beta} - n^{\beta\alpha});
\]
where \(\alpha\) and \(\beta\) denote an arbitrary pair of indices; we conclude from this that the equations
\[
n^{01} = n^{10}, \quad n^{02} = n^{20},
\]
and, generally,
\[
n^{\alpha\beta} = n^{\beta\alpha}, \quad \ldots
\]
necessarily imply
\[
N^{10} = N^{01};
\]
and, since the order of the unknowns is arbitrary, it is clear that under the supposition just made, we have in general
\[
N^{\alpha\beta} = N^{\beta\alpha}.
\]
Since the method set forth in this memoir should above all have useful application to geodetic calculations, we adjoin here several examples based on this part of science.

The conditioning equations which exist among the angles of a system of triangles can, in general, be divided into three categories:

I. The sum of the horizontal angles measured around the same peak and taking in the entire horizon, should be equal to four right angles.

II. The sum of the angles of each triangle can always be considered as known; for, even when the triangle is situated on a curved surface, the excess of the sum of its angles over two right angles can be calculated with such close approximation that it is permissible to consider the result absolutely exact.
III. Finally, one obtains a third type of relations by examining the ratios of the sides in the triangles which make up a closed grid. If, indeed, the triangles are so placed that the second triangle has a side $a$ in common with the first, and a side $b$ in common with the third, if the fourth triangle has two sides $c$ and $d$ in common with the third and fifth triangles respectively, and so on, with a final triangle having a side $k$ in common with the preceding one and a side $l$ in common with the first one of all, then the quotients

$$
\frac{a}{l}, \frac{b}{a}, \frac{c}{b}, \frac{d}{c}, \cdots, \frac{l}{k},
$$

can be calculated by means of the angles which are opposite to them in the triangle which includes the two sides which are compared. Since the product of these fractions is obviously one, one will have a relation between the sines of the various measured angles (diminished by a third of the spherical or spheroidal excess when one is operating on a curved surface). Furthermore, in somewhat complicated networks, it often happens that the equations of the second and third category are related with each other so that in consequence, their number must be reduced. On the other hand, it can happen, but only in rather rare cases, that several new equations will be joined to those of the second category; this is what will take place when the net contains polygons which are not divided into triangles; one can then introduce equations relating to the figures which have more than three sides. On some other occasion we shall reconsider in more detail these various circumstances whose examination at the moment would distract us from our goal. However, we cannot omit making a remark which will be vital to those who wish to make a rigorous application of our theory; we always suppose that the quantities denoted by $\nu, \nu', \nu'', \cdots$ each have been directly observed, or deduced from observations such that their determinations are independent of each other, or at least can be regarded as such. In the most common practice, one observes the angles which one may regard as being the elements of $\nu, \nu', \nu'', \cdots$ themselves. But one must not forget that if the system contains as well triangles whose angles have not been directly observed but have been obtained from those already known by additions or subtractions, the angles of these triangles should not be counted in the number of magnitudes determined by observation, and one must treat them in the calculation as functions of the elements from which they have been deduced. The situation is different if one adopts the method of observation of Struve (Astronomische Nachrichten, II, page 431), which consists in determining all the directions around one peak by measuring them all relative to one single arbitrary direction. The angles thus measured will then be taken for $\nu, \nu', \nu'', \cdots$ and the angles of the triangles will all arise as differences. In this case, the equations of the first category should be suppressed as superfluous for they will be identically satisfied. The procedure which I followed myself in triangulations made during the last few years, differs from the two preceding methods; one can however, as far as the result goes, treat it in the same way as the procedure of Struve, in the sense that at each station, one should regard $\nu, \nu', \nu'', \cdots$ as the angles between the directions from the station and one single arbitrarily chosen line.

We shall give two examples; the first relates to the first method of operation, and the second relates to observations made in accordance with the second method.
23.

The work of Krayenhof *Precis historique des operations trigonometriques faites en Holland* will furnish us the first example. We shall seek to adjust the part of the observations concerned with the terrain included between Harlingen, Sneek, Oldeholtpade, Ballum, Leeuwarden, Dockum, Drachten, Oosterwolde and Groningen. Between these points, nine triangles were formed, which, in the work cited, were numbered 121, 122, 123, 124, 125, 127, 128, 131, 132. The observed angles were the following:

**Triangle 121**

0. Harlingen 50° 58′ 15.238″
1. Leeuwarden 82 47 15.351
2. Ballum 46 14 27.202

**Triangle 122**

3. Harlingen 51° 5′ 39.717″
4. Sneek 70 48 33.445
5. Leeuwarden 58 5 48.707

**Triangle 123**

6. Sneek 49° 30′ 40.051″
7. Drachten 42 52 59.382
8. Leeuwarden 87 36 21.057

**Triangle 124**

9. Sneek 45° 36′ 7.492″
10. Oldeholtpade 67 52 0.048
11. Drachten 66 31 56.513

**Triangle 125**

12. Drachten 53° 55′ 24.745″
13. Oldeholtpade 47 48 52.580
14. Oosterwolde 78 15 42.347

**Triangle 127**

15. Leeuwarden 59° 24′ 0.645″
16. Dockum 76 34 9.021
17. Ballum 44 1 51.040

**Triangle 128**

18. Leeuwarden 72° 6′ 32.043″
19. Drachten 46 53 27.163
20. Dockum 61 0 4.494

**Triangle 131**

21. Dockum 57° 1′ 55.292″
22. Drachten 83 33 14.515
23. Groningen 39 24 52.397
The consideration of these triangles shows that the twenty-seven angles directly furnished by observation have thirteen necessary relations between them, that is, two of the first type, nine of the second, and two of the third. But there is no point writing here all these equations in their finite form, since for the calculation we need only the quantities denoted in the general theory by $A, a, a', a'', \ldots, B, b, b', b'', \ldots$, and for this reason we shall write immediately the equations (13) which display these quantities. In place of $\epsilon, \epsilon', \epsilon'', \ldots$, we shall simply write here $(0)$, $(1)$, $(2)$, etc. In this way, to the two equations of the first type, the following correspond:

$$\begin{align*}
(1) + (5) + (8) + (15) + (18) &= -2.197'' \\
(7) + (11) + (12) + (19) + (22) + (26) &= -0.436''.
\end{align*}$$

Next we find for the spheroidal excesses of the nine triangles:

$$\begin{align*}
1.749''; 1.147''; 1.243''; 1.698''; 0.873''; 1.167''; 1.104''; 2.161''; 1.403''.
\end{align*}$$

Then we have the conditioning equation of the second type

$$\nu^{(0)} + \nu^{(1)} + \nu^{(2)} - 180^\circ 0'1.748'' = 0,$$

and similarly with the others so that we obtain the nine equations following:

$$\begin{align*}
(0) + (1) + (2) &= -3.958'' \\
(3) + (4) + (5) &= +0.722 \\
(6) + (7) + (8) &= -0.753 \\
(9) + (10) + (11) &= +2.355 \\
(12) + (13) + (14) &= -1.201 \\
(15) + (16) + (17) &= -0.461 \\
(18) + (19) + (20) &= +2.596 \\
(21) + (22) + (23) &= +0.043 \\
(24) + (25) + (26) &= -0.616
\end{align*}$$

The conditioning equations of the third type are expressed more easily by means of logarithms; the first is

$$\begin{align*}
\log \sin(\nu^{(0)} - 0.583'') - \log \sin(\nu^{(2)} - 0.583'') - \log \sin(\nu^{(3)} - 0.382'') + \\
\log \sin(\nu^{(4)} - 0.382'') - \log \sin(\nu^{(6)} - 0.414'') + \log \sin(\nu^{(7)} - 0.414'') - \\
\log \sin(\nu^{(16)} - 0.389'') + \log \sin(\nu^{(17)} - 0.389'') - \log \sin(\nu^{(19)} - 0.368'') + \\
\log \sin(\nu^{(20)} - 0.368'') &= 0
\end{align*}$$

It does not seem useful to set down the other in its finite form. Corresponding to these two equations we obtain the following in which the coefficients refer to the seventh
decimal of the common logarithms\(^9\)

\[
17.068(0) - 20.174(2) - 16.993(3) + 7.328(4) - 17.976(6) + 22.672(7) - \\
\]

\[
17.976(6) - 0.880(8) - 20.617(9) + 8.564(10) - 19.082(13) + 4.375(14) + \\
\]

Since no reason leads us to attribute unequal weights to the various observations
we suppose
\[p^0 = p^1 = p^2 = \cdots = 1.\]

Denoting the correlatives of the conditioning equations in the same order in which
these equations have been written by

\[A, B, C, D, E, F, G, H, I, K, L, M, N,\]

we shall determine them by the following equations

\[
-2.197'' = 5A + C + D + E + H + I + 5.917N
\]
\[
-0.436 = 6B + E + F + G + I + K + L + 2.962M
\]
\[
-3.958 = A + 3C - 3.106M
\]
\[
+0.722 = A + 3D - 9.665M
\]
\[
-0.753 = A + B + 3E + 4.696M + 17.096N
\]
\[
+2.355 = B + 3F - 12.053N
\]
\[
-1.201 = B + 3G - 14.707N
\]
\[
-0.461 = A + 3H + 16.752M
\]
\[
-2.596 = A + B + 3I - 8.039M - 4.847N
\]
\[
+0.043 = B + 3K - 11.963N
\]
\[
-0.616 = B + 3L + 30.859N
\]
\[
\]
\[
-8.039I + 2902.27M - 459.33N
\]
\[
+370 = 5.917A + 17.096E - 12.053F - 14.707G - 4.874I
\]
\[
-11.963K + 30.859L - 459.33M + 3385.96N
\]

Solving these we obtain:

\[
A = -0.598 \quad E = -0.447 \quad K = +0.577
\]
\[
B = -0.255 \quad F = +1.351 \quad L = -1.351
\]
\[
C = -1.234 \quad G = +0.271 \quad M = -0.109792
\]
\[
D = +0.086 \quad H = +0.659 \quad N = +0.119681
\]
\[
I = +1.050
\]

\(^9\)These coefficients have all been multiplied, after differentiation by \(10''\), and divided by \(206264.8 = \frac{180}{60.60} = \frac{\pi}{180}\), to convert the errors to seconds.— J.B.
Finally, the most likely errors are given by the formulas

\[(0) = C + 17.068M\]
\[(1) = A + C\]
\[(2) = C - 20.174M\]
\[(3) = D - 16.993M\]

and we obtain the following numerical values; to which we add for purpose of comparison, the adjustments used by Krayenhof:

\[
\begin{array}{c|c|c|c}
\text{Krayenhof} & \text{Krayenhof} \\
(0) & -3.108 & -2.009 & (14) +0.795 \\
(1) & -1.832 & +0.116 & (15) +0.061 \\
(2) & +0.981 & -1.982 & (16) +1.211 \\
(3) & +1.952 & +1.722 & (17) -1.732 \\
(4) & -0.719 & +2.848 & (18) +2.126 \\
(5) & -0.512 & -3.848 & (19) +2.959 \\
(6) & +3.648 & -0.137 & (20) +1.628 \\
(7) & -3.221 & +1.000 & (21) +2.211 \\
(8) & -1.180 & -1.614 & (22) +0.322 \\
(9) & -1.116 & 0.0 & (23) -2.489 \\
(10) & +2.376 & +5.928 & (24) -1.709 \\
(11) & +1.096 & -3.570 & (25) +2.701 \\
(12) & +0.016 & +2.414 & (26) -1.606 \\
(13) & -2.013 & -6.014 & \\
\end{array}
\]

The sum of the squares of our corrections is 97.8845; the standard error as obtained from the 27 observed angles, is consequently

\[
\frac{97.8845}{13} = 2.7440''.
\]

The sum of the squares of Krayenhof’s corrections is 341.4201.

The triangles of the triangulation of Hanover, whose vertices were placed at Falkenberg, Breithorn, Hauselberg, Wulfsode and Wilsede will furnish our second example.

The following directions were observed

At the Station of Falkenberg

<table>
<thead>
<tr>
<th>Station</th>
<th>Angle (°')</th>
<th>Distance (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilsede</td>
<td>187° 47'</td>
<td>30.311''</td>
</tr>
<tr>
<td>Wulfsode</td>
<td>225 9</td>
<td>39.676</td>
</tr>
<tr>
<td>Hauselberg</td>
<td>266 13</td>
<td>56.239</td>
</tr>
<tr>
<td>Breithorn</td>
<td>274 14</td>
<td>43.634</td>
</tr>
</tbody>
</table>
At the Station of Breithorn

4. Falkenberg 94° 33' 40.755"
5. Hauselberg 122 51  23.054
6. Wilsede  150 18  35.100

At the Station of Hauselberg

7. Falkenberg 86° 29'  6.872"
8. Wilsede  154 37   9.624
9. Wulfsode 189 2    56.376
10. Breithorn 302 47   37.732

At the Station of Wulfsode

11. Hauselberg 9° 5' 36.593"
12. Falkenberg 45 27 33.556
13. Wilsede  118 44 13.159

At the Station of Wilsede

14. Falkenberg 7° 51' 1.027"
15. Wulfsode  298 29  49.519
16. Breithorn 330 3   7.392
17. Hauselberg 334 25  26.746

From these observations we can form seven triangles:

<table>
<thead>
<tr>
<th>Triangle</th>
<th>Falkenberg</th>
<th>Breithorn</th>
<th>Hauselberg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Triangle I</td>
<td>8° 0'</td>
<td>28 17</td>
<td>143 41</td>
</tr>
<tr>
<td>Triangle II</td>
<td>86° 27'</td>
<td>55 44</td>
<td>37 47</td>
</tr>
<tr>
<td>Triangle III</td>
<td>41° 4'</td>
<td>102 33</td>
<td>36 21</td>
</tr>
<tr>
<td>Triangle IV</td>
<td>78° 26'</td>
<td>68 8</td>
<td>35 25</td>
</tr>
<tr>
<td>Triangle V</td>
<td>37° 22'</td>
<td>73 16</td>
<td>69 21</td>
</tr>
<tr>
<td>Triangle VI</td>
<td>27° 27'</td>
<td>148 10</td>
<td>4 22</td>
</tr>
</tbody>
</table>
the spheroidal excess of the various triangles: I... I
... The side joining Wilsede to Wulfsode is 28877
excess of the seven triangles, and for that it is necessary to know the length of one side.
... to set down any of the first type); to obtain them we must find first the spheroidal
indicated above, labeled with the same indices, the angles of the first triangle will be
IV
... and consequently the first conditioning equation is

\[ \nu^{(3)} - \nu^{(2)} + \nu^{(5)} - \nu^{(4)} + 360 + \nu^{(7)} - \nu^{(10)} \]

and consequently the first conditioning equation is

\[-\nu^{(2)} + \nu^{(3)} - \nu^{(4)} + \nu^{(5)} + \nu^{(7)} - \nu^{(10)} + 179^{\circ} 59'59.798'' = 0.\]

The six remaining triangles give six analogous equations, but a little consideration will show that these equations are not independent; indeed, the second is identical with the sum of the first, the fourth and the sixth; the sum of the third and the fifth is identical with that of the fourth and the seventh; for this reason we shall omit the second and the fifth. In place of the remaining equations in finite form, we shall write here the corresponding equations (13), using the notation \( \nu^{(0)}, \nu^{(1)}, \nu^{(2)}, \nu^{(3)} \) etc. the angles which determine the directions indicated above, labeled with the same indices, the angles of the first triangle will be

\[ \nu^{(3)} - \nu^{(2)}, \quad \nu^{(5)} - \nu^{(4)}, \quad 360 + \nu^{(7)} - \nu^{(10)} \]

We have here seven conditioning equations of the second type (there is no need to set down any of the first type); to obtain them we must find first the spheroidal excess of the seven triangles, and for that it is necessary to know the length of one side. The side joining Wilsede to Wulfsode is 28877.94 meters. From this one obtains, for the spheroidal excess of the various triangles: I... 0.202''; II... 2.442''; III... 1.257''; IV... 1.919''; V... 1.957''; VI... 0.321''; VII... 1.295''.

If we denote by \( \nu^{(0)}, \nu^{(1)}, \nu^{(2)}, \nu^{(3)} \) etc. the angles which determine the directions, labeled with the same indices, the angles of the first triangle will be

\[ \nu^{(3)} - \nu^{(2)}, \quad \nu^{(5)} - \nu^{(4)}, \quad 360 + \nu^{(7)} - \nu^{(10)} \]

and consequently the first conditioning equation is

\[-\nu^{(2)} + \nu^{(3)} - \nu^{(4)} + \nu^{(5)} + \nu^{(7)} - \nu^{(10)} + 179^{\circ} 59'59.798'' = 0.\]

The six remaining triangles give six analogous equations, but a little consideration will show that these equations are not independent; indeed, the second is identical with the sum of the first, the fourth and the sixth; the sum of the third and the fifth is identical with that of the fourth and the seventh; for this reason we shall omit the second and the fifth. In place of the remaining equations in finite form, we shall write here the corresponding equations (13), using the notation \( \nu^{(0)}, \nu^{(1)}, \nu^{(2)}, \nu^{(3)} \) etc. of \( \epsilon, \epsilon', \epsilon'' \), etc.

\[ -1.368'' = - (2) + (3) - (4) + (5) + (7) - (10), \]
\[ +1.773'' = - (1) + (2) - (7) + (9) - (11) + (12), \]
\[ +1.042'' = - (0) + (2) - (7) + (8) + (14) - (17), \]
\[ -0.813'' = - (5) + (6) - (8) + (10) - (16) + (17), \]
\[ -0.750'' = - (8) + (9) - (11) + (13) - (15) + (17). \]

From the triangles of the system one can obtain eight equations of the third type, and for that it is allowable to combine three of the four triangles I, II, IV, VI, or of the triangles III, IV, V, VII; however, brief consideration shows that it suffices to consider two of them belonging to the two systems of triangles respectively and that these will include all the others.

Thus we shall have as sixth and seventh conditioning equations,

\[ \log \sin(\nu^{(3)} - \nu^{(2)} - 0.067'') - \log \sin(\nu^{(5)} - \nu^{(4)} - 0.067'') \]
\[ + \log \sin(\nu^{(14)} - \nu^{(17)} - 0.640'') - \log \sin(\nu^{(2)} - \nu^{(0)} - 0.640'') \]
\[ + \log \sin(\nu^{(16)} - \nu^{(5)} - 0.107'') - \log \sin(\nu^{(17)} - \nu^{(16)} - 0.107'') = 0 \]
\[ \log \sin(\nu^{(2)} - \nu^{(1)} - 0.419'') - \log \sin(\nu^{(12)} - \nu^{(11)} - 0.419'') \]
\[ + \log \sin(\nu^{(14)} - \nu^{(17)} - 0.640'') - \log \sin(\nu^{(2)} - \nu^{(0)} - 0.640'') \]
\[ + \log \sin(\nu^{(13)} - \nu^{(11)} - 0.432'') - \log \sin(\nu^{(17)} - \nu^{(15)} - 0.432'') = 0, \]

Triangle VII

<table>
<thead>
<tr>
<th>Hauselberg</th>
<th>34° 25'</th>
<th>46.752''</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wulfsode</td>
<td>109</td>
<td>38</td>
</tr>
<tr>
<td>Wilsede</td>
<td>35</td>
<td>55</td>
</tr>
</tbody>
</table>

202 meters. From this one obtains, for the spheroidal excess of the various triangles: I... 0.202''; II... 2.442''; III... 1.257''; IV... 1.919''; V... 1.957''; VI... 0.321''; VII... 1.295''.
to which correspond the equations
\[ +25 = +4.31(0) - 153.88(2) + 149.57(3) + 39.11(4) - 79.64(5) \\
+ 40.53(6) + 31.90(14) + 275.39(16) - 307.29(17), \]
\[ -3 = +4.31(0) - 24.16(1) + 19.85(2) + 36.11(11) - 28.59(12) \\
- 7.52(13) + 31.90(14) + 29.06(15) - 60.96(17). \]

If we attribute the same precision to the various directions, supposing \( p^{(0)} = p^{(1)} = p^{(2)} = \cdots = 1 \), the correlatives of the seven conditioning equations being denoted by \( A, B, C, D, E, F, G \), their determination will depend on the following equations.

\[ -1.368 = +6A - 2B - 2C - 2D + 184.72F - 19.85G, \]
\[ +1.773 = -2A + 6B + 2C + 2E - 153.88F - 20.69G, \]
\[ +1.042 = -2A + 2B + 6C - 2D - 2E + 181.00F + 108.40G, \]
\[ -0.813 = -2A - 2C + 6D + 2E - 462.51F - 60.96G, \]
\[ -0.750 = +2B - 2C + 2D + 6E - 307.29F - 133.65G, \]
\[ +25 = 184.72A - 153.88B + 181.00C - 462.51D \\
- 307.29E + 224868F + 16694.1G, \]
\[ -3 = -19.85A - 20.69B + 108.40C - 60.96D \\
- 133.65E + 16694.1F + 8752.39G. \]

Solving, we obtain
\[
\begin{align*}
A &= -0.225 & C &= -0.088 & E &= -0.323 \\
B &= +0.344 & D &= -0.171 & F &= +0.000215915 \\
G &= -0.005474620
\end{align*}
\]

and the most likely errors are given by the formulas
\[
\begin{align*}
(0) &= -C + 4.13F + 4.31G, \\
(1) &= -B - 24.16G, \\
(2) &= -A + B + C - 153.88F + 19.85G, \\
&\vdots
\end{align*}
\]

from which we obtain the following numerical values:
\[
\begin{align*}
(0) &= +0.065'' & (6) &= -0.162'' & (12) &= +0.501'' \\
(1) &= -0.212'' & (7) &= -0.481'' & (13) &= -0.282'' \\
(2) &= +0.339'' & (8) &= +0.406'' & (14) &= -0.256'' \\
(3) &= -0.193'' & (9) &= +0.021'' & (15) &= +0.164'' \\
(4) &= +0.233'' & (10) &= +0.054'' & (16) &= +0.230'' \\
(5) &= -0.071'' & (11) &= -0.219'' & (17) &= -0.139''
\end{align*}
\]
The sum of the squares of these errors is equal to 1.2288; consequently, the standard error resulting from the 18 observed directions is

$$\sqrt{\frac{1.2288}{7}} = 0.4190''$$


In order to give an example of the last part of our theory, let us see with what precision the side Falkenberg-Breithorn is determined from the adjusted observations, in terms of the side Wilsede-Wulfsode. The function $u$, by which it is expressed, is in this case,

$$u = \frac{\sin(\nu^{(18)} - \nu^{(12)} - 0.652''). \sin(\nu^{(14)} - \nu^{(16)} - 0.814'') \sin(\nu^{(6)} - \nu^{(4)} - 0.814'')}{\sin(\nu^{(4)} - \nu^{(0)} - 0.652''). \sin(\nu^{(6)} - \nu^{(4)} - 0.814'')} \times 22877.94\text{ meters}$$

and the value deduced from the corrected observations is 26766.68 meters. On differentiation, this equation gives

$$du = [0.16991\, d\nu^{(0)} - d\nu^{(1)}] + 0.08836(d\nu^{(4)} - d\nu^{(6)})$$

$$- [0.03899\, d\nu^{(12)} - d\nu^{(13)}] + 0.16731(d\nu^{(14)} - d\nu^{(16)})$$

where $d\nu^{(0)}$, $d\nu^{(1)}$, etc. are expressed in seconds; we deduce from this:

<table>
<thead>
<tr>
<th>$a1$</th>
<th>$d1$</th>
<th>$g1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-$0.08836</td>
<td>$+$0.07895</td>
<td>$+$10.99570</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$b1$</th>
<th>$e1$</th>
<th>$h1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$+$0.13092</td>
<td>$+$0.03899</td>
<td>$+$0.13238</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$c1$</th>
<th>$f1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-$0.00260</td>
<td>$-$40.13150</td>
</tr>
</tbody>
</table>

The methods indicated above give the value

$$\frac{1}{P} = 0.08329 \quad \text{or} \quad P = 12.006$$

taking the meter as unit length. We conclude from this that the standard error to be expected in the length of the side Falkenberg-Breithorn is 0.2886$m$ meters ($m$ denoting the standard error to be expected in the observed directions, expressed in seconds), and consequently, if we adopt the value of $m$ as stated above, the standard error to be expected is 0.1209 meters.

Finally, inspection of the system of triangles shows immediately that one could leave the station Hauselberg completely aside, without breaking the net which joins the four others. But it is not allowable to neglect the operations relative to this point, for they certainly contribute to the accuracy of the whole. To show more clearly the increase in precision which results from them, we shall finish by doing the calculation over again, excluding all the results which relate to Hauselberg. Of the 18 directions given above, 8 no longer enter the calculation, and the most likely errors on the remaining ones are

<table>
<thead>
<tr>
<th>$0$</th>
<th>$4$</th>
<th>$13$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$+$0.327''</td>
<td>$+$0.121''</td>
<td>$-$0.206''</td>
</tr>
<tr>
<td>$1$</td>
<td>$6$</td>
<td>$14$</td>
</tr>
<tr>
<td>$-$0.206''</td>
<td>$-$0.121''</td>
<td>$+$0.327''</td>
</tr>
<tr>
<td>$3$</td>
<td>$7$</td>
<td>$15$</td>
</tr>
<tr>
<td>$-$0.121''</td>
<td>$+$0.206''</td>
<td>$+$0.206''</td>
</tr>
<tr>
<td>$16$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$+$0.121''</td>
<td></td>
</tr>
</tbody>
</table>
The value of the side Falkenberg-Breithorn then becomes 26766.63 meters, a result which is not much different from that obtained above. But the calculation of the weight gives

\[ \frac{1}{P} = 0.13082 \quad \text{or} \quad P = 7.644 \]

and the standard error to be expected, in meters, is

\[ 0.36169m = 0.1515. \]

We see that by adding the operations relative to Hauselberg, the weight of the determination of the side Falkenberg-Breithorn is increased in the ratio of 7.644 to 12.006, that is to say in the ratio of 1 to 1.571.
NOTE I
Exposition of the Method of Least Squares
(Extract from *Theoria Motus Corporum celestium.*)

1.

Let us now take up a more general application which is one of the most fruitful in every application of calculation to natural phenomena. Let \( V, V', V'', \) etc. be \( \mu \) functions of the \( \nu \) unknowns \( p, q, r, s, \) etc. and let us suppose that direct observations have given the values

\[
V = M, \quad V' = M', \quad V'' = M''; \quad \ldots
\]

for these functions. In general, the calculation of the \( \nu \) unknowns will constitute a problem which is indeterminate, determinate, or over-determinate, according to whether one has

\[
\mu < \nu, \quad \mu = \nu, \quad \text{or} \quad \mu > \nu. \tag{10}
\]

We shall deal here only with the last case, in which it is obviously impossible to obtain an exact representation of all the observations, unless these observations are completely unaffected by error. But since this never happens in nature, one must consider as possible every system of values of the unknowns \( p, q, r, s, \) etc. for which the values for the functions \( V - M, V' - M', V'' - M'', \) which result from them do not exceed the limits of the errors which one may commit in the observations; one should not, however, consider all these possible systems as having the same degree of probability.

To begin with, let us suppose that in all the observations the state of things is such that there is no reason to consider any one of them as more exact than another, that is to say that one should consider equal errors in each of them as equally likely. The probability that an error \( \Delta \) will be made in one of the observations will be a function of \( \Delta \) which we shall denote \( \phi(\Delta) \). Although this function cannot be precisely determined, one can at least assert that it ought to be a maximum for \( \Delta = 0 \), should have in the majority of cases the same value for values of \( \Delta \) which are equal and of opposite sign, and finally, should vanish for values of \( \Delta \) greater than or equal to the maximum possible error; thus, properly speaking, \( \phi(\Delta) \) ought to be a discontinuous function, and if for ease of calculation we permit ourselves to substitute for it an analytic function, the latter must be chosen in such a way that it tends rapidly to zero outside of an interval of values of \( \Delta \) which includes zero, and that outside these limits one may consider it as zero. Now the probability that the error will fall between \( \Delta \) and a quantity \( \Delta + d\Delta \) which differs from it infinitesimally, will be expressed by \( \phi(\Delta)d\Delta \), and consequently, the probability that the error is included between \( D \) and \( D' \) is given by

\[
\int_D^{D'} \phi(\Delta)d\Delta
\]

\( ^{10}\)If, in the third case, \( \mu + 1 - \nu \) of the functions \( V, V', V'', \) etc. could be regarded as functions of all the others, the problem would become overdeterminate relative to these functions, but indeterminate relative to \( p, q, r, s, \) etc. Once could not then be able to deduce the values of the latter, even if the values of the functions \( V, V', V'', \) etc. were known absolutely exactly; we shall, however, exclude this special case from our investigations. —G.
This integral, taken between the greatest negative value of $\Delta$ and its greatest positive value, or more generally from $\Delta = -\infty$ to $\Delta = \infty$, must necessarily be equal to 1. Thus one has

$$\int_{-\infty}^{\infty} \phi(\Delta) d\Delta = 1.$$ 

Let us suppose then that we have given a set of values for the quantities $p, q, r, s$, etc.; the probability that observation will give the value $M$ for $V$, will be expressed by $\phi(M - V)$, after one has substituted into $V$ the values of $p, q, r, s$, etc.; similarly, $\phi(M' - V')$, $\phi(M'' - V'')$, etc. will express the probabilities that the observations give the values $M', M'', M'''$, etc. to the functions $V', V'', V'''$, etc. For this reason, in so far as one can consider all the observations as events which are independent of each other, the product

$$\phi(M - V) \phi(M' - V') \phi(M'' - V'') \cdots = \Omega,$$

will express the probability that all these values will result simultaneously from the observations.

2.

Just as, before any observations, the assumption of arbitrary values for the unknowns results in a definite probability for a set of values of the functions $V, V', V''$, etc., so, after definite values of these functions have been determined by observation, there results for each set of values of the unknowns a definite probability; for it is clear that one should consider those sets of values the most probable which give the greatest probability to the observed event. This probability may be evaluated by the following theorem.

If, on the assumption of a certain hypothesis $H$, the probability of a definite event $E$ is $h$, but on the assumption of another hypothesis $H'$, exclusive of the first and having the same a priori probability, the probability of the same event is $h'$; then after the event $E$ has taken place, the probability that $H$ is the true hypothesis is to the probability that $H'$ is the true hypothesis is as $h$ to $h'$.

To show this and to point out all the circumstances which may be involved, namely, that the hypothesis $H$ or $H'$ or some other may occur, and that an event $E$ or some other event may occur, let us draw up a table of the different possible cases which we shall regard as equally likely a priori (that is when it is unknown whether the event $E$ has occurred or not.) These cases can be thus distributed:

<table>
<thead>
<tr>
<th>Number of cases</th>
<th>hypothesis belonging to these cases</th>
<th>event which occurs in these cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m$</td>
<td>$H$</td>
<td>$E$</td>
</tr>
<tr>
<td>$n$</td>
<td>$H$</td>
<td>different from $E$</td>
</tr>
<tr>
<td>$m'$</td>
<td>$H'$</td>
<td>$E$</td>
</tr>
<tr>
<td>$n'$</td>
<td>$H'$</td>
<td>different from $E$</td>
</tr>
<tr>
<td>$m''$</td>
<td>different from $H$ and $H'$</td>
<td>$E$</td>
</tr>
<tr>
<td>$n''$</td>
<td>different from $H$ and $H'$</td>
<td>different from $E$</td>
</tr>
</tbody>
</table>
According to this one has

\[ h = \frac{m}{m+n}, \quad h' = \frac{m'}{m'+n'} \]

Now, before the event happens, the probability of the hypothesis \( H \) was

\[ \frac{m + n}{m + n + m' + n' + m'' + n''} \]

After the event, which excludes \( n + n' + n'' \) cases among those which are possible, this probability will be

\[ \frac{m}{m + m' + m''} \]

Similarly the probabilities for the hypothesis \( H' \) before and after the event happens are

\[ \frac{m + n}{m + n + m' + n' + m'' + n''} \quad \text{and} \quad \frac{m}{m + m' + m''} \]

respectively; but, since one has supposed that the hypotheses \( H \) and \( H' \) had the same probability before the event, one has

\[ m + n = m' + n' \]

and from this the truth of the theorem follows immediately.

If one supposes now that for determining the unknowns one has only the observations

\[ V = M, \quad V' = M', \quad V'' = M'', \ldots, \]

and that all the sets of values for the unknowns were equally likely before the observations were made, it is clear that the probability of a certain set of values after these observations will be proportional to \( \Omega \). That is to say that \( \lambda \Omega \, dp \, dq \, dr \ldots \) will express the probability that the values of the unknowns are contained in the infinitesimal intervals \( p \) to \( p + dp \), \( q \) to \( q + dq \), \( r \) to \( r + dr \), \ldots, respectively, where \( \lambda \) represents a quantity independent of \( p, q, r, s, \) etc.; and one will clearly have

\[ \frac{1}{\lambda} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \ldots \Omega \, dp \, dq \, dr \ldots \]

From this it naturally follows that the most probable set of values of \( p, q, r, s, \) etc. will correspond to the maximum of \( \Omega \), and will be obtained from the \( \nu \) equations

\[ \frac{d\Omega}{dp} = 0, \quad \frac{d\Omega}{dq} = 0, \quad \frac{d\Omega}{dr} = 0, \ldots; \]

if one puts

\[ V - M = \nu, \quad V' - M' = \nu', \quad V'' - M'' = \nu'', \ldots \quad \text{and} \quad \phi'(\Delta') = \frac{1}{\phi\Delta} \cdot \frac{d\phi}{d\Delta} \]
these equations take the following form:

\[
\frac{d\nu}{dp} \phi' (\nu) + \frac{d\nu'}{dp} \phi' (\nu') + \frac{d\nu''}{dp} \phi' (\nu'') + \cdots = 0, \\
\frac{d\nu}{dq} \phi' (\nu) + \frac{d\nu'}{dq} \phi' (\nu') + \frac{d\nu''}{dq} \phi' (\nu'') + \cdots = 0, \\
\vdots
\]

From this it follows that one could obtain a completely determined solution of the problem by solving these equations provided that the nature of the function \( \phi' \) were known. But since this function cannot be defined \textit{a priori}, let us attack the question from another point of view and seek a function tacitly accepted as basic in virtue of a simple principle which is generally admitted. Now it is usual to consider as an axiom the hypothesis that if a quantity has been obtained by several direct observations made with the same care in similar circumstances, the arithmetic mean of the observed values will be the most likely value of this quantity, if not absolutely exactly, at least with a very good approximation, so that the safest thing to do is always to take it as the result. Hence if one puts

\[ V = V' = V'' = \cdots = p, \]

and

\[ p = \frac{M + M' + M'' + \cdots}{\mu}, \]

one should have in general

\[ \phi'(M - p) + \phi'(M' - p) + \phi'(M'' - p) + \cdots = 0 \]

for every positive integral value of \( \mu \). Then if we set

\[ M' = M'' = \cdots = M - \mu N, \]

we have in general

\[ \phi'[(\mu - 1)N] = (1 - \mu)\phi'(-N), \]

from which it is easily deduced that \( \frac{\phi' (\Delta)}{\Delta} \) must in general be a constant \( k \). Thus one has

\[ \log \phi(\Delta) = \frac{1}{2} k \Delta^2 + \text{const.} = \frac{1}{2} k \Delta^2 + \log \kappa \]

and from this

\[ \phi(\Delta) = \kappa e^{\frac{1}{2} k \Delta^2}. \]

One easily sees that the constant \( k \) must be negative in order for \( \Omega \) to have a maximum; therefore let us put

\[ \frac{1}{2} k = -h^2, \]

and since, according to an elegant theorem of Laplace, one has

\[ \int_{-\infty}^{\infty} e^{-h^2 \Delta^2} d\Delta = \sqrt{\frac{\pi}{h}}, \]
our function will become

$$\phi(\Delta) = \frac{h}{\sqrt{\pi}} e^{-h^2 \Delta^2}$$

4.

The function which we have just found cannot express the probability of errors completely rigorously, since the possible errors are always bounded between certain limits, and the probabilities of greater errors must always be zero, while our function always has a non-zero value. However, this defect, which would be present in any other analytic function, has no importance in applications, because the value of a function decreases so rapidly as soon as $h\Delta$ reaches any considerable value, that one is completely safe in considering it as practically equivalent to zero. Moreover, the nature of the question never permits one to determine the limits of the errors with absolute rigor.

Finally, the constant $h$ can be regarded as serving to measure the precision of the observations. Indeed, if the probability of the errors present in one system of observations is expressed by

$$\frac{h}{\sqrt{\pi}} e^{-h^2 \Delta^2},$$

and in another system of observations more or less exact than the first by

$$\frac{h'}{\sqrt{\pi}} e^{-h'^2 \Delta^2},$$

the probability that in one observation of the first system the error will be included between the limits $-\delta$ and $+\delta$ will be given by

$$\int_{-\delta}^{\delta} \frac{h}{\sqrt{\pi}} e^{-h^2 \Delta^2} d\Delta,$$

and similarly the probability that the error of an observation of the second system will be included between the limits $-\delta'$ and $+\delta'$ will be given by

$$\int_{-\delta'}^{\delta'} \frac{h'}{\sqrt{\pi}} e^{-h'^2 \Delta^2} d\Delta;$$

and these integrals are obviously equal when one has

$$h\delta = h'\delta'$$

If, for instance, one has

$$h' = 2h,$$

an error of two units in the first system will be committed as easily as an error of one unit in the second, so that the latter observations have a degree of precision twice as great, to employ a commonly used expression.
5.
Let us now examine some consequences of this law. It is clearly necessary that for the product
\[ \Omega = h^\mu \cdot \pi ^{\frac{1}{2}} \mu \cdot e^{-h^2(\nu^2 + \nu'^2 + \nu''^2 + \cdots)} \]
to become a maximum, the sum
\[ \nu^2 + \nu'^2 + \nu''^2 + \cdots \]
must become a minimum. Hence the set of values of the unknowns \( p, q, r, s, \) etc.
which is most likely, corresponds to the case where the squares of the differences be-
tween the observed values and the calculated values of the quantities \( V, V', V'', \) etc.
give the smallest sum possible, provided that all the observations are assumed to be
equally precise.

This principle, which is extremely useful in all the applications of mathematics to
natural science, should be considered as an axiom, on the same basis as the principle
which makes us take the arithmetic mean of the observed values of a single quantity as
the most likely value for this quantity.

The principle extends without difficulty to the case of observations of an unequal
precision. For if the precisions of the observations by which one has found
\[ V = M, \quad V' = M', \quad V'' = M'', \quad \ldots, \]
are represented by \( h, h', h'', \) etc. respectively, that is to say if one supposes that errors
proportional to the reciprocals of these quantities are equally likely to be committed,
it is clear that this comes to the same thing as if from observations of equal precision
(with \( h = 1 \)) the values of functions \( hV, h'V', h''V'', \) etc. had been found equal to
\( hM, h'M', h''M', \) etc.; and this is why the most probable values for the set of quanti-
ties \( p, q, r, s, \) etc. will be that for which the sum
\[ h^2\nu^2 + h'^2\nu'^2 + h''^2\nu''^2 + \cdots, \]
is a minimum. That is to say where the sum of the squares of the differences between the
observed and calculated values, multiplied respectively by the squares of the numbers
which express the degree of precision, becomes a minimum. Consequently, it is not
even necessary for the functions \( V, V', V'', \) etc. to refer to homogeneous quantities,
but they can represent heterogeneous quantities (for example seconds of arc and time),
provided that one can estimate the ratio of the errors which, in each of these quantities,
are equally likely to be committed.

6.
The principle expounded in the preceding section recommends itself also by the fact
that it reduces this numerical calculation of the unknowns to a very rapid algorithm
when the functions \( V, V', V'', \) etc. are linear. Let us suppose
\[ V - M = \nu = -m + ap + bq + cr + ds + \cdots, \]
\[ V' - M' = \nu' = -m' + a'p + b'q + c'r + d's + \cdots, \]
\[ V'' - M'' = \nu'' = -m'' + a''p + b''q + c''r + d''s + \cdots, \]
\[ \vdots \]
Setting
\[ a\nu + a'\nu' + a''\nu'' + \cdots = P, \]
\[ b\nu + b'\nu' + b''\nu'' + \cdots = Q, \]
\[ c\nu + c'\nu' + c''\nu'' + \cdots = R, \]
\[ d\nu + d'\nu' + d''\nu'' + \cdots = S, \]
which are the equations of section 3 which determine the values of the unknowns will be
\[ P = 0, \quad Q = 0, \quad S = 0, \quad \ldots, \]
if we suppose the observations to be equally good. We can reduce all other cases to this one as we showed in the preceding section. Thus one has as many linear equations as there are unknowns, and can solve them by the usual method.

Let us now see whether this solution is always possible or whether it may give indeterminate values or be impossible. As a consequence of the theory of elimination it follows that the second or third case will occur if after omitting one of the equations
\[ P = 0, \quad Q = 0, \quad R = 0, \quad \ldots, \]
one can deduce from the remaining equations an equation identical or contradictory to that which has been omitted, or, what comes to the same thing, if one can determine a linear function
\[ \alpha P + \beta Q + \gamma R + \cdots, \]
which is either identically zero or does not contain any of the unknowns. Thus suppose one had
\[ \alpha P + \beta Q + \gamma R + \cdots = \kappa \]
one also has the identity
\[ (\nu + m)\nu + (\nu' + m')\nu' + (\nu'' + m'')\nu'' + \cdots = pP + qQ + rR + sS + \cdots \]
Suppose that on setting
\[ p = \alpha x, \quad q = \beta x, \quad r = \gamma x, \quad \ldots, \]
the functions \( \nu, \nu', \nu'', \) etc. become
\[ -m + \lambda x, \quad -m' + \lambda' x, \quad -m'' + \lambda'' x, \quad \ldots, \]
respectively, and one obtains the identity
\[ (\lambda^2 + \lambda'^2 + \lambda''^2 + \cdots)x^2 - (\lambda m + \lambda' m' + \lambda'' m'' + \cdots)x = \kappa x; \]
consequently
\[ \lambda^2 + \lambda'^2 + \lambda''^2 + \cdots = 0 \text{ and } \kappa + \lambda m + \lambda' m' + \lambda'' m'' + \cdots = 0, \]
from which it follows that

\[ \lambda = 0, \quad \lambda' = 0, \quad \lambda'' = 0, \quad \ldots, \]

and consequently

\[ \kappa = 0; \]

that is to say that the functions \( V, V', V'' \), etc. cannot change when \( p, q, r, s, \) etc. take on arbitrary increments proportional to the numbers \( \alpha, \beta, \gamma, \delta, \) etc. Such a situation, in which the determination of the unknown would not be possible even if one gave the true values of the functions, does not form a part of our subject, as we declared above.

Finally, one can easily reduce all cases to that in which the functions \( V, V', V'' \), etc. are linear. Let us denote by \( \pi, \chi, \rho, \sigma \), etc. approximate values for the unknowns \( p, q, r, s, \) etc. (which we obtain by using \( \nu \) of the \( \mu \) equations

\[ V = M, \quad V' = M', \quad V'' = M'', \quad \ldots \]

and let us set

\[ p = \pi + p', \quad q = \chi + q', \quad r = \rho + r', \quad s = \sigma + s', \quad \ldots; \]

it is clear that these new unknowns will be so small that their squares and products will be negligible, and that the equations will become linear as a consequence of the indicated substitutions. If it should happen that at the end of the calculation one finds, contrary to expectation, that the values of \( p', q', r', s', \) etc. which one obtains are too considerable, and it seems unsafe to neglect their squares and products, one can resolve this awkwardness by repeating the same operation (but taking as values for \( \pi, \chi, \rho, \sigma, \) etc. the corrected values of \( p, q, r, s, \) etc.).

7.

So long as one has only a single unknown \( p \), for the determination of which one has found that the functions,

\[ ap + n, \quad a'p + n', \quad a''p + n'', \quad \ldots, \]

have the respective values

\[ M, \quad M', \quad M'', \quad \ldots, \]

obtained from equally precise observations, the most likely value of \( p \) is

\[ A = \frac{am + a'm' + a''m'' + \cdots}{a^2 + a'^2 + a''^2 + \cdots} \]

where we have set

\[ m = M - n, \quad m' = M' - n', \quad m'' = M'' - n'', \quad \ldots \]
To estimate the degree of precision which should be attributed to this value, let us suppose that the probability of an error \( \Delta \) being committed in the observations is expressed by
\[
\frac{h}{\sqrt{\pi}} e^{-h^2 \Delta^2};
\]
it follows that the probability that the true value of \( p \) is \( A + p' \) is proportional to the function
\[
e^{-h^2 [(a^p - m)^2 + (a' \Delta - m')^2 + (a'' \Delta - m'')^2 + \cdots]}
\]
when we substitute
\[
p = A + p'.
\]
The exponent of this function can be reduced to the form
\[
-h^2 (a^2 + a'^2 + a''^2 + \cdots) (p^2 - 2pA + B),
\]
where \( B \) denotes a quantity independent of \( p \); consequently the function will be proportional to
\[
e^{-h^2 (a^2 + a'^2 + a''^2 + \cdots) p'^2}.
\]
One sees that the degree of precision which must be attributed to the value of \( A \) is the same as if this value had been found by a direct observation whose precision was to the precision of the original observations as
\[
h \sqrt{a^2 + a'^2 + a''^2 + \cdots} \text{ is to } h,
\]
or as
\[
\sqrt{a^2 + a'^2 + a''^2 + \cdots} \text{ is to } 1.
\]

Before investigating in the case of several unknowns the degree of precision to be attributed to each of them, it is necessary to study more closely the function
\[
\nu^2 + \nu'^2 + \nu''^2 + \cdots,
\]
which we shall denote by \( W \).

I. Let us put
\[
\frac{1}{2} \frac{dW}{dp} = p' = \lambda + \alpha p + \beta q + \gamma r + \delta s + \cdots
\]
and
\[
W - \frac{p'^2}{\alpha} = W';
\]
it clearly follows that
\[
p' = p;
\]
since one has
\[
\frac{dW'}{dp} = \frac{dW}{dp} - \frac{2p'}{\alpha} \frac{dp'}{dp} = 0,
\]
one sees that the function $W''$ will be independent of $p$. The coefficient 

$$= a^2 + a'^2 + a''^2 + \cdots,$$

will obviously always be a positive quantity.

II. Similarly let us set

$$\frac{1}{2} \frac{dW'}{dq} = q' = \lambda' + \beta' q + \gamma' r + \delta' s + \cdots$$

and

$$W' - \frac{q'^2}{\beta'} = W'';$$

one has

$$q' = \frac{1}{2} \frac{dW}{dq} - \frac{p'}{\alpha} \cdot \frac{dp'}{dq} = Q - \frac{\beta}{\alpha} p'$$

and

$$\frac{dW''}{dq} = 0.$$

Hence the function $W''$ is independent of both $p$ and $q$. These circumstances would not occur if it were possible to have

$$\beta' = 0.$$

But it is clear that $W'$ is obtained from

$$\nu^2 + \nu'^2 + \nu''^2 + \cdots,$$

by substituting into $\nu, \nu', \nu'', \text{etc.}$ the value of $p$ obtained from the equation

$$p' = 0;$$

hence $\beta'$ will be the sum of the coefficients of $q^2$ in $\nu^2, \nu'^2, \nu''^2, \text{etc.}$ after this substitution. But these coefficients are all squares and cannot all simultaneously vanish, unless it is the case which we exclude from our discussion, in which the unknowns are undetermined; thus $\delta'$ must be positive.

III. Finally, if one sets

$$\frac{1}{2} \frac{dW''}{dr} = r' = \lambda'' + \gamma'' r + \delta'' s + \cdots$$

and

$$W'' - \frac{r'^2}{\gamma''} = W''';$$

one will have

$$r' = R - \frac{\gamma}{\alpha} p' - \frac{\gamma'}{\beta'} q'.$$
and \( W^{'''\prime} \) will be independent of \( p \), of \( q \), and of \( r \). One can prove as above that the coefficient \( \gamma^{''} \) must be positive. One sees easily, indeed, that \( \gamma^{''} \) is the sum of the coefficients of \( r^2 \) in \( \nu^2, \nu'^2, \nu''^2 \), etc., after the quantities \( p \) and \( q \) have been eliminated from \( \nu^2, \nu'^2, \nu''^2 \), etc. by means of the equations
\[
p' = 0, \quad q' = 0.
\]

IV. Similarly, setting
\[
\frac{1}{2} \frac{dW^{'''}}{ds} = s' = \lambda^{'''\prime} + \delta^{'''2} s + \cdots, \quad W^{iv} = W^{'''\prime} - \frac{s'^2}{\delta^{'''\prime}}
\]
one has
\[
s' = S - \frac{\delta}{\alpha} p' - \frac{\delta'}{\beta} q' - \frac{\delta^{''\prime}}{\gamma^{''} r'}
\]
where \( W^{iv} \) is independent of \( p, q, r \) and \( s \) and \( \delta^{'''\prime} \) is a positive quantity.

V. If there are still more unknowns, one continues in the same way and finally obtains
\[
W = -\frac{1}{\alpha} p'^2 + \frac{1}{\beta} q'^2 + \frac{1}{\gamma^{''}} r'^2 + \frac{1}{\delta^{'''\prime}} s'^2 + \cdots + \text{const.}
\]
an expression in which \( \alpha, \beta', \gamma'', \delta''' \), etc. denote positive quantities.

VI. We have already seen that the probability for a set of values of \( p, q, r, s \), etc. is proportional to the function \( e^{hW} \); consequently, if the value of \( p \) remains undetermined, the probability for a given set of values for \( q, r, s \), etc. will be proportional to the integral
\[
\int_{-\infty}^{\infty} e^{-h^2W} dp
\]
which, according to the theorem of Laplace, is equal to
\[
h^{-1} \alpha^{-\frac{1}{2}} \pi^\frac{1}{2} e^{-h^2(\frac{1}{\alpha} p'^2)}
\]
and this probability would be proportional to the function
\[
e^{-h^2W}.
\]
Similarly, if one considers \( q \) also to be undetermined, the probability for a set of values for \( r, s \), etc. will be proportional to
\[
\int_{-\infty}^{\infty} e^{-h^2W'} dq,
\]
that is to say proportional to
\[
h^{-1} \beta^{-\frac{1}{2}} \pi^\frac{1}{2} e^{-h^2(\frac{1}{\beta} q'^2)}
\]
and, consequently, proportional to \( e^{-h^2W''} \). Similarly, if \( r \) also is considered to be undetermined, the probability of a given set of values for \( s \), etc. will be proportional to
\( e^{-h^2W'''} \), and so on. Let us suppose that there are only four unknowns; the conclusions are the same in the general case. The most likely value of \( s \) will be

\[
-\frac{\lambda'''}{\delta'''}
\]

and the probability that it differs by an amount \( \sigma \) from the true value will be proportional to

\[
e^{-hh''''\sigma^2};
\]

from this we conclude that \( \sqrt{\delta''''} \) measures the precision of this determination relative to the precision of the original observations.

9.

By the method of the preceding paragraph a certain degree of precision has been assigned to the single unknown which was left to the last in the process of elimination. To avoid this inconvenience, we shall now calculate \( \delta' \) in another way.

By solving the equations

\[
\begin{align*}
P &= p' \\
Q &= q' + \frac{\beta}{\alpha}p' \\
R &= r' + \frac{\gamma}{\beta}q' + \frac{\gamma'}{\alpha}p' \\
S &= s' + \frac{\delta''}{\gamma''}r'' + \frac{\delta'}{\beta''}q' + \frac{\delta}{\alpha}p'
\end{align*}
\]

to find \( p', q', r', s', \) one obtains

\[
\begin{align*}
p' &= P \\
q' &= Q + AP \\
r' &= R + B'Q + A'P \\
s' &= S + C''R + B''Q + A''P
\end{align*}
\]

so that the quantities \( A, A', A'', B', B'', C'' \) are determined. Then one has (restricting the number of unknowns to four)

\[
s = \frac{\lambda'''}{\delta'''} + \frac{A''}{\delta''}P + \frac{B''}{\delta''}Q + \frac{C''}{\delta''}R + \frac{1}{\delta''}S
\]

from which one obtains the following consequences. The values of the unknowns \( p, q, r, s, \) etc. which are to be derived from the equations

\[
P = 0, \quad Q = 0, \quad R = 0, \quad S = 0, \quad \ldots,
\]
are obviously expressed by linear functions of \( P, Q, R, S, \) etc., namely
\[
\begin{align*}
p &= L + AP + BQ + CR + DS + \cdots \\
q &= L' + A'P + B'Q + C'R + D'S + \cdots \\
r &= L'' + A''P + B''Q + C''R + D''S + \cdots \\
s &= L''' + A'''P + B'''Q + C'''R + D'''S + \cdots \\
\vdots
\end{align*}
\]

When this is done, the most likely values of the unknowns are \( L, L', L'', \) etc. respectively. The degrees of precision which should be attributed to these determinations are respectively
\[
\frac{1}{\sqrt{A}}, \frac{1}{\sqrt{B'}}, \frac{1}{\sqrt{C''}}, \frac{1}{\sqrt{D'''}}, \ldots,
\]

taking the precision of the original observations to be unity, for what we have said above concerning the unknown \( s \) (for which \( D''' \) corresponds to \( \frac{1}{\sigma'''^2} \)) applies to the other unknowns by a simple permutation.

10.

To illustrate the preceding discussion by an example, let us suppose that from observations which should be assumed to have equal precision, one has found
\[
\begin{align*}
p - q + 2r &= 3, \\
3p + 2q + 5r &= 5, \\
4p + q + 4r &= 21,
\end{align*}

but that by an observation to which a precision of one half should be assigned, one has found
\[
-2p + 6q + 6r = 28.
\]

In place of this last we shall substitute
\[
-p + 3q + 3r = 14,
\]

which we shall suppose to arise from an observation of the same precision as the first three. From this one obtains
\[
\begin{align*}
P &= 27p + 6q - 88 \\
Q &= 6p + 15q + r - 77 \\
R &= q + 54r - 107
\end{align*}
\]

and solving
\[
\begin{align*}
19899p &= 49154 + 809P - 32Q + 6R \\
737q &= 2617 - 12P + 54Q - R \\
6633r &= 12707 + 2P - 9Q + 123R
\end{align*}
\]
The most likely values for the unknowns are therefore

\[ p = 2.470, \quad q = 3.551, \quad r = 1.916, \]

with the degree of precision equal to

\[ \sqrt{\frac{19899}{809}} = 4.96, \quad \sqrt{\frac{737}{54}} = 3.69, \quad \sqrt{\frac{2211}{41}} = 7.34, \]

respectively.

The subject which we have treated up to now would give rise to elegant investigations for which we shall not stop, so as not to wander from our main object. For the same reason, we reserve for another occasion the exposition of artifices permitting the reduction of the calculation to a speedier algorithm. Let me, however, add one observation.

When the number of functions or equations involved is considerable, the calculation becomes particularly tedious because of the circumstance that the coefficients by which one should multiply the original equations to obtain \( P, Q, R, S, \) etc. are almost always complicated decimal fractions. Unless one considers it important in this case to calculate these products with the greatest care with the help of logarithms, it will most often suffice to substitute simple numbers for them which are not much different. Noticeable errors can result from this only when the precision of the unknowns becomes less than the precision of the original observations.

Finally, the principle according to which the sum of the squares of the differences between the observed and calculated quantities should be a minimum, can be established without reference to the calculus of probabilities, as follows.

When the number of unknowns is equal to the number of observations, one can determine the former so that they agree with the latter. But when the first number is the smaller of the two, one cannot obtain absolute agreement if the observations do not possess an absolute precision. Thus it is necessary in this case to establish the most satisfactory agreement, that is to say to proceed so that the differences are made as small as possible. But this idea itself has something vague about it. Indeed, although a set of values of the unknowns should without doubt be preferred to another for which all these differences were respectively greater, the choice between two sets for one of which the agreement is more satisfactory for some of the observations but less satisfactory for others, is to some degree arbitrary, and one can obviously propose several principles under which the first condition would be fulfilled. Using \( \Delta, \Delta', \Delta'', \) etc. to denote the differences between the calculations and the observations, one will satisfy this condition not only if \( \Delta^2 + \Delta'^2 + \Delta''^2 + \cdots, \) etc. becomes a minimum (which is our principle) but also if \( \Delta^4 + \Delta'^4 + \Delta''^4 + \cdots, \) or \( \Delta^6 + \Delta'^6 + \Delta''^6 + \cdots, \) or any sum of even powers in general, becomes a minimum. But of all these principles ours is the simplest, and all the others lead us into extremely complicated calculations. Finally,
this principle, which we have used since 1795, has been given recently by Legendre in his *Nouvelles méthodes pour la détermination des orbites des comètes*, Paris, 1806; one will find in this work several consequences which the desire for brevity has caused us to omit.

If the exponent of the even power which we were just speaking of, were infinite, we would be led to the system in which the maximum errors were less than in any other system.

For the solution of linear equations where there are more equations than unknowns, Laplace uses another principle, proposed originally by Boscovich, namely, that of making the sum of the absolute values of the differences a minimum. One can easily show that the set of values of the unknowns found by this principle alone, must necessarily\(^{11}\) satisfy as many equations, taken among those proposed, as there are unknowns, so that the other equations are used only to decide the choice which one should make.

For instance, if the equation \(V = M\) is one of those which are not satisfied, the set of values found by the principle in question would not be changed if in place of \(M\) one had observed another value of \(M\) such that \(n\) being the calculated value, the differences \(M - n\) and \(N - n\) were of the same sign. Finally, Laplace modifies this principle slightly by adding to it the new condition that the sum of the differences, taken with their signs, should be zero. From this it follows that the number of equations which is satisfied is one less than the number of the unknowns; but the remark which we have just made continues to apply when there are only two unknowns.

\(^{11}\text{Except for some special cases where there is indetermination.}\)
NOTE II
Application of the Method of Least Squares
to the Elements of the Planet Pallas

1.

In volume I of the Göttingen Nachrichten, Gauss gave the application of his method to
the correction of the elements of the planet Pallas. Since the illustrious mathematician
developed the algorithm outlined more briefly in his great work Theoria Motus Corpo-
rum coelestium (see the preceding Note) on this example, we felt we should translate
here this part of his Memoir. Since the first part requires an extensive knowledge of the
theory of planetary motion, we shall not reproduce it, and we take as starting point the
twelve equations which the six elements of the orbit ought to satisfy.

Denoting these corrections by
dL, dZ, dπ, dφ, dΩ, di

the equations obtained by Gauss are the following

100
\[
\begin{align*}
0 &= -183.93'' + 0.79363 \, dL + 143.66 \, dZ + 0.39493 \, d\pi \\
&\quad + 0.95920 \, d\phi - 0.18886 \, d\Omega + 0.17387 \, di; \\
0 &= -6.81'' - 0.02658 \, dL + 46.71 \, dZ + 0.02658 \, d\pi \\
&\quad - 0.20858 \, d\phi + 0.15946 \, d\Omega + 1.25782 \, di; \\
0 &= -0.06'' + 0.58880 \, dL + 358.12 \, dZ + 0.26208 \, d\pi \\
&\quad - 0.85234 \, d\phi + 0.14912 \, d\Omega + 0.17775 \, di; \\
0 &= -3.09'' + 0.01318 \, dL + 28.39 \, dZ - 0.01318 \, d\pi \\
&\quad - 0.07861 \, d\phi + 0.91704 \, d\Omega + 0.54365 \, di; \\
0 &= -0.02'' + 1.73436 \, dL + 1846.17 \, dZ - 0.54603 \, d\pi \\
&\quad - 2.05662 \, d\phi - 0.18833 \, d\Omega - 0.17445 \, di; \\
0 &= -8.98'' - 0.12606 \, dL - 227.42 \, dZ + 0.12606 \, d\pi \\
&\quad - 0.38939 \, d\phi + 0.17176 \, d\Omega - 1.35441 \, di; \\
0 &= -2.31'' + 0.99584 \, dL + 1579.03 \, dZ + 0.06456 \, d\pi \\
&\quad + 1.99545 \, d\phi - 0.06040 \, d\Omega - 0.33750 \, di; \\
0 &= +2.47'' - 0.08089 \, dL - 67.22 \, dZ + 0.08089 \, d\pi \\
&\quad - 0.09970 \, d\phi - 0.46359 \, d\Omega + 1.22803 \, di; \\
0 &= 0.01'' + 0.65311 \, dL + 1329.09 \, dZ + 0.38994 \, d\pi \\
&\quad - 0.08439 \, d\phi - 0.04305 \, d\Omega + 0.34268 \, di; \\
0 &= +38.12'' - 0.00218 \, dL + 38.47 \, dZ + 0.00218 \, d\pi \\
&\quad - 0.18710 \, d\phi + 0.47301 \, d\Omega - 1.14371 \, di; \\
0 &= -317.73'' + 0.69957 \, dL + 1719.32 \, dZ + 0.12913 \, d\pi \\
&\quad - 1.38787 \, d\phi + 0.17130 \, d\Omega - 0.08360 \, di; \\
0 &= +117.97'' - 0.01315 \, dL - 43.84 \, dZ + 0.01315 \, d\pi \\
&\quad + 0.02929 \, d\phi + 1.02138 \, d\Omega - 0.27187 \, di;
\end{align*}
\]
From the nature of the observations which furnished the tenth of these equations, it is judged to inspire too little confidence to make use of it, and the six unknowns will be determined only from the other eleven.

The following explanations are literally translated from Gauss’s Memoir.

J. Bertrand

2.

Since it is impossible for us to satisfy the eleven proposed equations exactly, that is to say, to make all the right hand sides zero, we shall seek to make the sum of their squares as small as possible.

One sees easily that if one considers the linear functions

\[ n + ap + bq + cr + ds + \ldots = w, \]
\[ n' + a'p + b'q + c'r + d's + \ldots = w', \]
\[ n'' + a''p + b''q + c''r + d''s + \ldots = w'', \]
\[ \vdots \]

the equations which must be solved in order to make

\[ \Omega = w^2 + w'^2 + w''^2 + \ldots \]

a minimum, are

\[ aw + a'w' + a''w'' + \ldots = 0, \]
\[ bw + b'w' + b''w'' + \ldots = 0, \]
\[ cw + c'w' + c''w'' + \ldots = 0, \]
\[ \vdots \]

or, defining the following abbreviations,

\[ an + a'n' + a''n'' + \ldots = (an), \]
\[ a^2 + a'^2 + a''^2 + \ldots = (aa), \]
\[ ab + a'b' + a''b'' + \ldots = (ab), \]
\[ \vdots \]
\[ b^2 + b'^2 + b''^2 + \ldots = (bb), \]
\[ bc + b'c' + b''c'' + \ldots = (bc), \]
\[ \vdots \]

\[ p, q, r, s, \text{ etc.} \] should be determined by the following equations

\[ (an) + (aa)p + (ab)q + (ac)r + \ldots = 0, \]
\[ (bn) + (ab)p + (bb)q + (bc)r + \ldots = 0, \]
\[ (cn) + (ac)p + (bc)q + (cc)r + \ldots = 0, \]
\[ \vdots \]
The process of solution, very tedious when the number of unknowns is considerable, can be simplified notably in the following way. Suppose that besides the coefficients \((an), (aa), \text{etc.}\) (of which the number is \(\frac{1}{2}(i^2 + 3i)\), if the number of unknowns is \(i\)) one has calculated the sum
\[N^2 + n^2 + n'^2 + \cdots = (nn);\]
one sees easily that one has
\[\Omega = (nn) + 2(an)p + 2(bn)q + 2(cn)r + \cdots + (aa)p^2 + 2(ab)pq + 2(ac)pr + \cdots + (bb)q^2 + 2(bc)qr + 2(bd)qs + \cdots + (cc)r^2 + 2(cd)rs + \cdots\]
and, denoting
\[(an) + (aa)p + (ab)q + \cdots\]
by \(A\), all the terms of \(\frac{A^2}{(aa)}\) which contain the factor \(p\), are found in the expression \(\Omega\), and consequently
\[\Omega - \frac{A^2}{(aa)}\]
is a function independent of \(p\). This is why, setting
\[(nn) - \frac{(an)^2}{(aa)} = (nn, 1),\]
\[(bn) - \frac{(an)(bn)}{(aa)} = (bn, 1),\]
\[(cn) - \frac{(an)(cn)}{(aa)} = (cn, 1),\]
\[\vdots\]
\[(bb) - \frac{(ab)^2}{(aa)} = (bb, 1),\]
\[(bc) - \frac{(ab)(ac)}{(aa)} = (bc, 1),\]
\[(bd) - \frac{(ab)(ad)}{(aa)} = (bd, 1),\]
\[\vdots\]
one has
\[-\frac{A^2}{(aa)} = (an, 1) + 2(bn, 1)q + 2(cn, 1)r + 2(dn, 1)s + \cdots + (bb, 1)q^2 + 2(bc, 1)qr + 2(bd, 1)qs + \cdots + (cc, 1)r^2 + 2(cd, 1)rs + \cdots + \cdots\]
We shall denote this function by $\Omega'$. 

Similarly setting 

$$(bn, 1) + (bb, 1)q + (bc, 1)r + \cdots = B,$$

the difference 

$$\Omega' - \frac{B^2}{(bb, 1)}$$

will be independent of $q$; we shall represent it by $\Omega''$. 

Similarly, setting 

$$(nn, 1) - (bn, 1)^2 = (nn, 2),$$

$$(cn, 1) - \frac{(bn, 1)(bc, 1)}{(bb, 1)} = (cn, 2),$$

$$(cc, 1) - \frac{(bc, 1)^2}{(bb, 1)} = (cc, 2),$$

$\vdots$

and 

$$(cn, 2) + (cc, 2)r + (cd, 2)s + \cdots = C,$$

the difference 

$$\Omega'' - \frac{C^2}{(cc, 2)}$$

will be a function independent of $r$. Continuing in this way we shall form a sequence of expressions $\Omega, \Omega', \Omega'', \ldots$, of which the last will be independent of the various unknowns and is denoted by $(nn, \mu)$, if $\mu$ denotes the number of these unknowns; then we shall have 

$$\Omega = \frac{A^2}{(aa)} + \frac{B^2}{(bb, 1)} + \frac{C^2}{(cc, 2)} + \frac{D^2}{(dd, 3)} + \cdots + (nn, \mu).$$

One can easily prove that since $\Omega$ is a sum of squares 

$$w^2 + w'^2 + w''^2 + \cdots,$$

and cannot become negative, the denominators $(aa), (bb, 1), (cc, 2), \ldots$ are all positive. For brevity we omit the details of the proof.) Accordingly, the minimum value of $\Omega$ obviously corresponds to the values of the unknowns for which 

$$A = 0, \quad B = 0, \quad C = 0, \quad \ldots,$$

and, by starting to solve the system with the last equation, which contains only one of them, one finds the values of $p, q, r, s, \ldots$ without having to carry out any elimination. The method gives at the same time the minimum value of $\Omega$, which is $(nn, \mu)$. 
Let us apply these principles to our example, in which \( p, q, r, s, \) etc. are replaced by \( dL, dZ, d\pi, d\phi, d\Omega, di. \) By careful calculation I have found

\[
\begin{align*}
(nn) &= 148848 \quad (ac) = -0.09344 \quad (cc) = +0.71917 \\
(an) &= -371.09 \quad (ad) = -2.28516 \quad (cd) = +1.13382 \\
(bn) &= -580104 \quad (ae) = -0.34664 \quad (ce) = +0.06400 \\
(cn) &= -133.45 \quad (af) = -0.18194 \quad (cf) = +0.26341 \\
(dn) &= +268.53 \quad (bb) = +1083425 \quad (dd) = +12.00340 \\
(en) &= +94.26 \quad (bc) = -49.06 \quad (de) = -0.37137 \\
(fn) &= -31.81 \quad (bd) = -3229.77 \quad (df) = -0.11762 \\
(aa) &= +5.91569 \quad (be) = -198.64 \quad (ee) = +2.28215 \\
(ab) &= +7203.91 \quad (bf) = -143.05 \quad (ef) = -0.36136 \\
(ef) &= +5.62456
\end{align*}
\]

from which one obtains

\[
\begin{align*}
(nn, 1) &= +125569 \quad (bc, 1) = +62.13 \quad (ef, 1) = +0.26054 \\
(bn, 1) &= -138534 \quad (bd, 1) = -510.58 \quad (dd, 1) = +11.12064 \\
(cn, 1) &= -119.31 \quad (be, 1) = +213.84 \quad (de, 1) = -0.50528 \\
(dn, 1) &= -125.18 \quad (bf, 1) = +73.45 \quad (df, 1) = -0.18790 \\
(en, 1) &= +72.52 \quad (cc, 1) = +0.71769 \quad (ee, 1) = +2.26185 \\
(fn, 1) &= -43.22 \quad (cd, 1) = +1.09773 \quad (ef, 1) = -0.37202 \\
(bb, 1) &= +2458225 \quad (ce, 1) = -0.05852 \quad (ff, 1) = +5.61905
\end{align*}
\]

Similarly,

\[
\begin{align*}
(nn, 2) &= +117763 \quad (cc, 2) = +0.71612 \quad (de, 2) = -0.46088 \\
(bn, 2) &= -115.81 \quad (cd, 2) = +1.11063 \quad (df, 2) = -0.17265 \\
(cn, 2) &= -153.95 \quad (ce, 2) = -0.06392 \quad (ee, 2) = +2.24325 \\
(dn, 2) &= +84.57 \quad (ef, 2) = +0.25868 \quad (ef, 2) = -0.37841 \\
(en, 2) &= -39.03 \quad (dd, 2) = +11.01466 \quad (ff, 2) = +5.61686
\end{align*}
\]

From which:

\[
\begin{align*}
(nn, 3) &= +99034 \quad (dd, 3) = +9.29213 \quad (ee, 3) = +2.23754 \\
(bn, 3) &= +25.66 \quad (de, 3) = -0.36175 \quad (ef, 3) = -0.35532 \\
(cn, 3) &= +74.23 \quad (df, 3) = -0.57384 \quad (ff, 3) = +5.52342 \\
(fn, 3) &= +2.75
\end{align*}
\]

Similarly,

\[
\begin{align*}
(nn, 4) &= +98963 \quad (fn, 4) = +4.33 \quad (ef, 4) = -0.37766 \\
(en, 4) &= +75.23 \quad (ee, 4) = +2.22346 \quad (ff, 4) = +5.48798
\end{align*}
\]

From which:

\[
\begin{align*}
(nn, 5) &= +96418 \quad (fn, 5) = +17.11 \quad (ff, 5) = +5.42383
\end{align*}
\]
From which we finally obtain

\[(nn, 6) = +96364\]

Thus we have the following six equations

\[
\begin{align*}
0 &= +17.11'' + 5.42383 \, di \\
0 &= +75.23'' + 2.22346 \, d\Omega - 0.37766 \, di \\
0 &= +25.66'' + 9.29213 \, d\phi - 0.36175 \, d\Omega - 0.57384 \, di \\
0 &= -115.81'' + 0.71612 \, d\pi + 1.11063 \, d\phi - 0.06392 \, d\Omega + 0.25868 \, di \\
0 &= -13854'' + 2458225 \, dZ + 62.13 \, d\pi - 0.51058 \, d\phi + 213.84 \, d\Omega + 73.45 \, di \\
0 &= -371.09'' + 5.91569 \, dL + 7203.91 \, dZ - 0.00344 \, d\pi - 2.20516 \, d\phi \\
&\quad - 0.34664 \, d\Omega - 0.18194 \, di
\end{align*}
\]

from which one obtains;

\[
\begin{align*}
\, di &= -3.15'' \\
\, d\Omega &= -34.37'' \\
\, d\phi &= -4.29'' \\
\, d\pi &= +166.44'' \\
\, dZ &= +0.054335'' \\
\, dL &= -3.06''
\end{align*}
\]

These are the corrections which should be applied to the elements originally found for the planet.
In order to establish the principles of the method of least squares, we have assumed that the probability of an error of observations $\Delta$ is given by the formula

$$\frac{h}{\sqrt{\pi}} e^{-h^2 \Delta^2}$$

where $\pi$ represents the half circumference of the unit circle, $e$ the base of hyperbolic logarithms, and $h$ a constant which one may consider (Theoria Motus Corporum coelestium, article 178) as measuring the precision of the observations. It is not necessary to know the value of $h$ in order to determine, by means of the method of least squares, the most likely values of the quantities on which the observations depend, since the ratio of the precision of these results to the precision of the observations is also independent of $h$.

Nonetheless, since knowledge of the quantity $h$ is interesting and instructive, I shall show how it may be determined from the observations.

Let us begin with some remarks which will clarify the question, and denote by $\theta(t)$ the definite integral

$$\int_0^t \frac{2e^{-t^2}}{\sqrt{\pi}} dt$$

Some specific values of this function will give an idea of its behavior.

<table>
<thead>
<tr>
<th>$t$</th>
<th>$\rho$</th>
<th>$\theta(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4769363</td>
<td>$\rho$</td>
<td>0.5</td>
</tr>
<tr>
<td>0.5951161</td>
<td>$\rho \times 1.247790$</td>
<td>0.6</td>
</tr>
<tr>
<td>0.7328691</td>
<td>$\rho \times 1.536618$</td>
<td>0.7</td>
</tr>
<tr>
<td>0.9061939</td>
<td>$\rho \times 1.900032$</td>
<td>0.8</td>
</tr>
<tr>
<td>1</td>
<td>$\rho \times 2.096716$</td>
<td>$\theta(t) = 0.8427008$</td>
</tr>
<tr>
<td>1.1630872</td>
<td>$\rho \times 2.348664$</td>
<td>0.9</td>
</tr>
<tr>
<td>1.8213864</td>
<td>$\rho \times 3.818930$</td>
<td>0.99</td>
</tr>
<tr>
<td>2.3276754</td>
<td>$\rho \times 4.880475$</td>
<td>0.999</td>
</tr>
<tr>
<td>2.7510654</td>
<td>$\rho \times 5.768204$</td>
<td>0.9999</td>
</tr>
<tr>
<td>$\infty$</td>
<td>$\infty$</td>
<td>1</td>
</tr>
</tbody>
</table>

The probability that the error of one observation is included in the limits $+\Delta$ and $-\Delta$, or disregarding the sign, that it does not exceed $\Delta$, will be equal to

$$\frac{h}{\sqrt{\pi}} \int_{-\Delta}^{\Delta} e^{-h^2 x^2} dx.$$
it will be equal to twice the integral in question if this is taken between the limits \( x = 0 \) and \( x = \Delta \), and consequently it will be equal to \( \theta(h\Delta) \).

Thus the probability that the error is not less than \( \frac{\theta}{2} \) is equal to \( \frac{1}{2} \), i.e., is equal to the probability of its contrary. Thus I shall call this quantity \( \frac{\theta}{2} \) the probable error, and shall denote it by \( r \).

On the other hand, the probability that the error exceeds \( 2.348664r \) is only one-tenth; the probability that the error exceeds \( 3.818390r \) is only one one-hundredth; and so on.

3.

Let us suppose that the errors actually made in \( m \) observations are \( \alpha, \beta, \gamma \), etc., and let us see what consequences one can derive regarding the values of \( h \) and \( r \).

By making two hypotheses on the exact value of \( h \), and supposing it equal to either \( H \) or \( H' \), the probabilities that the observations will be affected by the errors \( \alpha, \beta, \gamma \), etc. will be, for the two cases, in the ratio of

\[
He^{-H^2\alpha^2} \times He^{-H^2\beta^2} \times He^{-H^2\gamma^2} \times \cdots
\]

to

\[
H'e^{-H'^2\alpha^2} \times H'e^{-H'^2\beta^2} \times H'e^{-H'^2\gamma^2} \times \cdots,
\]

that is to say as

\[
H^m e^{-H^2(\alpha^2 + \beta^2 + \gamma^2 + \cdots)}
\]

is to

\[
H'^m e^{-H'^2(\alpha^2 + \beta^2 + \gamma^2 + \cdots)}
\]

It is clear that the probabilities that \( H \) or \( H' \) be the true values of \( H \) are in the same ratio (Theoria Motus Corporum coelestium, article 176); consequently the probability of an arbitrary value of \( H \) is proportional to

\[
H^m e^{-H^2(\alpha^2 + \beta^2 + \gamma^2 + \cdots)},
\]

and the most likely value of \( h \) is that for which this function becomes a maximum. But one finds by well known rules that \( h \) is then equal to

\[
\sqrt{\frac{m}{2(\alpha^2 + \beta^2 + \gamma^2 + \cdots)}};
\]

hence the most probable value of \( r \) will be

\[
\rho \sqrt{\frac{2(\alpha^2 + \beta^2 + \gamma^2 + \cdots)}{m}}
\]

or

\[
0.6744897 \times \sqrt{\frac{1}{m}(\alpha^2 + \beta^2 + \gamma^2 + \cdots)}
\]

This is a general result, whether \( m \) is large or small.
4.

It is easy to understand that the values found for $h$ and $r$ are less certain the smaller the number $m$.

Let us now find the degree of precision which one should assign to the values of $h$ and $r$ when $m$ is a fairly large number.

Let us denote by $H$ the most probable value of $h$, which we have found to be

$$
H = \sqrt{\frac{m}{2(\alpha^2 + \beta^2 + \gamma^2 + \ldots)}};
$$

and note that the probability that $H$ is the true value of $h$ is to the probability that $(H + \lambda)$ is the true value as

$$
H^m e^{-m} \text{ is to } (H + \lambda)^m e^{-m(H+\lambda)^2},
$$

or as

$$
1 \text{ is to } e^{\frac{\lambda^2 m}{2}} \left(1 - \frac{1}{2} \cdot \frac{\lambda}{H} + \frac{1}{4} \cdot \frac{\lambda^2}{H^2} - \frac{1}{5} \cdot \frac{\lambda^3}{H^3} + \ldots \right)
$$

The second term will bear a substantial ratio to the first only if $\frac{\lambda}{H}$ is a small fraction, and in this case, we may replace the indicated ratio by

$$
1 : e^{-\frac{\lambda^2 m}{2}}
$$

This means that the probability that the true value of $h$ is included between $(H + \lambda)$ and $(H + \lambda + d\lambda)$ is approximately equal to

$$
Ke^{-\frac{\lambda^2 m}{2}} d\lambda,
$$

where $K$ is a constant such that the integral

$$
\int Ke^{-\frac{\lambda^2 m}{2}} d\lambda
$$

taken between the admissible limits of $\lambda$, becomes equal to one.

In the present case because of the large value of $m$, $e^{-\frac{\lambda^2 m}{2}}$ becomes extremely small when $\frac{\lambda}{H}$ is different from a small fraction, and it will be permissible to take the integral from $-\infty$ to $\infty$; thus one obtains

$$
K = \frac{1}{H} \sqrt{\frac{m}{\pi}}.
$$

Consequently, the probability that the true value of $h$ lies between $(H - \lambda)$ and $(H+\lambda)$ will be equal to

$$
\theta \left(\frac{\lambda}{H} \sqrt{m}\right),
$$

and it will be equal to $\frac{1}{2}$ when $\frac{\lambda}{H} \sqrt{m} = \rho$. 
Thus it is an even bet that the true value of \( h \) lies between
\[
H \left( 1 + \frac{\rho}{\sqrt{m}} \right) \quad \text{and} \quad H \left( 1 - \frac{\rho}{\sqrt{m}} \right),
\]
or that the true value of \( r \) lies between
\[
\frac{R}{1 - \frac{\rho}{\sqrt{m}}} \quad \text{and} \quad \frac{R}{1 + \frac{\rho}{\sqrt{m}}},
\]
when \( R \) denotes the most likely value of \( r \) found in the preceding section. These limits may be called the probable limits of the true value of \( h \) and \( r \). It is clear that we may take here as probable limits of \( r \)
\[
R \left( 1 - \frac{\rho}{\sqrt{m}} \right) \quad \text{and} \quad R \left( 1 + \frac{\rho}{\sqrt{m}} \right).
\]

5.

In the preceding discussion, we considered \( \alpha, \beta, \gamma, \) etc., as given quantities, in order to evaluate the probability that the true value of \( h \) or \( r \) would lie between certain limits.

One may consider the question from another point of view, by assuming that the errors of the observations follow a given probability law, one can then evaluate the probability that the sum of the squares of \( m \) errors of observations will fall between certain limits. Laplace has already solved this problem in the case where \( m \) is a very large number, as well as the problem of determining the probability that the sum of \( m \) errors of observations lies between certain limits.

It is easy to generalize this investigation; I shall restrict myself here to indicating the result.

Let us denote by \( \phi(x) \) the probability of an error \( x \) in the observations, so that
\[
\int_{-\infty}^{\infty} \phi x \, dx = 1.
\]

Then let us denote by \( K_n \) the value of the integral
\[
\int_{-\infty}^{\infty} \phi x x^n \, dx.
\]

Finally let
\[
S_n = \alpha^n + \beta^n + \gamma^n + \cdots
\]
where \( \alpha, \beta, \gamma, \) etc. represent \( m \) arbitrary errors of observation; the terms of this sum will always be taken positively, even if \( n \) is odd.

Then \( mK_n \) will be the most likely value of \( S_n \), and the probability that the true value of \( S_n \) falls between the limits \((mK_n - \lambda)\) and \((mK_n + \lambda)\) will be equal to
\[
\frac{\lambda}{\sqrt{2m(K_{2n} - K_n^2)}};
\]
Consequently, the probable limits of $S_n$ will be

$$mK_n - \rho\sqrt{2m(K_{2n} - mK_n^2)}$$

and

$$mK_n + \rho\sqrt{2m(K_{2n} - mK_n^2)}$$

This result applies, in a general way, to any probability distribution for errors. Applying it to the particular case in which

$$\phi x = \frac{h}{\sqrt{\pi}} \cdot e^{-h^2x^2},$$

we find

$$K_n = \frac{\prod \frac{n}{2}(n-1)}{h^n\sqrt{\pi}}$$

where the symbol $\Pi$ is taken to have the same meaning as in *Disquisitiones generales circa seriem infinitam*\textsuperscript{12} (Comm. nov. Soc. Götting, Volume III, note 5, article 28).

Hence

$$K_1 = 1, \quad K_2 = \frac{1}{2h^2}, \quad K_3 = \frac{1}{h^3\sqrt{\pi}},$$

$$K_4 = \frac{1.3}{4h^3}, \quad K_5 = \frac{1.2}{h^5\sqrt{\pi}}, \quad K_6 = \frac{1.35}{8h^5}, \quad K_7 = \frac{1.23}{h^7\sqrt{\pi}},$$

and consequently the most likely value of $S_n$ will be

$$m\frac{\prod \frac{n}{2}(n-1)}{h^n\sqrt{\pi}}$$

and the most likely limits for the true value of $S_n$ will be

$$m\frac{\prod \frac{n}{2}(n-1)}{h^n\sqrt{\pi}} \cdot \left\{1 - \rho\sqrt{\frac{2}{m} \left[\frac{\prod \frac{n}{2}(n-1)\sqrt{\pi}}{\prod \frac{n}{2}(n-1)^2}\right] - 1}\right\}$$

and

$$m\frac{\prod \frac{n}{2}(n-1)}{h^n\sqrt{\pi}} \cdot \left\{1 + \rho\sqrt{\frac{2}{m} \left[\frac{\prod \frac{n}{2}(n-1)\sqrt{\pi}}{\prod \frac{n}{2}(n-1)^2}\right] - 1}\right\}$$

Thus if we put as above

$$\frac{\rho}{h} = r,$$

where $r$ represents the probable error of observation, the most likely value of

$$\rho \sqrt{\frac{S_n\sqrt{\pi}}{m\prod \frac{n}{2}(n-1)}}$$

\textsuperscript{12}$\Pi(n) = \Gamma(n + 1)$ is Gauss’s notation for the factorial function.
will obviously be \( r \); and the probable limits for the value of this quantity will be

\[
\begin{align*}
\rho \sqrt{\frac{2 \pi}{m \Pi^2(n-1)}} \left( 1 - \frac{\rho}{n} \sqrt{\frac{2}{m \left[ \Pi(n - \frac{1}{2}) \sqrt{\pi} \right]} - n} \right)
\end{align*}
\]

and

\[
\begin{align*}
\rho \sqrt{\frac{2 \pi}{m \Pi^2(n-1)}} \left( 1 + \frac{\rho}{n} \sqrt{\frac{2}{m \left[ \Pi(n - \frac{1}{2}) \sqrt{\pi} \right]} - n} \right)
\end{align*}
\]

Hence it is an even bet that \( r \) will lie between the limits

\[
\rho \sqrt{\frac{2 \pi}{m \Pi^2(n-1)}} \left( 1 - \frac{\rho}{n} \sqrt{\frac{2}{m \left[ \Pi(n - \frac{1}{2}) \sqrt{\pi} \right]} - 1} \right)
\]

and

\[
\rho \sqrt{\frac{2 \pi}{m \Pi^2(n-1)}} \left( 1 + \frac{\rho}{n} \sqrt{\frac{2}{m \left[ \Pi(n - \frac{1}{2}) \sqrt{\pi} \right]} - 1} \right)
\]

For \( n = 2 \), the limits will be

\[
\rho \sqrt{\frac{2 \pi}{m (1 - \frac{\rho}{m})}}
\]

and

\[
\rho \sqrt{\frac{2 \pi}{m (1 + \frac{\rho}{m})}}
\]

which agrees perfectly with those which we found in section 4.

In general, if \( n \) is even, one will have the limits

\[
\rho \sqrt{\frac{2 \pi}{m.1.3.5.7.\cdots(n-1)}} \left( 1 - \frac{\rho}{n} \sqrt{\frac{2}{m \left[ \Pi \left( n + 1 \right) \left( n + 3 \right) \cdots (2n - 1) \right]} - 1} \right)
\]

and

\[
\rho \sqrt{\frac{2 \pi}{m.1.3.5.7.\cdots(n-1)}} \left( 1 + \frac{\rho}{n} \sqrt{\frac{2}{m \left[ \Pi \left( n + 1 \right) \left( n + 3 \right) \cdots (2n - 1) \right]} - 1} \right)
\]

and, if \( n \) is odd, the following limits

\[
\rho \sqrt{\frac{2 \pi}{m.1.2.3.\cdots\left( \frac{n-1}{2} \right)}} \left( 1 - \frac{\rho}{n} \sqrt{\frac{1}{m \left[ \Pi \left( n - 1 \right) \pi \right] - 2} \right)
\]

and

\[
\rho \sqrt{\frac{2 \pi}{m.1.2.3.\cdots\left( \frac{n-1}{2} \right)}} \left( 1 + \frac{\rho}{n} \sqrt{\frac{1}{m \left[ \Pi \left( n - 1 \right) \pi \right] - 2} \right)
\]
6.

I add here the numerical values for the most simple cases:

<table>
<thead>
<tr>
<th>Probable values of $r$:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>I. $0.8453473 \times \frac{S_1}{m} \pi \left( 1 \pm \frac{0.5095841}{\sqrt{m}} \right)$</td>
<td></td>
</tr>
<tr>
<td>II. $0.6747897 \times \sqrt{\frac{S_2}{m} \pi} \left( 1 \pm \frac{0.4769363}{\sqrt{m}} \right)$</td>
<td></td>
</tr>
<tr>
<td>III. $0.5771897 \times \sqrt{\frac{S_3}{m} \pi} \left( 1 \pm \frac{0.4971987}{\sqrt{m}} \right)$</td>
<td></td>
</tr>
<tr>
<td>IV. $0.5125017 \times \sqrt{\frac{S_4}{m} \pi} \left( 1 \pm \frac{0.5507186}{\sqrt{m}} \right)$</td>
<td></td>
</tr>
<tr>
<td>V. $0.4655532 \times \sqrt{\frac{S_5}{m} \pi} \left( 1 \pm \frac{0.6355080}{\sqrt{m}} \right)$</td>
<td></td>
</tr>
<tr>
<td>VI. $0.4294972 \times \sqrt{\frac{S_6}{m} \pi} \left( 1 \pm \frac{0.7557764}{\sqrt{m}} \right)$</td>
<td></td>
</tr>
</tbody>
</table>

Thus one sees by this comparison that the second method of determining $r$ is the best; for 100 errors of observation, treated by this formula, will give a result as reliable as

| 114 | observations from formula I |
| 109 | " | III |
| 133 | " | IV |
| 178 | " | V |
| 251 | " | VI |

However formula I presents the advantage of lending itself better to numerical calculations; and since its degree of precision is only slightly inferior to that of formula II, one can always use it, unless one already knows the sum of the squares of the errors or one wishes to know it.

7.

The following procedure is more convenient but much less exact.

Let us arrange the absolute values of the $m$ errors of observation in order of magnitude, and denote by $M$ the center term if their number is odd or the arithmetic mean between the two center terms if their number is even.

One can show (which we shall not do here) that, for a large number of observations, the most likely value of $M$ is $r$, and that the probable limits of $M$ are

\[ r \left( 1 - e^{\sigma^2} \cdot \sqrt{\frac{\pi}{8m}} \right) \quad \text{and} \quad r \left( 1 + e^{\sigma^2} \cdot \sqrt{\frac{\pi}{8m}} \right), \]

or that the probable limits of $r$ are

\[ M \left( 1 - e^{\sigma^2} \cdot \sqrt{\frac{\pi}{8m}} \right) \quad \text{and} \quad M \left( 1 + e^{\sigma^2} \cdot \sqrt{\frac{\pi}{8m}} \right), \]

or, in numbers,

\[ M \left( 1 \pm \frac{0.7520974}{\sqrt{m}} \right) \]
Consequently this procedure is not much less exact than the application of formula VI, and it would be necessary to consider the errors in 249 observations in order to obtain a result as sure as in applying formula II to one hundred errors of observations.

8.

The application of these methods to the errors committed in 48 observations of the right ascension of the Pole Star by Bessel gave

\[
S_1 = 60.46'' \\
S_2 = 110.600'' \\
S_3 = 250.341118''
\]

From this one deduces the most probable values of \( r \):

According to formula I \( 1.065'' \) Probable error \( \pm 0.068'' \)

II \( 1.024'' \) \( \pm 0.070'' \)

III \( 1.001'' \) \( \pm 0.072'' \)

and according to section 7: \( 1.045'' \) \( \pm 0.113'' \)

an agreement of results which one could hardly hope for. Bessel gives 1.067'' and seems, consequently, to have calculated according to formula I.
As you requested, I am sending you the rules concerning the use of the method of least squares in the solution of the following problem:

To determine the position of a point from the horizontal angles observed from this point between other points whose position is known exactly.

This question, which is very elementary, can cause no difficulty to those who have grasped the spirit of the method of least squares. Nonetheless I shall develop the formulas to which this method leads for the benefit of those who may have to deal with the practical question without studying the theory.

Let $a$ and $b$ be the coordinates of one of the given points; we shall suppose that one takes the positive $x$ direction from north to south and the positive $y$ direction from west to east; let $x$ and $y$ be the approximate coordinates of the unknown point and $dx$, $dy$ the corrections, as yet unknown, which should be applied to them. Let us define two quantities $\phi$ and $r$ by the formulas

$$
\tan \phi = \frac{b - y}{a - x}, \quad r = \frac{a - x}{\cos \phi} = \frac{b - y}{\sin \phi}
$$

$\phi$ being taken in such a quadrant that the value of $r$ is positive.

Further, let us put

$$
\alpha = \frac{206265''(b - y)}{r^2}, \quad \beta = \frac{206265''(a - x)}{r^2}
$$

Then for an observer placed at the second point, the azimuth of the first (taking the azimuth of a line parallel to the $x$ axis to be zero) is

$$
\phi + \alpha dx + \beta dy,
$$

where the last two terms are expressed in seconds.

Let $\phi', \alpha', \beta'$ be the quantities corresponding to $\phi$, $\alpha$, $\beta$ relative to the second of the given points, $\phi''$, $\alpha''$, $\beta''$, those which refer to the third, and so on. Let us suppose that for the angular measurements taken at the point whose position is unknown, a theodolite has been used without repetition, with the lens turned successively towards the various known points without the position of the instrument itself being changed. If $h$, $h'$, $h''$ are the observed azimuths, one would have, assuming the observations to be rigorously exact and $dx$, $dy$ exactly known,

(1) \[ \phi - h + \alpha dx + \beta dy = \phi' - h' + \alpha' dx + \beta' dy = \phi'' - h'' + \alpha'' dx + \beta'' dy \cdots \]

Thus if one writes down that three of these differences have the same value, one will find the approximate value for $dx$ and $dy$; if only three points have been observed
there is nothing more to be done; but if the number of points considered is larger, the errors will be best compensated for by taking the average of the various expressions (1), setting the difference of each of them from the average equal to zero, and applying to these equations the method of least squares.

If all the measurements are independent of each other, each one of them furnishes an equation between $dx$ and $dy$, and it is necessary to combine these equations by the method of least squares, taking account, if one wishes, of the unequal precision of the observations.

For example, let $i$ be the angle between the first and the second point, $i'$ the angle between the second and the third, and so on, reckoning always from left to right; one obtains the equations

\[
\phi' - \phi - i + (\alpha' - \alpha)dx + (\beta' - \beta)dy = 0 \\
\phi'' - \phi' - i + (\alpha'' - \alpha')dx + (\beta'' - \beta')dy = 0
\]

If the various measurements have the same weight, one obtains from these equations two normal equations, by adding them after having successively multiplied each one by the coefficient of $dx$ or by the coefficient of $dy$.

If, on the other hand, the measurements of the angles are of unequal exactitude, and, for instance, the first is based on $\mu$ and the second on $\mu'$ repetitions, then it is necessary in the two cases to multiply the equations by $\mu$, $\mu'$, etc., before the addition; subsequently one finds $dx$, $dy$, etc. by elimination between the two normal equations so obtained.

(The preceding rules are only intended for persons to whom the method of least squares is still unknown and for whom it would perhaps be wise to recall that in the multiplications, the signs of $\alpha' - \alpha$, $\beta' - \beta$, etc. must be rigorously preserved. Finally, I remark again that we are considering only the compensation of errors committed in the angles, coordinates of the given points being supposed exact.)

Let us apply the preceding rules to the observations which we made together on the Holken's Bastion, at Copenhagen. I must warn you that the results here cannot be rigorously exact. Since the observed points were very close to the station, an inaccuracy of ten or twenty feet in their position would exert an influence much greater than the errors which are usually to be expected in measuring angles. Thus one should not be surprised that the best adjustment of the angles leaves the difference much larger than those one admits as possible in observations of such a nature. This application should be taken as an example of the procedure to follow in other cases.
Angles measured from the Holkenbastion

<table>
<thead>
<tr>
<th>Location 1</th>
<th>Location 2</th>
<th>Angle (°)</th>
<th>Prime (′)</th>
<th>Double (″)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Friedrichsberg</td>
<td>Petri</td>
<td>70</td>
<td>35</td>
<td>22.8</td>
</tr>
<tr>
<td>Petri</td>
<td>Erlosersturm</td>
<td>104</td>
<td>57</td>
<td>33.0</td>
</tr>
<tr>
<td>Erlosersturm</td>
<td>Friedrichsberg</td>
<td>181</td>
<td>27</td>
<td>5.0</td>
</tr>
<tr>
<td>Friedrichsberg</td>
<td>Frauenthurm</td>
<td>80</td>
<td>37</td>
<td>10.8</td>
</tr>
<tr>
<td>Frauenthurm</td>
<td>Friedrichsturm</td>
<td>101</td>
<td>11</td>
<td>50.8</td>
</tr>
<tr>
<td>Friedrichsturm</td>
<td>Friedrichsberg</td>
<td>178</td>
<td>11</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Coordinates of the various points in Paris feet, the origin being at the Copenhagen Observatory

<table>
<thead>
<tr>
<th>Point</th>
<th>Coordinates (feet)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Petri</td>
<td>487.7 + 1007.1</td>
</tr>
<tr>
<td>Frauenthurm</td>
<td>710.0 + 674.2</td>
</tr>
<tr>
<td>Friedrichsberg</td>
<td>2430.6 + 8335.0</td>
</tr>
<tr>
<td>Elossersturm</td>
<td>2940.0 − 3536.0</td>
</tr>
<tr>
<td>Friedrichsturm</td>
<td>3059.3 − 2231.2</td>
</tr>
</tbody>
</table>

The approximate coordinates of the Bastion are

\[ x = 2836.44 \]
\[ y = 444.33 \]

and thus we find the azimuths

<table>
<thead>
<tr>
<th>Point</th>
<th>Azimuth (°)</th>
<th>Prime (′)</th>
<th>Double (″)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Petri</td>
<td>166</td>
<td>30</td>
<td>42.56</td>
</tr>
<tr>
<td>Frauenthurm</td>
<td>173</td>
<td>33</td>
<td>50.54</td>
</tr>
<tr>
<td>Friedrichsberg</td>
<td>92</td>
<td>56</td>
<td>29.46</td>
</tr>
<tr>
<td>Elossersturm</td>
<td>271</td>
<td>29</td>
<td>25.38</td>
</tr>
<tr>
<td>Friedrichsturm</td>
<td>274</td>
<td>45</td>
<td>41.48</td>
</tr>
</tbody>
</table>

The angle under which one sees the distance from Petri to Friedrichsberg is consequently

\[ 73° 34′ 3.10″ = 6.15 dx + 81.70 dy; \]

and setting it equal to the observed angle, one has

\[ -79.70″ = 6.15 dx + 81.70 dy = 0 \]

Similarly one obtains the following equations

\[ 69.82″ = -71.71 dx - 84.39 dy = 0 \]
\[ 9.08 + 77.86 dx + 2.69 dy = 0 \]
\[ 0.28 - 15.27 dx + 94.44 dy = 0 \]
\[ 0.04 - 87.36 dx - 102.16 dy = 0 \]
\[ -3.42 + 102.63 dx + 7.72 dy = 0 \]
Supposing the observations to be equally precise, one deduces from this the normal equations
\[29640 dx + 14033 dy = 4168''\]
\[14033 dx + 33219 dy = 12383''\]
and consequently
\[dx = -0.05, \quad dy = 0.40;\]
the coordinates of the Bastion are therefore
\[
\begin{align*}
2836.39 \\
444.73.
\end{align*}
\]

The differences between the observed values of the angles and those which one calculates from the results are too big for one to be able to attribute them to errors of observation; they indicate, as we observed, a lack of precision in the determination of the known points.

The coordinates \(x\) and \(y\), taken as first approximations, were deduced directly from the fourth and fifth angle. Although the direct method should be considered an almost exhausted subject, I shall nonetheless indicate, for the sake of completeness, the method which I usually employ in such a case.

Let \(a\) and \(b\) be the coordinates of the first known point; those of the second will be of the form
\[
a + R \cos E, \quad b + R \sin E,
\]
and those of the third
\[
a + R' \cos E', \quad b + R' \sin E',
\]
Let
\[
a + \rho \cos \tau, \quad b + \rho \sin \tau
\]
be the desired coordinates of the point from which one is observing; let \(M\) be the observed angle (always from left to right) between the first and second point, and \(M'\) the angle observed between the first and third (supposing that, if necessary, one has subtracted 180 degrees); let
\[
\frac{R}{\sin M} = n, \quad \frac{R'}{\sin M'} = n',
\]
\[E - M = N, \quad E' - M' = N'.\]
Then one has the two equations
\[\rho = n \sin(\epsilon - N), \quad \rho' = n' \sin(\epsilon - N');\]
which, written in the following way,
\[
n = \frac{\rho}{\sin(\epsilon - N)}, \quad n' = \frac{\rho'}{\sin(\epsilon - N')}.
\]
are solved by the method set forth in *Theoria Motus Corporum coelestium*, page 82.

One of the solutions set forth in this place leads to the following rule. Let $n'$ be larger than, or at least no smaller than $n$, which is obviously permissible, since one may choose the second point arbitrarily; then setting

$$\frac{n}{n'} = \tan \zeta$$

$$\tan \frac{1}{2}(N' - N) = \tan \psi$$

and one will have

$$\epsilon = \frac{1}{2}(N' + N) + \psi$$

Since $\epsilon$ is known, one of the equations, or indeed both of them, will furnish the value of $\rho$.

In our example, if we consider Frauenthrum as the first point, Friedrichsberg as the second and Friedrichsthurm as the third, we have

| $a$     | 710.0 |
| $b$     | 684.2 |
| $E$     | $77^\circ 19'31.92''$ |
| $E'$    | $308^\circ 51'45.77''$ |
| $\log R$ | 3.8944205 |
| $\log R'$ | 3.5733549 |
| $M$     | $99^\circ 22'50.20''$ |
| $M'$    | $101^\circ 11'50.80''$ |
| $N$     | $337^\circ 56'42.72''$ |
| $N'$    | $207^\circ 39'54.97''$ |
| $\log n$ | 3.9002650 |
| $\log n'$ | 3.5817019 |

and, $n'$ being larger than $n$, we shall change the order and put

| $N$     | $207^\circ 39'54.97''$ |
| $N'$    | $337^\circ 56'42.72''$ |
| $\log n$ | 3.5817019 |
| $\log n'$ | 3.9002650 |

from this one obtains

$$\zeta = 19^\circ 39'3.87''$$

$$\psi = 80^\circ 45'31.69''$$

$$\epsilon = 353^\circ 33'50.53''$$

$$\log \rho = 3.3303990,$$

and as coordinates of the Holkensbastion,

2836.441 444.330
NOTE V
On the Chronometric Determination of Longitudes
(Astronomische Nachrichten, Volume V, page 227)

Let \( \theta, \theta', \theta'' \), etc. be the periods \( \pi \) in number, at which a chronometer has determined the differences \( a, a', a'' \), etc., with the times of places whose longitudes are \( x, x', x'' \), etc. \( \theta, \theta', \theta'' \), being supposed reduced to the time of a single place and \( u \) denoting the daily advance of the chronometer; one would have, if the instrument were perfectly regular, the equations

\[
a - \theta u - x = a' - \theta' u - x' = a'' - \theta'' u - x'' = \ldots
\]

In order that these equations suffice for the determination of the unknowns \( x, x', x'', \ldots \), \( u \), it is necessary, for one thing, to consider one of the longitudes as given, and for another, it is necessary that at least two observations have been made in the same place, so that at least two of the unknowns \( x, x', x'' \), etc. are equal to each other. If among these quantities there are only two which are identical, the problem is completely determined, in the contrary case it becomes indeterminate, and one should proceed to satisfy the equations

\[
0 = a - a' + (\theta' - \theta)u - x + x'
\]
\[
0 = a' - a'' + (\theta'' - \theta')u - x' + x''
\]
\[
0 = a'' - a''' + (\theta''' - \theta'')u - x'' + x'''
\]
\[
\vdots
\]

as exactly as possible, for the inevitable imperfections of the chronometer will never permit all of them to be satisfied rigorously. However, one should not assign to these equations equal weight, for the quantities

\[
a - a' + (\theta' - \theta)u - x + x'
\]
\[
a' - a'' + (\theta'' - \theta')u - x' + x''
\]
\[
\vdots
\]

represent the accumulations of all the variations in the motion of the chronometer in the intervals \( \theta' - \theta, \theta'' - \theta' \), etc. and if a good chronometer is involved to which one can truly attribute an average motion without a variation which keeps increasing in one direction, the average value to be expected for such a sum can be considered as proportional to the square root of the elapsed time.

Thus one should, in the application of the method of least squares, consider the preceding equations as having weights inversely proportional to the differences \( \theta' - \theta \), \( \theta'' - \theta' \), etc.

The solution then offers no difficulty, and furnishes the most likely values of \( x, x', x'' \), etc. as well as the weight of each determination.

However I shall add several remarks.
I. If the first and last observation have been made at the same place, the most probable
value of \( u \) is that which results simply from comparison of these extreme observations. The calculation then becomes very simple, for, by virtue of a theorem which is very easy to demonstrate, one may replace \( u \) in the equations by its most likely value, or, what comes to the same thing, one may use this value as if it were exact to correct the observations and to reduce them to those which would be made with a fictitious chronometer whose rate of gain was zero.

II. If one simply attributes to the various equations weights equal to

\[
\frac{1}{\theta' - \bar{\theta}} \quad \frac{1}{\theta'' - \bar{\theta'}} \quad \frac{1}{\theta''' - \bar{\theta''}}
\]

the unit of precision for the weights obtained will be the exactitude of that which one would obtain by the aid of the same chronometer observed only two times, and at one day’s interval; but in order to compare the results obtained by the aid of various chronometers of unequal precision, it is still necessary to introduce into the result a factor which depends on the greater or less perfection of each chronometer used.

To arrive at it I suppose that the expressions

\[
a' - a'' + (\theta' - \theta)u - x + x' \\
a'' - a''' + (\theta'' - \theta')u - x' + x'' \\
\vdots
\]

become \( \lambda, \lambda', \lambda'', \ldots \) respectively, when one substitutes for the unknowns their most probable values. Let

\[
\frac{\lambda^2}{\theta' - \bar{\theta}} + \frac{\lambda'^2}{\theta'' - \bar{\theta'}} + \frac{\lambda''^2}{\theta''' - \bar{\theta''}} + \ldots = S;
\]

if \( \nu \) is the number of unknowns and one puts

\[
m = \sqrt{\frac{S}{n - \lambda - 1}}
\]

the specific factor relating to each chronometer is proportional to \( \frac{1}{m^2} \) or to \( \frac{n - \nu - 1}{S} \) and one can consider \( m \) as the deviation of the average motion which is to be expected during a day.

III. The preceding rules are relevant to a chronometer whose motion is not subject to any noticeable irregularity which increases with time. If this hypothesis were untenable one might assume, when the observations do not include an excessively long period, a variation in the daily gain of the instrument, proportional to the time thus producing an additional unknown.

The equation would then take the following form:

\[
0 = a - a' + (\theta - \theta')u + (\theta'^2 - \bar{\theta}^2)\nu - x + x', \\
\vdots
\]

IV. Concerning the solution of the equations according to the method of least squares,
it is perhaps not unuseful to recollect that one should begin in most cases by calculating an approximate value for the unknowns, and then apply the method to the determination of the small corrections to which the values should be subjected.

It seemed useful to recall this general advice, because many calculators seem to have forgotten it and been led to calculations which were more laborious and perhaps less exact.

I have determined the behavior of the following chronometers

<table>
<thead>
<tr>
<th></th>
<th>h</th>
<th>m</th>
<th>s</th>
<th>Breguet 3056</th>
<th>Barraud 904</th>
<th>Kassel 1252</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greenwich</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30 June</td>
<td>3</td>
<td>22</td>
<td>−8</td>
<td>17.14&quot;</td>
<td>+1' 2.37&quot;</td>
<td></td>
</tr>
<tr>
<td>25 July</td>
<td>4</td>
<td>15</td>
<td>10</td>
<td>44.39</td>
<td>1 32.15</td>
<td>+30' 59.75&quot;</td>
</tr>
<tr>
<td></td>
<td>28</td>
<td></td>
<td></td>
<td>0.69</td>
<td>1 36.96</td>
<td>30 50.07</td>
</tr>
<tr>
<td>2 Aug.</td>
<td>4</td>
<td>11</td>
<td>0.69</td>
<td>13.69</td>
<td>40.24</td>
<td>50 39.69</td>
</tr>
<tr>
<td>17 Aug.</td>
<td>10</td>
<td>12</td>
<td>59.40</td>
<td>2 6.24</td>
<td>29 35.69</td>
<td>49 57.83</td>
</tr>
<tr>
<td>25</td>
<td>7</td>
<td>27</td>
<td>13</td>
<td>47.98</td>
<td>2 15.84</td>
<td>29 10.48</td>
</tr>
<tr>
<td>10 Sep.</td>
<td>7</td>
<td>40</td>
<td>15</td>
<td>24.47</td>
<td>2 40.36</td>
<td></td>
</tr>
<tr>
<td>Helgoland</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 July</td>
<td>3</td>
<td>40</td>
<td>−40</td>
<td>8.00</td>
<td>−30 26.84</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>12</td>
<td>42</td>
<td>2.02</td>
<td>30 3.89</td>
<td>−0 20.34</td>
<td>+16 47.39+18 48.39</td>
</tr>
<tr>
<td>5 Aug.</td>
<td>1</td>
<td>48</td>
<td>43</td>
<td>18.11</td>
<td>29 43.35</td>
<td>1 10.24</td>
</tr>
<tr>
<td>11</td>
<td>13</td>
<td>9</td>
<td>43</td>
<td>35.77</td>
<td>29 33.43</td>
<td>1 32.75</td>
</tr>
<tr>
<td>30</td>
<td>19</td>
<td>30</td>
<td>45</td>
<td>53.08</td>
<td>29 7.96</td>
<td>2 40.67</td>
</tr>
<tr>
<td>6 Sep.</td>
<td>3</td>
<td>6</td>
<td>46</td>
<td>51.56</td>
<td>28 58.94</td>
<td>3 4.55</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>42</td>
<td>46</td>
<td>38.72</td>
<td>28 56.71</td>
<td></td>
</tr>
<tr>
<td>Altona</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Aug.</td>
<td>5</td>
<td>55</td>
<td>−51</td>
<td>38.95</td>
<td>−37 55.76</td>
<td>−9 28.50</td>
</tr>
<tr>
<td>9</td>
<td>12</td>
<td>35</td>
<td>51</td>
<td>57.35</td>
<td>37 50.03</td>
<td>9 38.81</td>
</tr>
<tr>
<td>31</td>
<td>9</td>
<td>57</td>
<td>54</td>
<td>10.33</td>
<td>37 21.30</td>
<td>10 56.68</td>
</tr>
<tr>
<td>4 Sept.</td>
<td>22</td>
<td>12</td>
<td>54</td>
<td>39.16</td>
<td>37 15.21</td>
<td>11 15.36</td>
</tr>
<tr>
<td>Bremen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 Aug.</td>
<td>0</td>
<td>2</td>
<td>−47</td>
<td>50.65</td>
<td>−33 16.49</td>
<td>−5 23.37</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+14 21.86+16 5.83</td>
</tr>
</tbody>
</table>

Let us for example take Breguet’s chronometer 3056. Let zero be the longitude of Helgoland, −x that of Greenwich, y that of Altona. I do not take account here of that of Bremen, since having only one observation for this town, it is impossible to control it. I count the time from the first comparison of the chronometer No. 1 (Greenwich June 30 3h 22m). Substituting for Breguet’s chronometer a fictitious instrument with daily advance zero, we find
In the equations above, the unknowns $x$ and $y$ are separated, which facilitates their determinations; we find for $x$ four determinations:

<table>
<thead>
<tr>
<th>Weight</th>
<th>1889.40''</th>
<th>$\frac{1}{2.6}$ = 0.38</th>
</tr>
</thead>
<tbody>
<tr>
<td>1891.21</td>
<td>$\frac{1}{3.0}$ = 0.33</td>
<td></td>
</tr>
<tr>
<td>1889.78</td>
<td>$\frac{1}{5.9}$ = 0.17</td>
<td></td>
</tr>
<tr>
<td>1891.42</td>
<td>$\frac{1}{3.4}$ = 0.19</td>
<td></td>
</tr>
</tbody>
</table>

from which one obtains

$x = 1890.36''$  1.07

and similarly one finds

$y = 494.12''$  3.83

According to these values, the fictitious chronometer would indicate, in Helgoland time

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>22.4</th>
<th>60.20''</th>
<th>37.1</th>
<th>0.06''</th>
<th>61.6</th>
<th>0.75</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta$</td>
<td>25.0</td>
<td>1949.60</td>
<td>40.4</td>
<td>1.47</td>
<td>62.2</td>
<td>0.67</td>
</tr>
<tr>
<td>$\theta$</td>
<td>28.0</td>
<td>1950.87</td>
<td>42.4</td>
<td>0.75</td>
<td>66.8</td>
<td>0.49</td>
</tr>
<tr>
<td>$\theta$</td>
<td>32.9</td>
<td>1950.29</td>
<td>48.3</td>
<td>0.62</td>
<td>68.0</td>
<td>0.43</td>
</tr>
<tr>
<td>$\theta$</td>
<td>35.9</td>
<td>59.08</td>
<td>56.2</td>
<td>3.14</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

from which one obtains

$S = 6.00$

and the standard error to be expected is

for $x$  0.75'',  for $y$  0.40''

The results furnished by the five chronometers give

<table>
<thead>
<tr>
<th>Standard error to be expected</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breguet $x$ = 1890.36''</td>
<td>0.75</td>
</tr>
<tr>
<td>Kassel</td>
<td>1893.39</td>
</tr>
<tr>
<td>Barraud</td>
<td>1892.32</td>
</tr>
<tr>
<td>1</td>
<td>1892.39</td>
</tr>
<tr>
<td>4</td>
<td>1892.52</td>
</tr>
<tr>
<td>Average $x$ = 1892.35</td>
<td></td>
</tr>
</tbody>
</table>
Similarly one finds for \( y \)

<table>
<thead>
<tr>
<th></th>
<th>Standard error to be expected</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breguet</td>
<td>494.12</td>
<td>0.10</td>
</tr>
<tr>
<td>Kassel</td>
<td>493.89</td>
<td>0.36</td>
</tr>
<tr>
<td>Barraud</td>
<td>493.67</td>
<td>0.21</td>
</tr>
<tr>
<td>1</td>
<td>493.98</td>
<td>0.29</td>
</tr>
<tr>
<td>4</td>
<td>494.16</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The number placed under the heading of weight in the last column, is the reciprocal of the square of the standard error to be expected, taking as unit weight that which corresponds to observations giving a standard error to be expected of \( 1'' \), so that, for Altona, the standard error to be expected is \( \frac{1}{\sqrt{58.01}} = 0.13'' \); but it is preferable to consider the numbers in the last column as indicating merely ratios, and to deduce the absolute precision from the difference between the values of these final results found for \( x \) and \( y \) by means of each chronometer. The precision found in this way will be a little too large, since the determinations of time at Greenwich, at Helgoland and at Altona do not have an absolute precision, so that consequently whatever the number of chronometers, the errors arising from this source will always have some effect in each final result.

One may similarly, in the following way, obtain the longitude of Bremen.

Let this be longitude to the east of Helgoland, the comparison of the Breguet chronometer gives the position of the fictitious chronometer as

\[-164.52'' + z\]

and one deduces from comparison with previous results

\[z = 225.40'' \quad \frac{1}{1.4} = 0.7,\]

and others give

\[z = 224.76 \quad \frac{1}{4.5} = 0.2\]

\[225.24 \quad 0.9\]

The weight 0.9 should be multiplied by \( \frac{10}{6.000} \); the five chronometers give

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Breguet</td>
<td>225.24</td>
<td>1.5</td>
</tr>
<tr>
<td>Kassel</td>
<td>225.84</td>
<td>1.9</td>
</tr>
<tr>
<td>Barraud</td>
<td>225.39</td>
<td>3.6</td>
</tr>
<tr>
<td>1</td>
<td>226.04</td>
<td>2.9</td>
</tr>
<tr>
<td>4</td>
<td>224.86</td>
<td>4.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14.2</td>
</tr>
</tbody>
</table>
The longitude of Bremen, which according to this would be 208.54″ to the west of Altona, is naturally affected by errors in the determination of the time at Bremen, and this difference appears to be too small by several seconds. According to my triangulations, the tower of Ansgarius is 273.51″ of time to the west of Göttingen, and the observatory of Olbers 271.9″.