C · L · A · S · S · I · C · S
IN APPLIED MATHEMATICS

# Theoria Combinationis Observationum Erroribus Minimis Obnoxiae

*Pars Prior • Pars Posterior • Supplementum*

By Carl Friedrich Gauss

11

# Theory of the Combination of Observations Least Subject to Errors

*Part One • Part Two • Supplement*

Translated by G. W. Stewart

# Theoria
# Combinationis Observationum
# Erroribus Minimis Obnoxiae

# Theory of the
# Combination of Observations
# Least Subject to Errors

# Classics in Applied Mathematics

SIAM's Classics in Applied Mathematics series consists of books that were previously allowed to go out of print. These books are republished by SIAM as a professional service because they continue to be important resources for mathematical scientists.



## Classics in Applied Mathematics

Lin, C. C. and Segel, L. A., Mathematics Applied to Deterministic Problems in the Natural Sciences

Belinfante, Johan G. F. and Kolman, Bernard, A Survey of Lie Groups and Lie Algebras with Applications and Computational Methods

Ortega, James M., Numerical Analysis: *A Second Course*

Fiacco, Anthony V. and McCormick, Garth P., Nonlinear Programming: *Sequential Unconstrained Minimization Techniques*

Clarke, F. H., Optimization and Nonsmooth Analysis

Carrier, George F. and Pearson, Carl E., Ordinary Differential Equations

Breiman, Leo, Probability

Bellman, R. and Wing, G. M., An Introduction to Invariant Imbedding

Berman, Abraham and Plemmons, Robert J., Nonnegative Matrices in the Mathematical Sciences

Mangasarian, Olvi L., Nonlinear Programming

*Gauss, Carl Friedrich, Theory of the Combination of Observations Least Subject to Errors: *Part One, Part Two, Supplement.* Translated by G.W. Stewart

Bellman, Richard, Introduction to Matrix Analysis

*First time in print

# Theoria
# Combinationis Observationum
# Erroribus Minimis Obnoxiae

*Pars Prior ◆ Pars Posterior ◆ Supplementum*

By Carl Friedrich Gauss

# Theory of the
# Combination of Observations
# Least Subject to Errors

*Part One ◆ Part Two ◆ Supplement*

Translated by G. W. Stewart
*University of Maryland*

siam

Society for Industrial and Applied Mathematics
*Philadelphia 1995*

This SIAM edition includes works originally published as Theoria Combinationis Observationum Erroribus Minimis Obnoxiae, Pars Prior, Pars Posterior, Supplementum, Anzeigen.

10 9 8 7 6 5 4 3 2 1



Library of Congress Cataloging-in Publication Data

Gauss, Carl Friedrich, 1777-1855.
[Theoria combinationis observationum erroribus minimus obnoxiae. English]
Theory of the combination of observations least subject to error : part one, part two, supplement = Theoria combinationis observationum erroribus minimus obnoxiae : pars prior, pars posterior, supplementum / by Carl Friedrich Gauss ; translated by G.W. Stewart.
p. cm. -- (Classics in applied mathematics ; 11)
Includes bibliographical references (p. - ).
ISBN 0-89871-347-1 (pbk.)
1. Least squares. 2. Error analysis (Mathematics) I. Stewart, G. W. (Gilbert W.) II. Title. III. Series.
QA275.G37313 1995
511'.42--dc20 95-6589

# Contents



---

*The titles of the numbered articles are the translator's and are intended to help orient the reader. They do not appear in the numbered articles.

*This page intentionally left blank*

# Translator's Introduction

Although Gauss had discovered the method of least squares during the last decade of the eighteenth century and used it regularly after 1801 in astronomical calculations, it was Legendre who introduced it to the world in an appendix to an astronomical memoir. Legendre stated the principle of least squares for combining observations and derived the normal equations from which least squares estimates may be calculated. However, he provided no justification for the method, other than noting that it prevented extreme errors from prevailing by establishing a sort of equilibrium among all the errors, and he was content to refer the calculator to the methods of the day for solving linear systems.

In 1809, toward the end of his treatise on *The Theory of the Motion of Heavenly Bodies*, Gauss gave a probabilistic justification of the method, in which he essentially showed that if the errors are normal then least squares gives maximum likelihood estimates. However, his reasons for assuming normality were tenuous, and Gauss himself later rejected the approach. In other respects the treatment was more successful. It contains the first mention of Gaussian elimination (worked out in detail in a later publication), which was used to derive expressions for the precision of the estimates. He also described the Gauss–Newton method for solving nonlinear least squares problems and gave a characterization of what we would now call approximations in the $\ell_1$ norm.

Shortly thereafter, Laplace turned to the subject and derived the method of least squares from the principle that the best estimate should have the smallest mean error, by which he meant the mean of the absolute value of the error. Since the mean absolute error does not lead directly to the least squares principle, Laplace gave an asymptotic argument based on his central limit theorem.

In the 1820s Gauss returned to least squares in two memoirs, the first in two parts, published by the Royal Society of Gottingen under the common title *Theoria Combinationis Observationum Erroribus Minimis Obnoxiae*. In the *Pars Prior* of the first memoir, Gauss substituted the root mean square error for Laplace's mean absolute error. This enabled him to prove his minimum variance theorem: of all linear combinations of measurements estimating an unknown, the least squares estimate has the greatest precision. The remarkable thing about this theorem is that it does not depend on the distributions of the errors, and, unlike Laplace's result, it is not asymptotic.

The second part of the first memoir is dominated by computational considerations. Among other things Gauss gives several formulas for the residual sum of squares, a technique for adding and deleting an observation from an already solved problem, and new methods for computing variances. The second memoir, called *Supplementum*, is a largely self-contained work devoted to the application of the least squares principle to geodesy. The problem here is to adjust observations so that they satisfy certain constraints, and Gauss shows that the least squares solution is optimal in a very wide sense.

The following work is a translation of the *Theoria Combinationis Observationum* as it appears in Gauss's collected works, as well as the accompanying German notices (*Anzeigen*). The translator of Gauss, or of any author writing in Latin, must make some difficult choices. Historian and classicist Michael Grant quotes Pope's couplet†

> O come that easy Ciceronian style,
> So Latin, yet so English all the while.

and goes on to point out that Cicero and English have since diverged. Our language has the resources to render Gauss almost word for word into grammatically correct sentences. But the result is painful to read and does no justice to Gauss's style, which is balanced and lucid, albeit cautious.

In this translation I have aimed for the learned technical prose of our time. The effect is as if an editor had taken a blue pencil to a literal translation of Gauss: sentences and paragraphs have been divided; adverbs and adverbial phrases have been pruned; elaborate turns of phrase have been tightened. But there is a limit to this process, and I have tried never to abandon Gauss's meaning for ease of expression. Moreover, I have retained his original notation, which is not very different from ours and is sometimes revealing of his thought.

Regarding nomenclature, I have avoided technical terms, like "set," that have anachronistic associations. Otherwise I have not hesitated to use the modern term or phrase; e.g., "interval," "absolute value," "if and only if." Borderline cases are continuous for *continuus*, likelihood for *facilitas*, and estimate for *determinatio*. These are treated in footnotes at the appropriate places.‡

---

† "Translating Latin prose" in *The Translator's Art*, edited by William Radice and Barbara Williams, Viking Penguin, New York, 1987, p. 83.

‡ Translator's footnotes are numbered. Gauss's footnotes are indicated by *), as in his collected works.

The cost of all this is a loss of nuance, especially in tone, and historians who need to resolve fine points should consult the original, which accompanies the translation. For the rest, I hope I have produced a free but accurate rendering, which can read with profit by statisticians, numerical analysts, and other scientists who are interested in what Gauss did and how he set about doing it. In an afterword, I have attempted to put Gauss's contributions in historical perspective.

I am indebted to C. A. Truesdell, who made some very useful comments on my first attempt at a translation, and to Josef Stoer, who read the translation of the *Pars Prior*. Urs von Matt read the translation of Gauss's *Anzeigen*. Claudio Beccari kindly furnished his patterns for Latin hyphenation, and Charles Amos provided the systems support to use them. Of course, the responsibility for any errors in conception and execution is entirely mine.

I owe most to Woody Fuller, late professor of Germanic languages at the University of Tennessee and friend to all who had the good fortune to take his classes. He sparked my interest in languages and taught me that science is only half of human learning. This translation is dedicated to his memory.

College Park, Maryland
March 1995

*This page intentionally left blank*

# Theoria
# Combinationis Observationum
# Erroribus Minimis Obnoxiae

*Pars Prior*

# Theory of the
# Combination of Observations
# Least Subject to Errors

*Part One*

# Theoria Combinationis Observationum Erroribus Minimis Obnoxiae

## Pars Prior

### 1.

Quantacunque cura instituantur observationes, rerum naturalium magnitudinem spectantes, semper tamen erroribus maioribus minoribusve obnoxiae manent. Errores observatinum plerumque non sunt simplices, sed e pluribus fontibus simul originem trahunt: horum fontium duas species probe distinguere oportet. Quaedam errorum caussae ita sunt comparatae, ut ipsarum effectus in qualibet observatione a circumstantiis variabilibus pendeat, inter quas et ipsam observationem nullus nexus essentialis concipitur: errores hinc oriundi irregulares seu fortuiti vocantur, quatenusque illae circumstantiae calculo subiici nequeunt, idem etiam de erroribus ipsis valet. Tales sunt errores ab imperfectione sensuum provenientes, nec non a caussis extraneis irregularibus, e.g. a motu tremulo aeris visum tantillum turbante: plura quoque vitia instrumentorum vel optimorum huc trahenda sunt, e.g. asperitas partis interioris libellularum, defectus firmitatis absolutae etc. Contra aliae errorum caussae in omnibus observationibus ad idem genus relatis natura sua effectum vel absolute constantem exserunt, vel saltem talem, cuius magnitudo secundum legem determinatam unice a circumstantiis, quae tamquam essentialiter cum observatione nexae spectantur, pendet. Huiusmodi errores constantes seu regulares appellantur.

Ceterum perspicuum est, hanc distinctionem quodammodo relativam esse, et a sensu latiore vel arctiore, quo notio observationum ad idem genus pertinentium accipitur, pendere. E.g. vitia irregularia in divisione instrumentorum ad angulos mensurandos errorem constantem producunt, quoties tantummodo de observatione anguli determinati indefinite repetenda sermo est, siquidem hic semper eaedem devisiones vitiosae adhibentur: contra error ex illo fonte oriundus tamquam fortuitus spectari potest, quoties indefinite de angulis cuiusvis magnitudinis mensurandis agitur, siquidem tabula quantitatem erroris in singulis divisionibus exhibens non adest.

# Theory of the Combination of Observations Least Subject to Errors

## Part One

1.

However carefully one takes observations of the magnitudes of objects in nature, the results are always subject to larger or smaller errors. In general these errors are not simple but arise from many sources acting together. Two of these must be carefully distinguished.

Certain causes of error are such that their effect on any one observation depends on varying circumstances that seem to have no essential connection with the observation itself. Errors arising in this way are called irregular or random, and they are no more subject to calculation than the circumstances on which they depend. Such errors come from the imperfections of our senses and random external causes, as when shimmering air disturbs our fine vision. Many defects in instruments, even the best, fall in this category; e.g., a roughness in the inner part of a level, a lack of absolute rigidity, etc.

On the other hand, other sources of error by their nature have a constant effect on all observations of the same class. Or if the effect is not absolutely constant, its size varies regularly with circumstances that are essentially connected with the observations. These errors are called constant or regular.

Now it is clear that this distinction is to some extent relative and depends on how broadly we take the notion of observations of the same class. For example, consider irregularities in the graduations of an instrument for measuring angles. If we need to measure a given angle again and again, then the irregularities produce constant errors, since the same defective graduation is used repeatedly. On the other hand, the same errors can be regarded as random when one is measuring unknown angles of arbitrary magnitude, since there is no table of errors in the individual graduations.

2.

Errorum regularium consideratio proprie ab instituto nostro excluditur. Scilicet observatoris est, omnes caussas, quae errores constantes producere valent, sedulo investigare, et vel amovere, vel saltem earum rationem et magnitudinem summo studio perscrutari, ut effectus in quavis observatione determinata assignari, adeoque haec ab illo liberari possit, quo pacto res eodem redit, ac si error omnino non affuisset. Longe vero diversa est ratio errorum irregularium, qui natura sua calculo subiici nequeunt. Hos itaque in observationibus quidem tolerare, sed eorum effectum in quantitates ex observationibus derivandas per scitam harum cominationem quantum fieri potest extenuare oportet. Cui argumento gravissimo sequentes disquisitiones dicatae sunt.

3.

Errores observatinum ad idem genus pertinentium, qui a caussa simplici determinata oriuntur, per rei naturam certis *limitibus* sunt circumscripti, quos sine dubio exacte assignare liceret, si indoles ipsius caussae *penitus* esset perspecta. Pleraeque errorum fortuitorum caussae ita sunt comparatae, ut secundum legem continuitatis omnes errores intra istos limites comprehensi pro possibilibus haberi debeant, perfectaque caussae cognitio etiam doceret, utrum omnes hi errores aequali facilitate gaudeant an inaequali, et quanta probabilitas relative, in casu posteriore, cuivis errori tribuenda sit. Eadem etiam respectu erroris totalis, e pluribus erroribus simplicibus conflati, valebunt, puta inclusus erit certis limitibus (quorum alter aequalis erit aggregato omnium limitum superiorum partialium, alter aggregato omnium limitum inferiorum); omnes errores intra hos limites possibiles quidem erunt, sed prout quisque infinitis modis diversis ex erroribus partialibus componi postest, qui ipsi magis minusve probabiles sunt, alii maiorem, alii minorem facilitatem tribuere debebimus, eruique poterit lex probabilitatis relativae, si leges errorum simplicium cognitae supponuntur, salvis difficultatibus analyticis in colligendis omnibus combinationibus.

Exstant utique quaedam errorum caussae, quae errores non secundum legem continuitatis progredientes, sed discretos tantum, producere possunt, quales sunt errores divisionis instrumentorum (siquidem illos erroribus fortuitis annumerare placet): divisionum enim multitudo in quovis instrumento determinato est finita. Manifest autem, hoc non obstante si modo non omnes errorum caussae errores discretos producant, complexus omnium errorum totalium possibilium constituet serium secundum legem continuitatis progredientem, sive plures eiusmode series

2.

We explicitly exclude the consideration of regular errors from this investigation. Of course, it is up to the observer to ferret out all sources of constant error and remove them. Failing that, he should at least scrutinize their origins and magnitudes, so that their effects on any given observation can be determined and removed, after which it will be the same as if the errors had never occurred.

Irregular errors are essentially different, since by their nature they are not subject to calculation. For this reason we have to put up with them in the observations themselves; however, we should reduce their effects on derived quantities as far as possible by using judicious combinations of the observations. The following inquiry is devoted to this very important subject.

3.

Observation errors of the same class arising from a simple cause naturally lie within fixed limits. If we really knew the nature of the cause, we could determine the limits exactly. Most causes of random error obey a continuous law, so that all errors within the limits must be regarded as possible. From a perfect knowledge of the cause we could learn whether or not these errors were equally likely; or if not, then how to determine their relative probabilities.

The same is true of sums of simple errors. They lie within fixed limits (one of which is the sum of the upper limits of the simple errors and the other the sum of the lower limits). All errors within the limits are possible; but they are not equally likely since they can be formed from any number of combinations of their component errors, which themselves are more or less likely. Moreover, if the laws determining the relative probabilities of the simple errors are known, we can derive the laws for the compound errors — setting aside the analytic difficulties in enumerating all the combinations of the simple errors.

There are, of course, certain causes of error that produce discrete errors instead of continuous ones. Errors in the graduations of instruments — provided we agree to regard them as random — are of this kind, since the number of graduations in any one instrument is finite. In spite of this, it is clear that if at least one of the component errors is continuous then the total error will form an interval, or perhaps several disjoint intervals, within which it obeys a continuous law. The case of disjoint intervals occurs when some difference between two adjacent discrete errors (ordered by size) is greater than the difference of the limits of the total continuous error. But this case will almost never occur in practice unless a interruptas, si forte, omnibus erroribus discretis possibilibus secundum magnitudinem ordinatis, una alterave differentia inter binos terminos proximos maior evadat, quam differentia inter limites errorum totalium, quatenus e solis erroribus continuis demanant. Sed in praxi casus posterior vix umquam locum habebit, nisi divisio vitiis crassioribus laboret.

4.

Designando facilitatem relativam erroris totalis $x$, in determinato observationum genere, per characteristicam $\varphi x$, hoc, propter errorum continuitatem, ita intelligendum erit, probabilitatem erroris inter limites infinite proximos $x$ et $x+\mathrm{d}x$ esse $= \varphi x.\mathrm{d}x$. Vix, ac ne vix quidem, umquam in praxi possible erit, hanc functionem a priori assignare: nihilominus plura generalia eam spectantia stabilire possunt, quae deinceps proferemus. Obvium est, functionem $\varphi x$ eatenus ad functiones discontinuas referendam esse, quod pro omnibus valoribus ipsius $x$ extra limites errorum possibilium iacentibus esse debet $= 0$; intra hos limites vero ubique valorem positivum nanciscetur (omittendo casum, de quo in fine art. praec. locuti sumus. In plerisque casibus errores positivos et negativos eiusdem magnitudinis aeque faciles supponere licebit, quo pacto erit $\varphi(-x) = \varphi x$. Porro quum errores leviores facilius committantur quam graviores, plerumque valor ipsius $\varphi x$ erit maximus pro $x = 0$, continuoque decrescet, dum $x$ augetur.

Generaliter autem valor integralis $\int \varphi x.\mathrm{d}x$, ab $x = a$ usque ad $x = b$ extensi exprimet probabilitatem, quod error aliquis nondum cognitus iaceat inter limites $a$ et $b$. Valor itaque istius integralis a limite inferiore omnium errorum possibilium usque ad limitem superiorem semper erit $= 1$. Et quum $\varphi x$ pro omnibus valoribus ipsius $x$ extra hos limites iacentibus semper sit $= 0$, manifesto etiam

*valor integralis* $\int \varphi x.\mathrm{d}x$ *ab* $x = -\infty$ *usque ad* $x = +\infty$ *extensi semper fit* $= 1$.

5

Consideremus porro integrale $\int x\varphi x.\mathrm{d}x$ inter eosdem limites, cuius valorem statuemus $= k$. Si omnes errorum caussae simplices ita sunt comparatae, ut nulla adsit ratio, cur errorum aequalium sed signis oppositis affectorum alter facilius producatur quam alter, hoc etiam respectu erroris totalis valebit, sive erit $\varphi(-x) = \varphi x$, et proin necessario $k = 0$. Hinc colligimus, quoties $k$ non evanescat, sed e.g. sit quantitas positiva, necessario adesse debere unam alteramve errorum caussam, quae vel errores positivos tantum producere possit, vel certe positivos facilius quam negativos. Haecce quantitas $k$, quae revera est medium graduation suffers a gross deviation.

4.

Let $\varphi x$ denote the relative likelihood of a total error $x$ in a fixed class of observations.[1] By the continuity of the error, this means that the probability of an error lying between two infinitely close limits $x$ and $x + \mathrm{d}x$ is $\varphi x.\mathrm{d}x$. In practice we will seldom, if ever, be able to determine $\varphi$ a priori. Nonetheless, some general observations can be made.

First of all, it is clear that the function $\varphi x$ must be regarded as discontinuous,[2] because it is zero outside the limits of possible errors while it is positive within those limits (here we disregard the case mentioned at the end of the preceding article). In most instances, we may assume that positive and negative errors of the same magnitude are equally likely, so that $\varphi(-x) = \varphi x$. Moreover, since small errors are more likely to occur than large ones, $\varphi x$ will generally be largest for $x = 0$ and will decrease continuously with increasing $x$.

In general the value of the integral $\int \varphi x.\mathrm{d}x$ taken from $x = a$ to $x = b$ represents the probability that an error, as yet unknown, will lie between $a$ and $b$. Hence the value of this integral taken from the lower limit of the possible errors to the upper limit will be one. Since $\varphi x$ is zero outside these limits, it is clear that

*the value of the integral* $\int \varphi x.\mathrm{d}x$ *from* $x = -\infty$ *to* $x = +\infty$ *is always one.*

5.

Let us now consider the integral $\int x\varphi x.\mathrm{d}x$ between the above limits. We will denote its value by $k$. If all the simple causes of error are such that equal errors of opposite sign are equally likely, then the same is true of the total error and $\varphi(-x) = \varphi x$. It follows that $k = 0$. From this we see that if $k$ does not vanish, say it is positive, then some cause of error must produce only positive errors, or at least produce positive errors with greater likelihood than negative errors.[3]

---

[1]There are some tricky points of terminology here. Gauss calls $\varphi x$ both the relative likelihood (*facilitas*) of $x$ and the relative probability (*probabilitas*) of $x$. Since $\varphi$ is a density function, its values are not probabilities, at least for continuous distributions. Fortunately, Gauss's *facile* and *facilitas* are so near our own "likely" and "likelihood" and are so near the modern notion of likelihood that we can use them freely.

[2]Gauss's use of the word continuous is not ours and varies according to context. Here discontinuous probably means something like not analytic.

[3]This statement is true of the median, but not the mean.

omnium errorum possibilium, seu valor medius ipsius $x$, commode dici potest erroris pars constans. Ceterum facile probari potest, partem constantem erroris totalis aequalem esse aggregato partium constantium, quas continent errores e singulis caussis simplicibus prodeuntes. Quodsi quantitas $k$ nota supponitur, a quavis observatione resecatur, errorque observationis ita correctae designatur per $x'$, ipsiusque probabilitas per $\varphi' x'$, erit $x' = x - k$, $\varphi' x' = \varphi x$, ac proin $\int x'\varphi' x'.\mathrm{d}x' = \int x\varphi x.\mathrm{d}x - \int k\varphi x.\mathrm{d}x = k - k = 0$, i. e. errores observationum correctarum partem constantem non habebunt, quod et per se clarum est.

6.

Perinde ut integrale $\int x\varphi x.\mathrm{d}x$, seu valor medius ipsius $x$, erroris constantis vel absentiam vel praesentiam et magnitudinem docet, integrale

$$\int xx\varphi x.\mathrm{d}x$$

ab $x = -\infty$ usque ad $x = +\infty$ extensum (seu valor medius quadrati $xx$) aptissimum videtur ad incertitudinem observationum in genere definiendam et dimetiendam, ita ut e duobus observationum systematibus, quae quoad errorum facilitatem inter se differunt, eae praecisione praestare conseantur, in quibus integrale $\int xx\varphi x.\mathrm{d}x$ valorem minorem obtinet. Quodsi quis hanc rationem pro arbitrio, nulla cogente necessitate, electam esse obiiciat, lubenter assentiemur. Quippe quaestio haec per rei naturam aliquid vagi implicat, quod limitibus circumscribi nisi per principium aliquatenus arbitrarium nequit. Determinatio alicuius quantitatis per observationem errori maiori minorive obnoxiam, haud inepte comparature ludo, in quo solae iacturae, lucra nulla, dum quilibet error metuendus iacturae affinis est. Talis ludi dispendium aestimatur e iactura probabili, puta ex aggregato productorum singularum iacturarum possibilium in porbabilitates respectivas. Quantae vero iacturae quemlibet observationis errorem aequiparare conveniat, neutiquam per se clarum est; quin potius haec determinatio aliqua ex parte ab arbitrio nostro pendet. Iacturam ipsi errori aequalem statuere manifesto non licet; si enim errores positivi pro iacturis acciperentur, negativi lucra repraesentare deberent. Magnitudo iacturae potius per talem erroris functionem exprimi debet, quae natura sua semper fit positiva. Qualium functionum quum varietas sit infinita, simplicissima, quae hac proprietate gaudet, prae ceteris eligenda videtur, quae absque lite est quadratum: hoc pacto principium supra prolatum prodit.

Ill. LAPLACE simili quidem modo rem consideravit, sed errorem ipsum semper positive acceptum tamquam iacturae mensuram adoptavit. At ni fallimur haecce

We may call $k$ the *constant part* of the error, since in fact it is the center of all possible errors, that is, the mean value of $x$. Moreover, we can easily show that the constant part of a total error is equal to the sum of the constant parts of the errors produced by its simple causes. But now suppose we know the error and remove it from each observation. If we denote the corrected observation by $x'$ and its probability by $\varphi'$, then $x' = x - k$ and $\varphi' x' = \varphi x$. From this we have $\int x'\varphi' x'.\mathrm{d}x' = \int x\varphi x.\mathrm{d}x - \int k\varphi x.\mathrm{d}x = k - k = 0$. Thus the errors in the corrected observations have no constant part, a fact which is actually self-evident.

6.

The integral $\int x\varphi x.\mathrm{d}x$, i.e., the mean value of $x$, indicates the presence or absence of constant error, as well as its magnitude. Similarly, the integral

$$\int xx\varphi x.\mathrm{d}x$$

taken from $x = -\infty$ to $x = +\infty$ (the mean square of $x$) seems most appropriate to generally define and quantify the uncertainty of the observations. Thus, given two systems of observations which differ in their likelihoods, we will say that the one for which the integral $\int xx\varphi x.\mathrm{d}x$ is smaller is the more precise.

Now if someone should object that this convention has been chosen arbitrarily with no compelling necessity, I will gladly agree. In fact, the problem has a certain intrinsic vagueness about it that can only be resolved by a more or less arbitrary principle. It is not out of place to compare the estimation of a quantity by means of an observation subject to larger or smaller errors with a game of chance.[4] Since any error to be feared in an observation is connected with a loss, the game is one in which nobody wins and everybody loses. We estimate the outcome of such a game from the probable loss: namely, from the sum of the products of the individual losses with their respective probabilities.

It is by no means self-evident how much loss should be assigned to a given observation error. On the contrary, the matter depends in some part on our own judgment. Clearly we cannot set the loss equal to the error itself; for if positive errors were taken as losses, negative errors would have to represent gains. The size of the loss is better represented by a function that is naturally positive. Since

[4]Gauss's *determinatio* usually has the meaning of the calculation of a numerical quantity or the quantity so calculated. In many instances, however, the quantity estimates some unknown quantity, and in these cases it is appropriate to translate *determinatio* by "estimation" or "estimate."

ratio saltem non minus arbitraria est quam nostra: utrum enim error duplex aeque tolerabilis putetur quam simplex bis repetitus, an aegrius, et proin utrum magis conveniat, errori duplici momentum duplex tantum, an maius, tribuere, quaestio est neque per se clara, neque demonstrationibus mathematicis decidenda, sed libero tantum arbitrio remittenda. Praeterea negari non potest, ista ratione continuitatem laedi: et propter hanc ipsam caussam modus ille tractationi analyticae magis refragatur, dum ea, ad quae principium nostrum perducit, mira tum simplicitate tum generalitate commendantur.

7.

Statuendo valorem integralis $\int xx\varphi x.\mathrm{d}x$ ab $x=-\infty$ usque ad $x=+\infty$ extensi $=mm$, quantitatem $m$ vocabimus *errorem medium metuendum,* sive simpliciter *errorem medium* observationum, quarum errores indefiniti $x$ habent probabilitatem relativam $\varphi x$. Denominationem illam non ad observationes immediatas limitabimus, sed etiam ad determinationes qualescunque ex observationibus derivatas extendemus. Probe autem cavendum est, ne error medius confundatur cum medio arithmetico omnium errorum, de quo in art. 5 locuti sumus.

Ubi plura observationum genera, seu plures determinationes ex observationibus petitae, quibus haud eadem praecisio concedenda est, comparantur, *pondus* earum relativum nobis erit quantitas ipsi $mm$ reciproce proportionalis, dum *praecisio* simpliciter ipsi $m$ reciproce proportionalis habetur. Quo igitur pondus per numerum exprimi possit, pondus certi observationum generis pro unitate acceptum esse debet.

8.

Si observationum errores partem constantem implicant, hanc auferendo error medius minuitur, pondus et praecisio augentur. Retinendo signa art. 5, designandoque per $m'$ errorem medium observationum correctarum, erit

$$\begin{aligned} m'm' &= \int x'x'\varphi' x'.\mathrm{d}x' = \int (x-k)^2\varphi x.\mathrm{d}x \\ &= \int xx\varphi x.\mathrm{d}x - 2k\int x\varphi x.\mathrm{d}x + kk\int \varphi x.\mathrm{d}x = mm - 2kk + kk = mm - kk. \end{aligned}$$

Si autem loco partis constantis veri $k$ quantitas alia $l$ ab observationibus ablata esset, quadratum erroris medii novi evaderet $= mm - 2kl + ll = m'm' + (l-k)^2$.

the number of such functions is infinite, it would seem that we should choose the simplest function having this property. That function is unarguably the square, and the principle proposed above results from its adoption.

LAPLACE has also considered the problem in a similar manner, but he adopted the absolute value of the error as his measure of loss. Now if I am not mistaken, this convention is no less arbitrary than mine. Should an error of double size be considered as tolerable as a single error twice repeated or worse? Is it better to assign only twice as much influence to a double error or more? The answers are not self-evident, and the problem cannot be resolved by mathematical proofs, but only by an arbitrary decision. Moreover, it cannot be denied that LAPLACE'S convention violates continuity and hence resists analytic treatment, while the results that my convention leads to are distinguished by their wonderful simplicity and generality.

## 7.

Let $mm$ denote the integral $\int xx\varphi x.\mathrm{d}x$ from $x = -\infty$ to $x = +\infty$. We will call $m$ the *mean error* or the *mean error to be feared* in observations whose errors have relative probability $\varphi x$. We will not restrict this terminology to just the observations but will extend it to any quantities derived from them. However, one should take care not to confuse this mean error with the arithmetic mean, which was treated in Art. 5.

When we compare several classes of observations—or several quantities derived from the observations—not having the same precision, we will take their relative *weights* to be quantities proportional to the reciprocals of $mm$. Likewise their *precisions* will be proportional to the reciprocals of $m$. In order to represent weights numerically, the weight of one of the classes of observations should be set to one.

## 8.

If the observation errors have a constant part, removing it reduces their mean error and increases their weight and precision. In the notation of Art. 5, if $m'$ denotes the mean error of the corrected observation, then

$$\begin{aligned} m'm' &= \int x'x'\varphi'x'.\mathrm{d}x' = \int (x-k)^2\varphi x.\mathrm{d}x \\ &= \int xx\varphi x.\mathrm{d}x - 2k\int x\varphi x.\mathrm{d}x + kk\int \varphi x.\mathrm{d}x = mm - 2kk + kk = mm - kk. \end{aligned}$$

However, if instead of the true constant part $k$ we remove another quantity $l$ from the observations, the new mean square error becomes $mm - 2kl + ll = m'm' + (l-k)^2$.

9.

Denotante $\lambda$ coëfficientem determinatum, atque $\mu$ valorem integralis $\int \varphi x.dx$ ab $x = -\lambda m$ usque ad $x = +\lambda m$, erit $\mu$ probabilitas, quod error alicuius observationis sit minor quam $\lambda m$ (sine respectu signi), nec non $1 - \mu$ probabilitas erroris maioris quam $\lambda m$. Si itaque valor $\mu = \frac{1}{2}$ respondet valori $\lambda m = \rho$, error aeque facile infra $\rho$ quam supra $\rho$ cadere potest, quocirca $\rho$ commode dici potest error probabilis. Relatio quantitatum $\lambda$, $\mu$ manifesto pendet ab indole functionis $\varphi x$, quae plerumque incognita est. Operae itaque pretium erit, istam relationem pro quibusdam casibus specialibus propius considerare.

I. Si limites omnium errorum possibilium sunt $-a$ et $+a$, omnesque errores intra hos limites aeque probabiles, erit $\varphi x$ inter limites $x = -a$ et $x = +a$ constans, et proin $= \frac{1}{2a}$. Hinc $m = a\sqrt{\frac{1}{3}}$, nec non $\mu = \lambda\sqrt{\frac{1}{3}}$, quamdiu $\lambda$ non maior quam $\sqrt{3}$; denique $\rho = m\sqrt{\frac{3}{4}} = 0.8660254\,m$, probabilitasque, quod error prodeat errore medio non maior, erit $= \sqrt{\frac{1}{3}} = 0.5773503$.

II. Si ut antea $-a$ et $+a$ sunt errorum possibilium limites, errorumque ipsorum probabilitas inde ab errore 0 utrimque in progressione arthmetica decrescere supponitur, erit

$$\varphi x = \frac{a-x}{aa}, \text{ pro valoribus ipsius } x \text{ inter 0 et } +a$$

$$\varphi x = \frac{a+x}{aa}, \text{ pro valoribus ipsius } x \text{ inter 0 et } -a$$

Hinc deduciter $m = a\sqrt{\frac{1}{6}}$, $\mu = \lambda\sqrt{\frac{2}{3}} - \frac{1}{6}\lambda\lambda$, quamdiu $\lambda$ est inter 0 et $\sqrt{6}$, denique $\lambda = \sqrt{6} - \sqrt{6 - 6\mu}$, quamdiu $\mu$ inter 0 et 1, et proin

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

Probabilitas erroris medium non superantis erit in hoc casu

$$\sqrt{\tfrac{2}{3}} - \tfrac{1}{6} = 0.6498299$$

III. Si functionem $\varphi x$ proportionalem statuimus huic $e^{-\frac{xx}{hh}}$ (quod quidem in rerum natura proxime tantum verum esse potest), esse debebit

$$\varphi x = \frac{e^{-\frac{xx}{hh}}}{h\sqrt{\pi}}$$

denotante $\pi$ semiperipheriam circuli pro radio 1, unde porro deducimus

$$m = h\sqrt{\tfrac{1}{2}}$$

9.

For any value of $\lambda$, let $\mu$ be the value of the integral $\int \varphi x.dx$ from $x = -\lambda m$ to $x = +\lambda m$. Then $\mu$ is the probability that the error (disregarding signs) in any one observation will be less than $\lambda m$, and $1 - \mu$ is the probability that the error will be greater than $\lambda m$. Thus if the value $\mu = \frac{1}{2}$ corresponds to the value $\lambda m = \rho$, the error is equally likely to be less than $\rho$ or greater than $\rho$. For this reason, it is appropriate to call $\rho$ the *probable error.* The relation between the quantities $\lambda$ and $\mu$ obviously depends on the function $\varphi x$, which is usually unknown. It is therefore worthwhile to examine this relation more closely for certain special cases.

I. Suppose the limits of all possible errors are $-a$ and $+a$, and all errors between these limits are equally probable. Then $\varphi x$ will be constant between $x = -a$ and $x = +a$ and in fact will be equal to $\frac{1}{2a}$. Hence $m = a\sqrt{\frac{1}{3}}$, and $\mu = \lambda\sqrt{\frac{1}{3}}$, as long as $\lambda$ is not greater that $\sqrt{3}$. Finally, $\rho = m\sqrt{\frac{3}{4}} = 0.8660254\,m$, and the probability that an error will not exceed the mean error is $\sqrt{\frac{1}{3}} = 0.5773503$.

II. As above suppose that the limits of possible errors are $-a$ and $+a$, and suppose that the probability of the error decreases arithmetically on both sides of zero. Then

$$\varphi x = \tfrac{a-x}{aa} \text{ for } x \text{ between } 0 \text{ and } +a,$$

$$\varphi x = \tfrac{a+x}{aa} \text{ for } x \text{ between } 0 \text{ and } -a.$$

From this we find that $m = a\sqrt{\frac{1}{6}}$ and $\mu = \lambda\sqrt{\frac{2}{3}} - \frac{1}{6}\lambda\lambda$, as long as $\lambda$ is between 0 and $\sqrt{6}$. Finally, $\lambda = \sqrt{6} - \sqrt{6 - 6\mu}$, as long as $\mu$ is between 0 and 1, and

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

For this case the probability of not exceeding the mean error is

$$\sqrt{\tfrac{2}{3}} - \tfrac{1}{6} = 0.6498299.$$

III. If we take the function $\varphi x$ to be proportional to $e^{-\frac{xx}{hh}}$ (which can be only approximately true in real life), then we must have

$$\varphi x = \frac{e^{-\frac{xx}{hh}}}{h\sqrt{\pi}},$$

where $\pi$ denotes half the perimeter of a circle of radius 1. From this we find that

$$m = h\sqrt{\tfrac{1}{2}}$$

(V. *Disquis. generales circa seriem infinitam* etc. art. 28). Porro si valor integralis

$$\frac{2}{\sqrt{\pi}}\int e^{-zz}\mathrm{d}z$$

a $z=0$ inchoati denotatur per $\Theta z$, erit

$$\mu=\Theta(\lambda\sqrt{\tfrac{1}{2}})$$

Tabula sequens exhibet aliquot valores huius quantitatis:

| $\lambda$ | $\mu$ |
|---|---|
| 0.6744897 | 0.5 |
| 0.8416213 | 0.6 |
| 1.0000000 | 0.6826895 |
| 1.0364334 | 0.7 |
| 1.2815517 | 0.8 |
| 1.6448537 | 0.9 |
| 2.5758293 | 0.99 |
| 3.2918301 | 0.999 |
| 3.8905940 | 0.9999 |
| $\infty$ | 1 |

10.

Quamquam relatio inter $\lambda$ et $\mu$ ab indole functionis $\varphi x$ pendet, tamen quaedam generalia stabilire licet. Scilicet qualiscunque sit haec functio, si modo ita est comparata, ut ipsius valor, crescente valore absoluto ipsius $x$, semper decrescat, vel saltem non crescat, certo erit

$\lambda$ minor vel saltem non maior quam $\mu\sqrt{3}$, quoties $\mu$ est minor quam $\frac{2}{3}$;

$\lambda$ non maior quam $\frac{2}{3\sqrt{1-\mu}}$, quoties $\mu$ est maior quam $\frac{2}{3}$.

Pro $\mu=\frac{2}{3}$ uterque limes coincidit, puta $\lambda$ nequit esse maior quam $\sqrt{\frac{4}{3}}$.

Ut hoc insigne theorema demonstremus, denotemus per $y$ valorem integralis $\int\varphi z.\mathrm{d}z$ ab $z=-x$ usque ad $z=+x$ extensi, quo pacto $y$ erit probabilitas, quod error aliquis contentus sit intra limites $-x$ et $+x$. Porro statuamus

$$x=\psi y,\quad \mathrm{d}\psi y=\psi' y.\mathrm{d}y,\quad \mathrm{d}\psi' y=\psi'' y.\mathrm{d}y$$

Erit taque $\psi 0=0$, nec non

$$\psi' y=\frac{1}{\varphi x+\varphi(-x)}$$

(see *Disquis. generales circa seriem infinitam* ..., Art. 28). Moreover, if $\Theta z$ denotes the integral

$$\frac{2}{\sqrt{\pi}} \int e^{-zz} \mathrm{d}z$$

taken from $z = 0$, then

$$\mu = \Theta(\lambda\sqrt{\tfrac{1}{2}}).$$

The following table gives some values of these quantities.

| $\lambda$ | $\mu$ |
|---|---|
| 0.6744897 | 0.5 |
| 0.8416213 | 0.6 |
| 1.0000000 | 0.6826895 |
| 1.0364334 | 0.7 |
| 1.2815517 | 0.8 |
| 1.6448537 | 0.9 |
| 2.5758293 | 0.99 |
| 3.2918301 | 0.999 |
| 3.8905940 | 0.9999 |
| $\infty$ | 1 |

10.

Although the relation between $\lambda$ and $\mu$ depends on the nature of the function $\varphi x$, we can still establish some general facts about it. Specifically, if such function decreases (or at least does not increase) as the absolute value of $x$ increases, then

$\lambda$ is less than or equal to $\mu\sqrt{3}$, whenever $\mu$ is less than $\frac{2}{3}$;

$\lambda$ is not greater than $\frac{2}{3\sqrt{1-\mu}}$, whenever $\mu$ is greater than $\frac{2}{3}$.

For $\mu = \frac{2}{3}$ the two bounds coincide and $\lambda$ cannot be greater than $\sqrt{\frac{4}{3}}$.

To prove this remarkable theorem, let $y$ denote the value of the integral $\int \varphi z.\mathrm{d}z$ taken from $z = -x$ to $z = +x$, so that $y$ is the probability that an error will be bounded by the limits $-x$ and $+x$. Further, set

$$x = \psi y, \quad \mathrm{d}\psi y = \psi' y.\mathrm{d}y, \quad \mathrm{d}\psi' y = \psi'' y.\mathrm{d}y.$$

Then $\psi 0 = 0$ and

$$\psi' y = \frac{1}{\varphi x + \varphi(-x)}.$$

quare per hyp. $\psi' y$ ab $y = 0$ usque ad $y = 1$ semper crescet, saltem nullibi decrescet, sive, quod idem est, valor ipsius $\psi'' y$ semper erit positivus, vel saltem non negativus. Porro habemus $\mathrm{d}.y\psi' y = \psi' y.\mathrm{d}y + y\psi'' y.\mathrm{d}y$, adeoque

$$y\psi' y - \psi y = \int y\psi'' y.\mathrm{d}y$$

integratione ab $y = 0$ inchoata. Valor expressionis $y\psi' y - \psi y$ itaque semper erit quantitas positive, saltem non negative, adeoque

$$1 - \tfrac{\psi y}{y\psi' y}$$

quantitas positive unitate minor. Sit $f$ eius valor pro $y = \mu$, i.e. quum habeature $\psi\mu = \lambda m$, sit

$$f = 1 - \tfrac{\lambda m}{\mu\psi'\mu} \quad \text{sive} \quad \psi'\mu = \tfrac{\lambda m}{(1-f)\mu}$$

His ita praeparatis, consideremus functionem ipsius $y$ hanc

$$\tfrac{\lambda m}{(1-f)\mu}(y - \mu f)$$

quam statuemus $= Fy$, nec non $\mathrm{d}Fy = F' y.\mathrm{d}y$. Perspicuum est, feiri

$$F\mu = \lambda m = \psi\mu$$

$$F'\mu = \tfrac{\lambda m}{(1-f)\mu} = \psi'\mu$$

Quare quum $\psi' y$, aucta ipsa $y$, continuo crescat (saltem non decrescat, quod semper subintelligendum), $F' y$ vero constans sit, differentia $\psi' y - F' y = \frac{\mathrm{d}(\psi y - Fy)}{\mathrm{d}y}$ erit positiva pro valoribus ipsius $y$ maioribus quam $\mu$, negativa pro minoribus. Hinc facile colligitur, $\psi y - Fy$ semper esse quantitatem positivam, adeoque $\psi y$ semper erit absolute maior, saltem non minor, quam $Fy$, certe quamdiu valor ipsius $Fy$ est positivus, i.e. ab $y = \mu f$ usque at $y = 1$. Hinc valor integralis $\int (Fy)^2 \mathrm{d}y$ ab $y = \mu f$ usque ad $y = 1$ erit minor valore integralis $\int (\psi y)^2 \mathrm{d}y$ inter eosdem limites, adeoque a potiori minor valore huius integralis ab $y = 0$ usque ad $y = 1$, qui fit $mm$. At valor integralis prioris invenitur

$$= \tfrac{\lambda\lambda mm(1-\mu f)^3}{3\mu\mu(1-\mu f)^2}$$

Hence by hypothesis $\psi' y$ always increases from $y = 0$ to $y = 1$, or at least it never decreases. Equivalently, the value of $\psi'' y$ is always positive, or at least nonnegative.

Next we have $\mathrm{d}.y\psi' y = \psi' y.\mathrm{d}y + y\psi'' y.\mathrm{d}y$, so that

$$y\psi' y - \psi y = \textstyle\int y\psi'' y.\mathrm{d}y,$$

where the integral starts at $y = 0$. Thus the expression $y\psi' y - \psi y$ will always be positive, or at least nonnegative, and

$$1 - \tfrac{\psi y}{y\psi' y}$$

will be positive and less than one. Let $f$ be the value of this last expression when $y = \mu$. Since $\psi\mu = \lambda m$, we have

$$f = 1 - \tfrac{\lambda m}{\mu\psi'\mu} \quad \text{or} \quad \psi'\mu = \tfrac{\lambda m}{(1-f)\mu}.$$

With these facts established, we now consider the function

$$Fy = \tfrac{\lambda m}{(1-f)\mu}(y - \mu f)$$

and its differential $\mathrm{d}Fy = F' y.\mathrm{d}y$. It is clear that

$$F\mu = \lambda m = \psi\mu,$$

$$F'\mu = \tfrac{\lambda m}{(1-f)\mu} = \psi'\mu.$$

It follows that since $\psi' y$ increases with $y$ (or at least does not decrease—which is always to be understood) and since $F' y$ is a constant, the difference $\psi' y - F' y = \frac{\mathrm{d}(\psi y - Fy)}{\mathrm{d}y}$ is positive when $y$ is greater than $\mu$ and negative when $y$ is less than $\mu$. From this it follows easily that $\psi y - Fy$ is always positive. Hence the absolute value of $\psi y$ is greater than or equal to the absolute value of $Fy$ whenever $Fy$ is positive, i.e., from $y = \mu f$ to $y = 1$. Hence the integral $\int (Fy)^2 \mathrm{d}y$ from $y = \mu f$ to $y = 1$ is less than the integral $\int (\psi y)^2 \mathrm{d}y$ between the same limits and is therefore less than the latter integral taken from $y = 0$ to $y = 1$, which is $mm$.

The value of the first integral is

$$\tfrac{\lambda\lambda mm(1-\mu f)^3}{3\mu\mu(1-\mu f)^2}.$$

unde colligimus, $\lambda\lambda$ esse minorem quam $\frac{3\mu\mu(1-f)^2}{(1-\mu f)^3}$, ubi $f$ est quantitas inter 0 et 1 iacens. Iam valor fractionis $\frac{3\mu\mu(1-f)^2}{(1-\mu f)^3}$, cuius differentiale, si $f$ tamquam quantitas variabilis consideratur, fit =

$$-\frac{3\mu\mu(1-f)}{(1-\mu f)^4}\cdot(2-3\mu+\mu f)\mathrm{d}f$$

continue decrescit, dum $f$ a valore 0 usque ad valorem 1 transit, quoties $\mu$ minor est quam $\frac{2}{3}$, adeoque valor maximus possibilis erit is, qui valori $f=0$ respondet, puta $=3\mu\mu$, ita ut in hoc casu $\lambda$ certo fiat minor vel non maior quam $\mu\sqrt{3}$. Q.E.P. Contra quoties $\mu$ maior est quam $\frac{2}{3}$, valor istius fractionis erit maximus pro $2-3\mu+\mu f=0$, i.e. pro $f=3-\frac{2}{\mu}$, unde ille fit $\frac{4}{9(1-\mu)}$, adeoque in hoc casu $\lambda$ non maior quam $\frac{2}{3\sqrt{1-\mu}}$. Q.E.S.

Ita e.g. pro $\mu=\frac{1}{2}$ certo $\lambda$ nequit esse maior quam $\sqrt{\frac{3}{4}}$, i.e. error probabilis superare nequit limitem $0.8660254m$, cui in exemplo primo art. 9 aequaelis inventus est. Porro facile e theoremate nostro concluditur, $\mu$ non esse minorem quam $\lambda\sqrt{\frac{1}{3}}$, quamdiu $\lambda$ minor sit quam $\sqrt{\frac{4}{3}}$, contra $\mu$ non esse minorem quam $1-\frac{4}{9\lambda\lambda}$, pro valore ipsius $\lambda$ minor quam $\sqrt{\frac{4}{3}}$.

## 11.

Quum plura problemata infra tractanda etiam cum valore integralis $\int x^4\varphi x.\mathrm{d}x$ nexa sint, operae pretium erit, eum pro quibusdam casibus specialibus evolvere. Denotabimus valorem huius integralis ab $x=-\infty$ usque ad $x=+\infty$ extensi per $n^4$.

I. Pro $\varphi x=\frac{1}{2a}$, quatenus $x$ inter $-a$ et $+a$ continetur, habemus $n^4=\frac{1}{5}a^4=\frac{9}{5}m^4$.

II. In casu secundo art. 6, ubi $\varphi x=\frac{a\mp x}{aa}$, pro valoribus ipsius $x$ inter 0 et $\pm a$, fit $n^4=\frac{1}{15}a^4=\frac{12}{5}m^4$.

III. In casu tertio, ubi

$$\varphi x=\frac{e^{-\frac{xx}{hh}}}{a\sqrt{\pi}}$$

invenitur per eq, quae in commentatione supra citata exponuntur $n^4=\frac{3}{4}h^4=3m^4$.

Ceterum demonstrari potest, valorem ipsius $\frac{n^4}{m^4}$ certo non esse minorem quam $\frac{9}{5}$, si modo suppositio art. praec. locum habeat.

From this we see that $\lambda\lambda$ is less than $\frac{3\mu\mu(1-f)^2}{(1-\mu f)^3}$, where $f$ is a number lying between 0 and 1. Now the fraction $\frac{3\mu\mu(1-f)^2}{(1-\mu f)^3}$ regarded as a function of $f$ has the differential

$$-\tfrac{3\mu\mu(1-f)}{(1-\mu f)^4} \cdot (2 - 3\mu + \mu f)\mathrm{d}f.$$

When $\mu$ is less that $\frac{2}{3}$, the fraction decreases as $f$ ranges from 0 to 1. Hence it assumes its greatest value, namely, $3\mu\mu$, at $f = 0$. In this case $\lambda$ is certainly less than or equal to $\mu\sqrt{3}$, which is the first inequality. On the other hand, if $\mu$ is greater than $\frac{2}{3}$, the fraction is largest for $2 - 3\mu + \mu f = 0$, i.e., for $f = 3 - \frac{2}{\mu}$. Its value at this point is $\frac{4}{9(1-\mu)}$, and in this case $\lambda$ is not greater than $\frac{2}{3\sqrt{1-\mu}}$, which is the second inequality.

As an example, when $\mu = \frac{1}{2}$, $\lambda$ cannot be greater than $\sqrt{\frac{4}{3}}$; i.e., the probable error cannot exceed $0.8660254m$, a limit which is attained in the first example of Art. 9. Moreover, from our theorem it is easily established that $\mu$ is not less than $\lambda\sqrt{\frac{1}{3}}$ as long as $\lambda$ is less than $\sqrt{\frac{4}{3}}$. On the other hand, $\mu$ is not less than $1 - \frac{4}{9\lambda\lambda}$ for $\lambda$ greater than $\sqrt{\frac{4}{3}}$.[5]

## 11.

Since several problems to be treated later depend on the integral $\int x^4\varphi x.\mathrm{d}x$, it will be worthwhile to derive its value for our special cases. We will denote the value of this integral for $x = -\infty$ to $x = +\infty$ by $n^4$.

I. For $\varphi x = \frac{1}{2a}$ when $x$ lies between $-a$ and $+a$, we have $n^4 = \frac{1}{5}a^4 = \frac{9}{5}m^4$.

II. In the second case of Art. 6, where $\varphi x = \frac{a\mp x}{aa}$ for $x$ between 0 and $\pm a$, we have $n^4 = \frac{1}{15}a^4 = \frac{12}{5}m^4$.

III. In the third case, where

$$\varphi x = \frac{e^{-\frac{xx}{hh}}}{a\sqrt{\pi}},$$

we find from the reference cited above that $n^4 = \frac{3}{4}h^4 = 3m^4$.

Moreover, under the assumption of the preceding article it can be shown that $\frac{n^4}{m^4}$ is never less than $\frac{9}{5}$.

[5]The text reads *pro valore ipsius* $\lambda$ *minor quam* $\sqrt{\frac{4}{3}}$, an obvious misprint.

## 12.

Denotantibus $x$, $x'$, $x''$, etc. indefinite errores observationum eiusdem generis ab invicem independentes, quorum probabilitates relativas exprimit praefixa characteristica $\varphi$; nec non $y$ functionem datam rationalem indeterminatarum $x$, $x'$, $x''$ etc.; integral multiplex (I)

$$\int \varphi x.\varphi x'.\varphi x'' \ldots \mathrm{d}x.\mathrm{d}x'.\mathrm{d}x'' \ldots$$

extensum per omnes valores indeterminatarum $x$, $x'$, $x''$, pro quibus valor ipsius $y$ cadit intra limites datos 0 et $\eta$, exprimit probabilitatem valoris ipsius $y$ indefinite intra 0 et $\eta$ siti. Manifesto hoc integrale erit functio ipsius $\eta$, cuius differential statuemus $= \psi\eta.\mathrm{d}\eta$, ita ut integrale ipsum fiat aequale integrali $\int \psi\eta.\mathrm{d}\eta$ ab $\eta = 0$ incepto. Hoc pacto simul characteristica $\psi y$ probabilitatem relativam cuiusvis valoris ipsius $y$ exprimere censenda est. Quum $x$ considerari possit tamquam functio indeterminatarum $y$, $x'$, $x''$, etc., quam statuemus

$$= f(y, x', x'', \ldots)$$

integrale (I) fiet

$$= \int \varphi.f(y, x', x'', \ldots).\tfrac{\mathrm{d}f(y,x',x'',\ldots)}{\mathrm{d}y}.\varphi x'.\varphi x'' \ldots \mathrm{d}y.\mathrm{d}x'.\mathrm{d}x'' \ldots$$

ubi $y$ extendi debet ab $y = 0$ usque ad $y = \eta$, indeterminatae reliquae vero per omnes valores, quibus respondet valor realis ipsius $f(y, x', x'', \ldots)$. Hinc colligitur

$$\psi y = \int \varphi.f(y, x', x'', \ldots).\tfrac{\mathrm{d}f(y,x',x'',\ldots)}{\mathrm{d}y}.\varphi x'.\varphi x'' \ldots \mathrm{d}x'.\mathrm{d}x'' \ldots$$

integratione, in qua $y$ tamquam constans considerari debet, extensa per omnes valores indeterminatarum $x'$, $x''$, etc., qui ipsi $f(y, x', x'', \ldots)$ valorem realem conciliant.

## 13.

Ad hanc integrationem reipsa exsequendam cognitio functionis $\varphi$ requireretur, quae plerumque incognita est: quin adeo, etiamsi haec functio cognita esset, in plerisque casibus integratio vires analyseos superaret. Quae quum ita sint, probabilitatem quidem singulorum valorum ipsius $y$ assignare non poterimus: at secus res se habebit, si tantummodo desideratur valor medius ipsius $y$, qui oritur ex integration $\int y\psi y.\mathrm{d}y$ per omnes valores ipsius $y$, quos quidem assequi potest, extensa.

12.

Let $x$, $x'$, $x''$, etc. denote mutually independent random errors in observations of the same class, and let $\varphi$ denote their relative probability. Let $y$ denote a rational function of the unknowns $x$, $x'$, $x''$, etc. The probability that the function $y$ will lie between 0 and $\eta$ is represented by the multiple integral (I)

$$\int \varphi x.\varphi x'.\varphi x'' \ldots \mathrm{d}x.\mathrm{d}x'.\mathrm{d}x'' \ldots,$$

taken over all values of the unknowns $x$, $x'$, $x''$, for which the value of $y$ falls between 0 and $\eta$. Obviously, this integral is a function of $\eta$. We will set its differential equal to $\psi\eta.\mathrm{d}\eta$, so that the integral is equal to $\int \psi\eta.\mathrm{d}\eta$ taken from $\eta = 0$. Thus $\psi y$ represents the relative probability of a value of $y$.

Since $x$ can be regarded as a function of the unknowns $y$, $x'$, $x''$, etc., a function which we will call

$$f(y, x', x'', \ldots),$$

the integral (I) becomes

$$\int \varphi.f(y, x', x'', \ldots) \cdot \frac{\mathrm{d}f(y,x',x'',\ldots)}{\mathrm{d}y}.\varphi x'.\varphi x'' \ldots \mathrm{d}y.\mathrm{d}x'.\mathrm{d}x'' \ldots,$$

where $y$ ranges from $y = 0$ to $y = \eta$ and the other unknowns range over all values for which the value of $f(y, x', x'', \ldots)$ is real.[6] From this we see that

$$\psi y = \int \varphi.f(y, x', x'', \ldots).\frac{\mathrm{d}f(y,x',x'',\ldots)}{\mathrm{d}y}.\varphi x'.\varphi x'' \ldots \mathrm{d}x'.\mathrm{d}x'' \ldots,$$

in which $y$ is constant and the integration extends over all values of $x'$, $x''$, etc. that give a real value to $f(y, x', x'', \ldots)$.

13.

To actually perform this integration, we have to know the function $\varphi$, which is usually unknown. Even if it were known, the integration would, in most cases, be beyond our analytic powers. This being so, we will not be able to assign probabilities to individual values of $y$. But it is a different matter if we only want

[6]This interchange of $x$ and $y$ is not always possible; for example, consider $y = x^2$. It is possible that Gauss intends to allow $f$ to be multivalued and his integrals to be computed by summing the integrals of the branches of $f$. It should be noted that Gauss is trying to establish a proposition—namely, that the mean of $y$ is $\int y\varphi x.\varphi x'.\varphi x'' \ldots \mathrm{d}x.\mathrm{d}x'.\mathrm{d}x'' \ldots$—which is best approached via measure theory and the Lebesgue integral.

Et quum manifesto pro omnibus valoribus, quos $y$ assequi nequit, vel per naturam functionis, quam exprimit (e.g. pro negativis, si esset $y = xx + x'x' + x''x'' + \text{etc.}$), vel ideo, quod erroribus ipsis $x$, $x'$, $x''$, etc. certi limites sunt positi, statuere oporteat $\varphi y = 0$, manifest res perinde se habebit, si integratio illa extendatur per omnes valores reales ipsius $y$, puta ab $y = -\infty$ usque ad $y = +\infty$. Iam integrale $\int y\psi y.\mathrm{d}y$ inter limites determinatos, puta ab $y = \eta$ usque ad $y = \eta'$ sumtum aequale est integrali

$$\int y\varphi.f(y, x', x'', \ldots).\tfrac{\mathrm{d}f(y,x',x'',\ldots)}{\mathrm{d}y}.\varphi x'.\varphi x'' \ldots \mathrm{d}y.\mathrm{d}x'.\mathrm{d}x'' \ldots$$

integratione extensa ab $y = \eta$ usque ad $y = \eta'$, atque per omnes valores indeterminatarum $x'$, $x''$, etc., quibus respondet valor realis ipsius $f(y, x', x'', \ldots)$, sive quod idem est, valori integralis

$$\int y\varphi x.\varphi x'.\varphi x'' \ldots \mathrm{d}x.\mathrm{d}x'.\mathrm{d}x'' \ldots$$

adhibendo in hac integratione pro $y$ eius valorem per $x$, $x'$, $x''$, etc. expressum, extendendoque eam per omnes harum indeterminatarum valores, quibus respondet valor ipsius $y$ inter $\eta$ et $\eta'$ situs. Hinc colligimus, integrale $\int y\psi y.\mathrm{d}y$ per omnes valores ipsius $y$, ab $y = -\infty$ usque ad $y = +\infty$ extensum obtineri ex integratione

$$\int y\varphi x.\varphi x'.\varphi x'' \ldots \mathrm{d}x.\mathrm{d}x'.\mathrm{d}x'' \ldots$$

per omnes valores reales ipsarum $x$, $x'$, $x''$, etc. extensa, puta ab $x = -\infty$ usque ad $x = +\infty$, ab $x' = -\infty$ usque ad $x' = +\infty$, etc.

14.

Reducta itaque functione $y$ ad formam aggregati talium partium

$$Ax^\alpha x'^\beta x''^\gamma \ldots$$

valor integralis $\int y\psi y.\mathrm{d}y$ per omnes valores ipsius $y$ extensi, seu valor medius ipsius $y$ aequalis erit aggrato partium

$$A \times \int x^\alpha \varphi x.\mathrm{d}x \times \int x'^\beta \varphi x'.\mathrm{d}x' \times \int x''^\gamma \varphi x''.\mathrm{d}x'' \ldots$$

ubi integrationes extendendae sunt ab $x = -\infty$ usque ab $x = +\infty$, ab $x' = -\infty$ useque ab $x' = +\infty$, etc.; sivi quod eodem redit, aggreato partium quae oriuntur, dum pro singulis potestatibus $x^\alpha$, $x'^\beta$, $x''^\gamma$, etc. ipsarum valores medii substituuntur, cuius theorematis gravissimi veritas etiam ex aliis considerationibus facili derivari potuisset.

the mean value of $y$, which comes from the integral $\int y\psi y.\mathrm{d}y$ taken over all the values that $y$ can attain. Now it is obvious that we should set $\psi y = 0$ for all values that $y$ cannot attain—either because of the nature of the function (e.g., negative values of $y$ when $y = xx + x'x' + x''x'' + \text{etc.}$) or because the errors $x$, $x'$, $x''$, etc. lie within fixed limits. But this amounts to extending the integral to all values of $y$, that is from $y = -\infty$ to $y = +\infty$. Now the integral $\int y\psi y.\mathrm{d}y$ between fixed limits, say from $y = \eta$ to $y = \eta'$, is equal to the integral

$$\int y\varphi.f(y, x', x'', \ldots).\tfrac{\mathrm{d}f(y,x',x'',\ldots)}{\mathrm{d}y}.\varphi x'.\varphi x'' \ldots \mathrm{d}y.\mathrm{d}x'.\mathrm{d}x'' \ldots,$$

where the integration extends from $y = \eta$ to $y = \eta'$ and over all values of $x'$, $x''$, etc. corresponding to real values of $f(y, x', x'', \ldots)$. This is the same as the integral

$$\int y\varphi x.\varphi x'.\varphi x'' \ldots \mathrm{d}x.\mathrm{d}x'.\mathrm{d}x'' \ldots,$$

in which $y$ takes the value determined by $x$, $x'$, $x''$, etc., which in turn range over all values corresponding to values of $y$ between $\eta$ and $\eta'$. From this we conclude that the integral $\int y\psi y.\mathrm{d}y$ taken over all values of $y$ from $y = -\infty$ to $y = +\infty$ may be obtained from the integral

$$\int y\varphi x.\varphi x'.\varphi x'' \ldots \mathrm{d}x.\mathrm{d}x'.\mathrm{d}x'' \ldots,$$

taken from $x = -\infty$ to $x = +\infty$, from $x' = -\infty$ to $x' = +\infty$, etc.

## 14.

If the function $y$ reduces to a sum of terms of the form

$$Ax^{\alpha}x'^{\beta}x''^{\gamma} \ldots,$$

then the value of the integral $\int y\psi y.\mathrm{d}y$ taken over all values of $y$—the mean value of $y$—is equal to the sum of the terms

$$A \times \int x^{\alpha}\varphi x.\mathrm{d}x \times \int x'^{\beta}\varphi x'.\mathrm{d}x' \times \int x''^{\gamma}\varphi x''.\mathrm{d}x'' \ldots,$$

where the integration extends from $x = -\infty$ to $x = +\infty$, from $x' = -\infty$ to $x' = +\infty$, etc. Equivalently, it is the sum of terms in which the powers $x^{\alpha}$, $x'^{\beta}$, $x''^{\gamma}$, etc. have been replaced by their mean values. The truth of this important theorem could easily have been established from other considerations.

## 15.

Applicemus ea, quae in art. prae. exposuimus, ad casum specialem, ubi

$$y = \frac{xx+x'x'+x''x''+\text{etc.}}{\sigma}$$

denotante $\sigma$ multitudinem partium in numeratore. Valor medius ipsius $y$ hic illico invenitur $= mm$, accipiendo characterem $m$ in eadem significatione ac supra. Valor verus quidem ipsius $y$ in casu determinato maior minorve evadere potest medio, perinde ac valor versus termine simplicis $xx$: sed probabilitas quod valor fortuitus ipsius $y$ a medio $mm$ haud sensibiliter aberret, continuo magis ad certitudinem appropinquabit crescente multitudine $\sigma$. Quod quo clarius eluceat, quum probabilitatem ipsam exacte determinare non sit in potestate, investigemus errorm medium metuendum, dum supponiums $y = mm$. Manifest per principia art. 6 hic error erit radix quadrata valoris medii functionis

$$\left(\frac{xx+x'x'+x''x''+\text{etc.}}{\sigma} - mm\right)^2$$

ad quem eruendum sufficit observare, valorem medium termine talis $\frac{x^4}{\sigma\sigma}$ esse $\frac{n^4}{\sigma\sigma}$ (utendo charactere $n$ in significatione art. 11), valorem medium autem termini talis $\frac{2xxx'x'}{\sigma\sigma}$ fieri $\frac{2m^4}{\sigma\sigma}$, unde facillime deducitur valor medius istius functionis

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

Hinc discimus, si copia satis magna errorum fortuitorum ab invicem independentium $x$, $x'$, $x''$, etc. in promtu sit, magna certitudine inde peti posse valorem approximatum ipsius $m$ per formulam

$$m = \sqrt{\frac{xx+x'x'+x''x''+\text{etc.}}{\sigma}}$$

erroremque medium in hac determinatione metuendum, respectu quadrati $mm$, esse

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

Ceterum, quum posterior formula implicet quantitatem $n$, si id tantum agitur, ut idea qualiscunque de gradu praecisionis istius determinationis formari possit, sufficiet, aliquam hypothesin respectu functionis $\varphi$ amplecti. E.g. in hypothesi tertia art. 9, 11 iste error fit $= mm\sqrt{\frac{2}{\sigma}}$. Quod si minus arridet, valor approxmatus ipsius $n^4$ ex ipsis erroribus adiumento formula

$$\frac{x^4+x'^4+x''^4+\text{etc.}}{\sigma}$$

15.

Let us apply the results of the preceding article to the special case

$$y = \frac{xx+x'x'+x''x''+\text{etc.}}{\sigma},$$

where $\sigma$ is the number of terms in the numerator. The mean value of $y$ is readily found to be $mm$, where $m$ has its usual meaning. In any particular case the actual value of $y$ may turn out to be larger or smaller than its mean value—just as the actual value of the simple term $xx$. But as $\sigma$ grows, the probability that a random value of $y$ will not deviate appreciably from its mean $mm$ increases toward certainty. Because we are unable to determine the exact probabilities, let us clarify the matter by investigating the error to be feared when we assume that $y = mm$. According to the principles of Art. 6, this error is the square root of the mean value of the function

$$\left(\frac{xx+x'x'+x''x''+\text{etc.}}{\sigma} - mm\right)^2.$$

To compute this mean value, it is sufficient to observe that the mean value of a term like $\frac{x^4}{\sigma\sigma}$ is $\frac{n^4}{\sigma\sigma}$ (here $n$ has the same meaning as in Art. 11) and the mean value of a term like $\frac{2xxx'x'}{\sigma\sigma}$ is $\frac{2m^4}{\sigma\sigma}$. From this we easily find that the mean value of the function is

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

Thus if we have a sufficiently large number of mutually independent random errors $x$, $x'$, $x''$, etc., we are quite safe in using the formula

$$m = \sqrt{\frac{xx+x'x'+x''x''+\text{etc.}}{\sigma}}$$

to obtain an approximate value of $m$. The mean error to be feared in this estimate (with respect to the square $mm$) is

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

This last formula contains the quantity $n$. If we merely wish to form a rough idea of the precision of the estimate, it suffices to make some hypothesis about the function $\varphi$. For example, under the third hypotheses in Arts. 9 and 11, the error is $mm\sqrt{\frac{2}{\sigma}}$. If this is unsatisfactory, one can instead approximate the value of $n^4$ by the formula

$$\frac{x^4+x'^4+x''^4+\text{etc.}}{\sigma}.$$

peti poterit. Generaliter autem affirmare possumus, praecisionem duplicatam in ista determinatione requirere errorum copiam quadruplicatam, sive pondus determinationis ipsi multitudini $\sigma$ esse proportionale.

Prorsus simili modo, si observationum errores partem constantem involvunt, huius valor approximatus eo tutius e medio arithmetico multorum errorum colligi poterit, quo maior horum multitudo fuerit. Et quidem error medius in hac determinatione metuendus exprimetur per

$$\sqrt{\frac{mm-kk}{\sigma}}$$

si $k$ designat partem constantem ipsam atque $m$ errorem medium observationum parte constante nondum purgatarum, sive simpliciter per $\frac{m}{\sqrt{\sigma}}$, si $m$ denotat errorem medium observationum a parte constante liberatarum (v. art. 8).

16.

In artt. 12–15 supposuimus, errores $x$, $x'$, $x''$, etc. ad idem observationum geneus pertinere, ita ut singulorum probabilitates per eandem functionem exprimantur. Sed sponte patet, disquisitionem generalem artt. 12–14 aeque facile ad casum generaliorem extendi, ubi probabilitates errorem $x$, $x'$, $x''$, etc. per functioniones diversas $\varphi x$, $\varphi' x'$, $\varphi'' x''$, etc. exprimantur, i.e. ubi errores illi pertineant ad observationes praecisionis seu incertitudinis diversae. Supponamus, $x$ esse errorem observationis talis, cuius error medius metuendus sit $= m$; nec non $x'$, $x''$, etc. esse errores aliarum observationum, quarum errores medii metuendi resp. sint $m'$, $m''$, etc. Tunc valor medius aggretati $xx + x'x' + x''x'' +$ etc. erit $mm + m'm' + m''m'' +$ etc. Iam si aliunde constat, quantitates $m$, $m'$, $m''$, etc. esse in ratione data, puta numeris 1, $\mu'$, $\mu''$, etc. resp. proportionales, valor medius expressionis

$$\frac{xx+x'x'+x''x''+\text{etc.}}{1+\mu'\mu'+\mu''\mu''+\text{etc.}}$$

erit $= mm$. Si vero valorem eiusdem expressionis determinatum, prout fors errors $x$, $x'$, $x''$, etc. offert, ipsi $mm$ aequalem ponimus, error medius, cui haec determinatio obnoxia manet, simili ratione ut in art. praec. invenitur

$$= \frac{\sqrt{n^4+n'^4+n''^4+\text{etc.}-m^4-m'^4-m''^4-\text{etc.}}}{1+\mu'\mu'+\mu''\mu''+\text{etc.}}$$

ubi $n'$, $n''$, etc. respectu observationum, ad quas pertinent errores $x'$, $x''$, etc., idem denotare supponuntur, atque $n$ respectu observationis primae. Quodsi itaque

In general we can say that to get twice the precision in an estimate of this kind, we need four times the number of errors. In other words, the weight of the estimate is proportional to the number $\sigma$.

Similarly, when the observation errors have a constant part, its value can be approximated by the arithmetic mean of the errors. The larger the number of observations, the more reliable this approximation will be. In particular, if $k$ is the constant part of the observations and $m$ is the mean error of the observations from which the constant part has not yet been removed, then the mean error in the arithmetic mean is given by

$$\sqrt{\frac{mm-kk}{\sigma}}.$$

If $m$ is the mean error in the observations from which the constant part has been removed (see Art. 8), then it is just $\frac{m}{\sqrt{\sigma}}$.

16.

In Arts. 12–15 we assumed that the errors $x$, $x'$, $x''$, etc. were all of the same class, so that their individual probabilities could be represented by the same function. But clearly the derivations of Arts. 12–14 are easily extended to the case where the probabilities of the errors $x$, $x'$, $x''$, etc. are represented by different functions $\varphi x$, $\varphi' x'$, $\varphi'' x''$, etc., that is, to the case where the errors come from observations of different precision or uncertainty.

Let us suppose that $x$ is a observation error whose mean error is $m$. Similarly, suppose that $x'$, $x''$, etc. are observation errors whose mean errors are $m'$, $m''$, etc. Then the mean value of the sum $xx+x'x'+x''x''+\text{etc.}$ is $mm+m'm'+m''m''+\text{etc.}$ If in addition we know from some source that the quantities $m$, $m'$, $m''$, etc. are in fixed ratios, say 1, $\mu'$, $\mu''$, etc., then the mean value of the expression

$$\frac{xx+x'x'+x''x''+\text{etc.}}{1+\mu'\mu'+\mu''\mu''+\text{etc.}}$$

will be $mm$. If the value of this expression is regarded as equal to $mm$, then the value determined by the random the errors $x$, $x'$, $x''$, etc. is subject to the mean error

$$\frac{\sqrt{n^4+n'^4+n''^4+\text{etc.}-m^4-m'^4-m''^4-\text{etc.}}}{1+\mu'\mu'+\mu''\mu''+\text{etc.}}.$$

Here $n'$, $n''$, etc. denote the same quantity with respect to $x'$, $x''$, etc. as $n$ does with respect to $x$, and the derivation is similar to that of the preceding article.

numeros $n$, $n'$, $n''$, etc. ipsis $m$, $m'$, $m''$, etc. proportionales supponere licet, error ille metuendus medius fit

$$=\frac{\sqrt{n^4-m^4}\sqrt{1+\mu'^4+\mu''^4+\text{etc.}}}{1+\mu'\mu'+\mu''\mu''+\text{etc.}}$$

At haecce ratio, valorem approximatum ipsius $m$ determinandi non est ea, quae maxime ad rem facit. Quod quo clarius ostendamus, consideremus expressionem generaliorem

$$y=\frac{xx+\alpha'x'x'+\alpha''x''x''+\text{etc.}}{1+\alpha'\mu'\mu'+\alpha''\mu''\mu''+\text{etc.}}$$

cuius valor medius quoque erit $= mm$, quomodocunque eligantur coëfficientes $\alpha'$, $\alpha''$, etc. Error autem medius metuendus, dum valorem determinatum ipsius $y$, prout fors errores $x$, $x'$, $x''$, etc. offert, ipsi $mm$ aequalem supponimus, invenitur per principia supra tradita

$$=\frac{\sqrt{n^4-m^4+\alpha'\alpha'(n'^4-m'^4)+\alpha''\alpha''(n''^4-m''^4)+\text{etc.}}}{1+\alpha'\mu'\mu'+\alpha''\mu''\mu''+\text{etc.}}$$

Ut hic error medius fiat quam minimus, statuere oportebit

$$\begin{aligned}\alpha' &= \frac{n^4-m^4}{n'^4-m'^4}\cdot\mu'\mu'\\ \alpha'' &= \frac{n^4-m^4}{n''^4-m''^4}\cdot\mu''\mu'' \quad \text{etc.}\end{aligned}$$

Manifesto hi valores evolvi nequeunt, nisi insuper ratio quantitatum $n$, $n'$, $n''$, etc. ad $m$, $m'$, $m''$, etc. aliunde nota fuerit; qua cognitione exacta deficiente, saltem tutissimum videtur,*) illas his proportionales supponere (v. art. 11), unde prodeunt valores

$$\alpha'=\frac{1}{\mu'\mu'},\quad \alpha''=\frac{1}{\mu''\mu''}\ \text{etc.}$$

i.e., coëfficientes $\alpha'$, $\alpha''$, etc. aequales statui debent ponderibus relativis observationum, ad quas pertinent errores $x'$, $x''$, etc., assumto pondere observationis, ad quam pertinet error $x$, pro unitate. Hoc pacto, designante ut supra $\sigma$ multitudinem errorum propositorum, habebimus valorem medium expressionis

$$\frac{xx+\alpha'x'x'+\alpha''x''x''+\text{etc.}}{\sigma}$$

---

*)Scilicet cognitionem quantitatum $\mu$, $\mu'$, etc. in eo solo casu in potestate esse concipimus, ubi per rei naturam errores $x'$, $x''$, etc. ipsis 1, $\mu$, $\mu'$, etc. proportionales, aeque probabiles censendi sunt, aut potius ubi

$$\varphi x=\mu'\varphi'(\mu'x)=\mu''\varphi''(\mu''x)\ \text{etc.}$$

Thus, if the numbers $n$, $n'$, $n''$, etc. are assumed to be proportional to the numbers $m$, $m'$, $m''$, etc., then the mean error becomes

$$\frac{\sqrt{n^4-m^4}\sqrt{1+\mu'^4+\mu''^4+\text{etc.}}}{1+\mu'\mu'+\mu''\mu''+\text{etc.}}.$$

However, the above quotient is not the best estimate of $m$. To make this clearer, consider the general expression

$$y = \frac{xx+\alpha'x'x'+\alpha''x''x''+\text{etc.}}{1+\alpha'\mu'\mu'+\alpha''\mu''\mu''+\text{etc.}},$$

whose mean value is also $mm$, however the coefficients $\alpha'$, $\alpha''$, etc. are chosen. If we assume that the value of $y$ is $mm$, then by the above principles the mean error in the value determined by the random values of $x$, $x'$, $x''$, etc. is

$$\frac{\sqrt{n^4-m^4+\alpha'\alpha'(n'^4-m'^4)+\alpha''\alpha''(n''^4-m''^4)+\text{etc.}}}{1+\alpha'\mu'\mu'+\alpha''\mu''\mu''+\text{etc.}}.$$

In order to make this error as small as possible, we must set

$$\alpha' = \frac{n^4-m^4}{n'^4-m'^4}\cdot\mu'\mu',$$
$$\alpha'' = \frac{n^4-m^4}{n''^4-m''^4}\cdot\mu''\mu'', \quad \text{etc.}$$

Obviously these values cannot be calculated unless we also know the ratios of $n$, $n'$, $n''$, etc. to $m$, $m'$, $m''$, etc. If they are unknown, it would seem safest*) to assume that the $m$'s are proportional to the $n$'s (see Art. 11), in which case

$$\alpha' = \frac{1}{\mu'\mu'}, \quad \alpha'' = \frac{1}{\mu''\mu''}, \quad \text{etc.};$$

i.e., the coefficients $\alpha'$, $\alpha''$, etc. should be set equal to the relative weights of the observations having errors $x'$, $x''$, etc., assuming the weight of the observation corresponding to $x$ is one. In this case if $\sigma$ is the number of errors, the mean value of the expression

$$\frac{xx+\alpha'x'x'+\alpha''x''x''+\text{etc.}}{\sigma}$$

*)The only imaginable case where we can know the quantities $\mu$, $\mu'$, etc. is when by the nature of things the errors $x$, $x'$, $x''$, etc., which are proportional to 1, $\mu$, $\mu'$, etc., can be regarded as equally probable, or in other words when

$$\varphi x = \mu'\varphi'(\mu'x) = \mu''\varphi''(\mu''x) \text{ etc.}$$

$= mm$, atque errorem medium metuendum, dum valorem fortuito determinatum huius expressionis pro valore vero ipsius $mm$ adoptamus

$$\frac{\sqrt{n^4+\alpha'\alpha' n'^4+\alpha''\alpha'' n''^4+\text{etc.}-\sigma m^4}}{\sigma}$$

et proin, siquidem licet $n$, $n'$, $n''$, etc. ipsis $m$, $m'$, $m''$ proportionales supponere,

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

quae formula identica est cum ea, quam supra pro casu observation eiusdem generis inveneramus.

17.

Si valor quantitatis, quae ab alia quantitate incognita pendent, per obervationem praecisione absoluta non gaudentem determinata est, valor incognitae hinc calculatus etiam errori obnoxius erit, sed nihil in hac determinatione arbitrio relinquiter. At si *plures* quantitates ab eadem incognita pendentes per observationes haud absolute exactas innotuerunt, valorum incognitae vel per quamlibet harum observationum eruere possumus, vel per aliquam plurium observationum combinationem, quod infinitis modis diveris fieri potest. Quamquam vero valor incognitae tali modo prodiens errori semper obnoxius manet, tamen in alia combinatione maior, in alia minor error metuendus erit. Similiter res se habebit, si plures quantitates a pluribus incognitis simul pendentes sunt observatae: prout observationum multitudo multitudini incognitarum vel aequalis, vel hac minor, vel maior fuerit, problema vel determinatum, vel indeterminatum, vel plus quam determinatum erit (generaliter saltem loquendo), et in casu tertio ad incognitarum determinationem observationes infinitis modis diversis combinari poterunt. E tali combinationum varietate eas eligere, quai maxime ad rem faciant, i.e. quae incognitarum valores erroribus minimis obnoxios suppeditent, problema sane est in applicatione matheseos ad philosophiam naturalem longe gravissimum.

In Theoria motus corporum coelestium ostendimus, quomodo valores incognitarum *maxime probabiles* eruendi sint, si lex probabilitatis errorum observationum cognita sit; et quum haec lex natura sua in omnibus fere casibus hypothetica maneat, theorem illam ad legem maxime plausibilem applicavimus, ubi probabilitas erroris $x$ quantiti exponentiale $e^{-hhxx}$ proportionalis supponitur, unde methodus a nobis dudum in calculis praesertim astronomicis, et nunc quidem a plerisque calculatoribus sub nomine methodi quadratorum minimorum usitata demanavit.

is $mm$. The mean error to be feared when we accept the random value of the expression in place of the true value $mm$ is

$$\frac{\sqrt{n^4+\alpha'\alpha'n'^4+\alpha''\alpha''n''^4+\text{etc.}-\sigma m^4}}{\sigma}.$$

If we may assume that $n$, $n'$, $n''$, etc. are proportional to $m$, $m'$, $m''$, etc., then the mean error is

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

This formula is the same as the one we derived for observations of the same class.

## 17.

Suppose a quantity that depends on another unknown quantity is estimated by an observation that is not absolutely precise. If the unknown is calculated from this observation, it will also be subject to error, and there will be no freedom in this estimate of it. But if *several* quantities depending on the same unknown have been determined by inexact observations, we can recover the unknown either from one of the observations or from any of an infinite number of combinations of the observations. Although the value of an unknown determined in this way is always subject to error, there will be less error in some combinations than in others.

A similar situation occurs when we observe several quantities depending on several unknowns. The number of observations may be equal to, less than, or greater than the number of unknowns. In the first case the problem is well determined; in the second it is indeterminate. In the third case the problem is (generally speaking) overdetermined, and the observations can be combined in an infinite number of ways to estimate the unknowns. One of the most important problems in the application of mathematics to the natural sciences is to choose the best of these many combinations, i.e., the combination that yields values of the unknowns that are least subject to the errors.

In my *Theory* of the motion of heavenly bodies I showed how to calculate *most probable* values of the unknowns, provided the probability law of the observation errors is known. But in almost all cases this law can only be hypothetical, and for this reason I applied the theory to the most plausible law, in which the probability of an error $x$ is proportional to $e^{-hhxx}$. From this supposition came a method which I had already used for some time, especially in astronomical calculations. It is now used by many calculators under the name of the method of least squares.

Postea ill. LAPLACE, rem alio modo aggressus, idem principium omnibus aliis etiamnum praeferendum esse docuit, quaecunque fuerit lex probabilitatis errorum, si modo multitudo sit permagna. At pro multitudine observationum modica, res intacta mansit, ita ut si lex nostra hypothetica respuatur, methodus quadratorum minimorum eo tantum nomine prae aliis commendabilis habenda sit, quod calculorum concinnitati maxime est adaptata.

Geometris itaque gratum fore speramus, si in hac nova argumenti tractatione docuerimus, methodum quadratorum minimorum exhibere combinationem ex omnibus optimam, non quidem proxime, sed absolute, quaecunque fuerit lex probabilitatis errorum, quaecunque observationum multitudo, si modo notionem erroris medii non ad menterm ill. LAPLACE set ita, ut in artt. 5 et 6 a nobis factum est, stabiliamus.

Ceterum expressis verbis hic praemonere convenit, in omnibus disquisitionibus sequentibus tantummodo de erroribus irregularibus atque a parte constante liberis sermonem esse, quum proprie ad perfectam artem observandi pertineat, omnes errorum constantium caussas summo studio amovere. Quaenam vero subsidia calculator tales observationes tractare suscipiens, quas ab erroribus constantibus non liberas esse iusta suspicio adest, ex ipso calculo probabilium petere possit, disquisitioni peculiari alia occasione promulgandae reservamus.

## 18.

PROBLEMA. *Designante $U$ functionem datam quantitatum incognitarum $V$, $V'$, $V''$, etc., quaeritur error medius $M$ in determinatione valoris ipsius $U$ metuendus, si pro $V$, $V'$, $V''$, etc. adoptentur non valores veri, sed ii, qui ex observationibus ab invicem independentibus, erroribus mediis $m$, $m'$, $m''$, etc. resp. obnnoxiis prodeunt.*

*Sol.* Denotatis erroribus in valoribus observatis ipsarum $V$, $V'$, $V''$, etc. per $e$, $e'$, $e''$, etc., error inde redundans in valorem ipsius $U$ exprimi poterit per functionem linearem

$$\lambda e + \lambda' e' + \lambda'' e'' + \text{etc.} = E$$

ubi $\lambda$, $\lambda'$, $\lambda''$, etc. sunt valores quotientium differentialium $\frac{\mathrm{d}U}{\mathrm{d}V}$, $\frac{\mathrm{d}U}{\mathrm{d}V'}$, $\frac{\mathrm{d}U}{\mathrm{d}V''}$, etc. pro valoribus veris ipsarum $V$, $V'$, $V''$, etc., siquidem observationes satis exactae sunt, ut errorum quadrata productaque negligere liceat. Hinc primo sequitur, quoniam observationum errores a partibus constantibus liberi supponuntur, valorem medium ipsius $E$ esse $= 0$. Porro error medius in valore ipsius $U$ metuendus erit

Later LAPLACE attacked the problem from a different angle and showed that if the number of observations is very large then the method of least squares is to be preferred, whatever the probability law of the errors. But for a modest number of observations, things are as they were, and if one rejects my hypothetical law, the only reason for recommending the method of least squares over other methods is that it lends itself to easy calculation.

I therefore hope that mathematicians will be grateful if in this new treatment of the subject I show that the method of least squares gives the best of all combinations — not approximately, but absolutely, whatever the probability law of the errors and whatever the number of observations — provided only that we take the notion of mean error not in the sense of LAPLACE but as in Arts. 5 and 6.

Here we should say that in the sequel we will be concerned only with random errors having no constant part, since the craft of taking observation requires that we take pains to remove all causes of constant errors. On another occasion I will give a special treatment about what help a calculator can expect from the calculus of probabilities when he undertakes to treat observations he suspects are not free of constant errors.

18.

PROBLEM. *Given a function $U$ of the unknown quantities $V$, $V'$, $V''$, etc., find the mean error $M$ to be feared in estimating $U$ when, instead of the true values of $V$, $V'$, $V''$, etc. one uses independently observed values having mean errors $m$, $m'$, $m''$, etc.*

*Solution.* Let $e$, $e'$, $e''$, etc. denote the errors in the observed values of $V$, $V'$, $V''$, etc., and let $\lambda$, $\lambda'$, $\lambda''$, etc. be the differential quotients $\frac{dU}{dV}$, $\frac{dU}{dV'}$, $\frac{dU}{dV''}$, etc. at the true values of $V$, $V'$, $V''$, etc. Then the resulting error in $U$ can be represented by the linear function

$$\lambda e + \lambda' e' + \lambda'' e'' + \text{etc.} = E,$$

provided the observations are precise enough so that we can neglect squares and products of the errors. From this it follows first that the mean value of $E$ is zero, since the observation errors are assumed to have no constant parts. Moreover, the mean error to be feared in this value of $U$ is the square root of the of the mean radix quadrata e valore medio ipsius $EE$, sive $MM$ erit valor medius aggregati

$$\lambda\lambda ee + \lambda'\lambda' e'e' + \lambda''\lambda'' e''e'' + \text{etc.} + 2\lambda\lambda' ee' + 2\lambda\lambda'' ee'' + 2\lambda'\lambda'' e'e'' + \text{etc.}$$

At valor medius ipsius $\lambda\lambda ee$ fit $\lambda\lambda mm$, valor medius ipsius $\lambda'\lambda' e'e'$ fit $\lambda'\lambda' m'm'$, etc.; denique valores medii productorum $2\lambda\lambda' ee'$, etc. omnes fiunt $= 0$. Hinc itaque colligimus

$$M = \sqrt{\lambda\lambda mm + \lambda'\lambda' m'm' + \lambda''\lambda'' m''m'' + \text{etc.}}$$

Huic solutioni quasdam annotationes adiicere conveniet.

I. Quatenus spectando observationum errores tamquam quantitates primi ordinis, quantitates ordinum altiorum negliguntur, in formula nosta pro $\lambda$, $\lambda'$, $\lambda''$, etc. etiam valores eos quotientium $\frac{\mathrm{d}U}{\mathrm{d}V}$, etc. adoptare licebit, qui prodeunt e valoribus observatis quantitatum $V$, $V'$, $V''$, etc. Quoties $U$ est functio linearis, manifesto nulla prorsus erit differentia.

II. Si loco errorum mediorum observationum, harum pondera introducere malumus, sint haec, secundum unitatem arbitrariam, resp. $p$, $p'$, $p''$, etc., atque $P$ pondus determinationis valoris ipsius $U$ e valoribus observatis quantitatum $V$, $V'$, $V''$, etc. prodeuntis. Ita habebimus

$$P = \frac{1}{\frac{\lambda\lambda}{p} + \frac{\lambda'\lambda'}{p'} + \frac{\lambda''\lambda''}{p''} + \text{etc.}}$$

III. Si $T$ is functio alia data quantitatum $V$, $V'$, $V''$, etc. atque, pro harum valoribus veris

$$\frac{\mathrm{d}T}{\mathrm{d}V} = \kappa, \quad \frac{\mathrm{d}T'}{\mathrm{d}V'} = \kappa', \quad \frac{\mathrm{d}T''}{\mathrm{d}V''} = \kappa'' \ \text{ etc.}$$

error in determinatione valor ipsius $T$, e valoribus observatis ipsarum $V$, $V'$, $V''$, etc. petita, erit $= \kappa e + \kappa' e' + \kappa'' e'' + \text{etc.}, = E'$ atque error medius in ista determinatione metuendus $= \sqrt{\kappa\kappa mm + \kappa'\kappa' m'm' + \kappa''\kappa'' m''m'' + \text{etc.}}$. Errores $E$, $E'$ vero manifesto ab invicem iam non erunt independentes, valorque medius producti $EE'$, secus ac valor medius producti $ee'$, non erit $= 0$, sed $= \kappa\lambda mm + \kappa'\lambda' m'm' + \kappa''\lambda'' m''m'' + \text{etc.}$

IV. Problema nostrum etiam ad casum eum extendere licet, ubi valores quantitatum $V$, $V'$, $V''$, etc. non immediate per observationes inveniuntur, sed quomodocunque ex observationum combinationibus derivantur, si modo singularum determinationes ab invicem sunt independentes, i.e. observationibus diversis superstructae: quoties autem haec conditio locum non habet, formula pro $M$ erronea evaderet. E.g. si una alterave observatio, quae ad determinationem valoris ipsius value of $EE$; that is, $MM$ is the mean value of the sum

$$\lambda\lambda ee + \lambda'\lambda' e'e' + \lambda''\lambda'' e''e'' + \text{etc.} + 2\lambda\lambda' ee' + 2\lambda\lambda'' ee'' + 2\lambda'\lambda'' e'e'' + \text{etc.}$$

Now the mean value of $\lambda\lambda ee$ is $\lambda\lambda mm$, the mean value of $\lambda'\lambda' e'e'$ is $\lambda'\lambda' m'm'$, etc. The mean values of the products $2\lambda\lambda' ee'$, etc. are all zero. Hence it follows that

$$M = \sqrt{\lambda\lambda mm + \lambda'\lambda' m'm' + \lambda''\lambda'' m''m'' + \text{etc.}}$$

It is appropriate to append some comments to this solution.

I. Since we have taken the observation errors to be quantities of the first order and have neglected quantities of higher orders, we may use the values of the differential quotients $\frac{\mathrm{d}U}{\mathrm{d}V}$, etc. that come from the observed quantities $V$, $V'$, $V''$, etc. to evaluate our formula instead of $\lambda$, $\lambda'$, $\lambda''$, etc. Obviously this substitution makes no difference at all when $U$ is a linear function.

II. Let $p$, $p'$, $p''$, etc. be the weights of the observation errors with respect to an arbitrary unit, and let $P$ be the weight of the estimate of $U$ derived from the observed quantities $V$, $V'$, $V''$, etc. If we prefer to work in terms of these quantities rather than the mean errors, then we have

$$P = \frac{1}{\frac{\lambda\lambda}{p} + \frac{\lambda'\lambda'}{p'} + \frac{\lambda''\lambda''}{p''} + \text{etc.}}.$$

III. Let $T$ be another function of the quantities $V$, $V'$, $V''$, etc., and for the true value of these quantities let

$$\frac{\mathrm{d}T}{\mathrm{d}V} = \kappa, \quad \frac{\mathrm{d}T'}{\mathrm{d}V'} = \kappa', \quad \frac{\mathrm{d}T''}{\mathrm{d}V''} = \kappa'', \quad \text{etc.}$$

Then the error in the estimate for $T$ obtained from the observed values $V$, $V'$, $V''$, etc. is $E' = \kappa e + \kappa' e' + \kappa'' e'' + \text{etc.}$, and the error to be feared in this estimate is $\sqrt{\kappa\kappa mm + \kappa'\kappa' m'm' + \kappa''\kappa'' m''m'' + \text{etc.}}$. The errors $E$ and $E'$ are clearly not independent, and, unlike the products $ee'$, the mean value of $EE'$ is not zero but $\kappa\lambda mm + \kappa'\lambda' m'm' + \kappa''\lambda'' m''m'' + \text{etc.}$

IV. Our problem also extends to the case where the quantities $V$, $V'$, $V''$, etc. are not obtained directly from observations but are derived from arbitrary combinations of observations. However, the individual quantities must be mutually independent, i.e., based on different observations. If this condition does not hold, then the formula for $M$ will be in error. For example, if some observation that

$V$ inserviit, etiam ad valorem ipsius $V'$ determinandum adhibita esset, errores $e$ et $e'$ haud amplius ab invicem independentes forent, neque adeo producti $ee'$ valor medius $= 0$. Si vero in tali casu nexus quantatitatum $V$, $V'$ cum observationibus simplicibus, e quibus deductae sunt, rite perpenditur, valor medius producte $ee'$ adiumento annotationis III. assignari, atque sic formula pro $M$ completa reddi poterit.

## 19.

Sint $V$, $V'$, $V''$, etc. functiones incognitarum $x$, $y$, $z$, etc., multitudo illarum $= \pi$, multitudo incognitarum $= \rho$, supponamusque, per observationes vel immediate vel mediate valores functionum inventos esse $V = L$, $V' = L'$, $V'' = L''$, etc., ita tamen ut hae determinationes ab invicem fuerint independentes. Si $\rho$ maior est quam $\pi$ incognitarum evolutio manifest fit problema indeterminatum; si $\rho$ ipsi $\pi$ aequalis est, singulae $x$, $y$, $z$, etc. in formam functionum ipsarum $V$, $V'$, $V''$, etc. redigi vel redactae concipi possunt, ita ut ex harum valoribus observatis valores istarum inveniri possint, simulque adiumento art. praec. praecisionem relativam singulis his determinationibus tribuendam assignare liceat; denique si $\rho$ minor est quam $\pi$, sungulae $x$, $y$, $z$, etc. infinitis modis diversis in formam functionum ipsarum $V$, $V'$, $V''$, etc. redigi, adeoque illarum valores infinitis modis diversis erui poterunt. Quae determinationes exacte quidem quadrare deberent, si observationes praecisione absoluta gauderent; quod quum secus se habeat, alii modi alios valores suppeditabunt, nec minus determinationes e combinationibus diversis petitae inaequali praecisione instructae erunt.

Ceterum si in casu secundo vel tertio functiones $V$, $V'$, $V''$, etc. ita comparatae essent, ut $\pi - \rho + 1$ ex ipsis, vel plures, tamquam functiones reliquarum spectare liceret, problema respectu posteriorum functionum etiamnum plus quam determinatum esset, respectu incognitarum $x$, $y$, $z$, etc. autem indeterminatum; harum scilicet valores ne tunc quidem determinare liceret, quando valores functionum $V$, $V'$, $V''$, etc. absolute exacti dati essent: sed hunc casum a disquisitione nostra excludemus.

Quoties $V$, $V'$, $V''$, etc. per se non sunt functiones *lineares* indeterminatarum suarum, hoc efficietur, si loco incognitarum primitivarum introducuntur ipsarum differentiae a valoribus approximatis, quos aliunde cognitos esse suponere licet. Errores medios in determinationibus $V = L$, $V' = L'$, $V'' = L''$, etc. metuendos resp. denotabimus per $m$, $m'$, $m''$, etc., determinationumque pondera per $p$, $p'$, $p''$, etc., ita ut sit $pmm = p'm'm' = p''m''m''$ etc. Rationem, quam inter se tenent was involved in the calculation of $V$ is also used in the calculation of $V'$, the errors $e$ and $e'$ will no longer be independent, and the mean value of the product $ee'$ will not be zero. However, if we can ascertain the relation of $V$ and $V'$ with the simple observations from which they were derived, we can determine the mean value of the product $ee'$ by the methods of comment III and hence give the correct formula for $M$.

## 19.

Let $V$, $V'$, $V''$, etc. be $\pi$ functions of $\rho$ unknowns $x$, $y$, $z$, etc. Suppose that by means of observations we have, directly or indirectly, found mutually independent values $V = L$, $V' = L'$, $V'' = L''$, etc. of the functions. If $\rho$ is greater than $\pi$, the problem of extracting the unknowns is obviously indeterminate. If $\rho$ is equal to $\pi$, we can, at least conceptually, reduce $x$, $y$, $z$, etc. to functions of $V$, $V'$, $V''$, etc., so that the unknowns may be found from observations. At the same time we can determine the relative precisions of each unknown by the methods of the preceding article. Finally, if $\rho$ is less than $\pi$, then there are infinitely many was of reducing $x$, $y$, $z$, etc. to functions of $V$, $V'$, $V''$, etc. and thereby calculating their values. All these calculations would agree exactly if the observations were absolutely precise; otherwise, different expressions will give different values, and the estimates from these various combinations will have different precisions.

Now in the second and third cases, if $\pi - \rho + 1$ or more of the functions $V$, $V'$, $V''$, etc. could be regarded as functions of the rest, then the problem would still be overdetermined with respect the functions $V$, $V'$, $V''$, etc. but would be indeterminate with respect to the unknowns $x$, $y$, $z$, etc., whose values could not be determined even if the values of the functions $V$, $V'$, $V''$, etc. were given exactly. However, we will exclude this case from our inquiry.

Whenever $V$, $V'$, $V''$, etc. are not *linear* functions of the unknowns, we can make them effectively so by replacing the original unknowns with their differences from approximations obtained from other sources. We will denote the mean errors to be feared in the estimates $V = L$, $V' = L'$, $V'' = L''$, etc. by $m$, $m'$, $m''$, etc., and the corresponding weights by $p$, $p'$, $p''$, etc., so that $pmm = p'm'm' = p''m''m''$, etc. We will assume that the ratios among the mean errors are known, so that errores medii, cognitam supponemus, ita ut pondera, quorum unum ad lubitum accipi potest, sint nota. Denique statuemus

$$(V-L)/\sqrt{p}=v, \quad (V'-L')/\sqrt{p'}=v', \quad (V''-L'')/\sqrt{p''}=v'' \quad \text{etc.}$$

Manifesto itaque res perinde se habebit, ac si observationes immediatae, aequali pracisione gaudentes, puta quarum error medius $= m\sqrt{p} = m'\sqrt{p'} = m''\sqrt{p''}$, etc., sive quibus pondus $= 1$ tribuitur, suppeditavissent

$$v=0, \quad v'=0, \quad v''=0, \quad \text{etc.}$$

20.

PROBLEMA. *Designantibus $v$, $v'$, $v''$, etc. functiones lineares indeterminatarum $x$, $y$, $z$, etc. sequentes*

$$\left.\begin{aligned} v &= ax + by + cz + \text{etc.} + l \\ v' &= a'x + b'y + c'z + \text{etc.} + l' \\ v'' &= a''x + b''y + c''z + \text{etc.} + l'' \text{ etc.} \end{aligned}\right\} \text{(I)}$$

*ex omnibus systematibus coëfficientium $\kappa$, $\kappa'$, $\kappa''$, etc., qui indefinite dant*

$$\kappa v + \kappa' v' + \kappa'' v'' + \text{etc.} = x - k$$

*ita ut $k$ sit quantitas determinata i.e. ab $x$, $y$, $z$, etc. independens, eruere id, pro quo* $\kappa\kappa + \kappa'\kappa' + \kappa''\kappa'' + \text{etc.}$ *nanciscatur valorem minimum.*

*Solutio.* Statuamus

$$\left.\begin{aligned} av + a'v' + a''v'' + \text{etc.} &= \xi \\ bv + b'v' + b''v'' + \text{etc.} &= \eta \\ cv + c'v' + c''v'' + \text{etc.} &= \zeta \end{aligned}\right\} \text{(II)}$$

etc.: eruntque etiam $\xi$, $\eta$, $\zeta$, etc. functiones lineares ipsarum $x$, $y$, $z$, etc., puta

$$\left.\begin{aligned} \xi &= x\Sigma aa + y\Sigma ab + z\Sigma ac + \text{etc.} + \Sigma al \\ \eta &= x\Sigma ab + y\Sigma bb + z\Sigma bc + \text{etc.} + \Sigma bl \\ \zeta &= x\Sigma ac + y\Sigma bc + z\Sigma cc + \text{etc.} + \Sigma cl \text{ etc.} \end{aligned}\right\} \text{(III)}$$

(ubi $\Sigma aa$ denotat aggretum $aa + a'a' + a''a'' + \text{etc.}$, ac perinde de reliquis) multitudoque ipsarum $\xi$, $\eta$, $\zeta$, etc. multitudini indeterminatarum $x$, $y$, $z$, etc. aequalis, their weights are known whenever the weight of one is fixed. Finally we set

$$(V-L)/\sqrt{p}=v, \quad (V'-L')/\sqrt{p'}=v', \quad (V''-L'')/\sqrt{p''}=v'', \quad \text{etc.}$$

Obviously this is the same as if the direct observations, all having the same precisions $m\sqrt{p}=m'\sqrt{p'}=m''\sqrt{p''}$ etc. or the same weight one, provided the equations

$$v=0, \quad v'=0, \quad v''=0, \quad \text{etc.}$$

## 20.

PROBLEM. *Let $v$, $v'$, $v''$, etc. denote the following functions of the unknowns $x$, $y$, $z$, etc.:*

$$\left.\begin{aligned} v\ &= ax + by + cz + \text{etc.} + l \\ v' &= a'x + b'y + c'z + \text{etc.} + l' \\ v'' &= a''x + b''y + c''z + \text{etc.} + l'' \text{ etc.} \end{aligned}\right\} \quad \text{(I)}$$

*Of all systems of coefficients $\kappa$, $\kappa'$, $\kappa''$, etc. for which the general relation*

$$\kappa v + \kappa' v' + \kappa'' v'' + \text{etc.} = x - k$$

*holds for some constant $k$ independent of $x$, $y$, $z$, etc. find the one for which $\kappa\kappa + \kappa'\kappa' + \kappa''\kappa''$ + etc. is a minimum.*

*Solution.* Set

$$\left.\begin{aligned} av + a'v' + a''v'' + \text{etc.} &= \xi \\ bv + b'v' + b''v'' + \text{etc.} &= \eta \\ cv + c'v' + c''v'' + \text{etc.} &= \zeta \end{aligned}\right\} \quad \text{(II)}$$

etc. Then $\xi$, $\eta$, $\zeta$, etc. are linear functions of $x$, $y$, $z$, etc.; namely,

$$\left.\begin{aligned} \xi &= x\Sigma aa + y\Sigma ab + z\Sigma ac + \text{etc.} + \Sigma al \\ \eta &= x\Sigma ab + y\Sigma bb + z\Sigma bc + \text{etc.} + \Sigma bl \\ \zeta &= x\Sigma ac + y\Sigma bc + z\Sigma cc + \text{etc.} + \Sigma cl \text{ etc.} \end{aligned}\right\} \quad \text{(III)}$$

where $\Sigma aa$ denotes the sum $aa+a'a'+a''a''+$etc., and similarly for the rest. The number of unknowns $x$, $y$, $z$, etc. is the same as the number of quantities $\xi$, $\eta$, $\zeta$, puta $= \rho$. Per eliminationem itaque elici poterit aequatio talis*)

$$x = A + [\alpha\alpha]\xi + [\alpha\beta]\eta + [\alpha\gamma]\zeta + \text{etc.}$$

in qua substituendo pro $\xi$, $\eta$, $\zeta$, etc. valores earum ex III, aequatio identica prodire debet. Quare statuendo

$$\left.\begin{array}{l} a[\alpha\alpha] + b[\alpha\beta] + c[\alpha\gamma] + \text{etc.} = \alpha \\ a'[\alpha\alpha] + b'[\alpha\beta] + c'[\alpha\gamma] + \text{etc.} = \alpha' \\ a''[\alpha\alpha] + b''[\alpha\beta] + c''[\alpha\gamma] + \text{etc.} = \alpha'' \text{ etc.} \end{array}\right\} \text{(IV)}$$

necessario erit indefinite

$$\alpha v + \alpha' v' + \alpha'' v'' + \text{etc.} = x - A \quad \text{(V)}$$

Haec aequatio docet, inter systemata valorum coëfficientium $\kappa$, $\kappa'$, $\kappa''$, etc. certo etiam referendos esse hos $\kappa = \alpha$, $\kappa' = \alpha'$, $\kappa'' = \alpha''$, etc., nec non, pro systemate quocunque, fieri debere indefinite

$$(\kappa - \alpha)v + (\kappa' - \alpha')v' + (\kappa'' - \alpha'')v'' + \text{etc.} = A - k$$

quae aequatio implicat sequentes

$$\begin{array}{l} (\kappa - \alpha)a + (\kappa' - \alpha')a' + (\kappa'' - \alpha'')a'' + \text{etc.} = 0 \\ (\kappa - \alpha)b + (\kappa' - \alpha')b' + (\kappa'' - \alpha'')b'' + \text{etc.} = 0 \\ (\kappa - \alpha)c + (\kappa' - \alpha')c' + (\kappa'' - \alpha'')c'' + \text{etc.} = 0 \text{ etc.} \end{array}$$

Multiplicando has aequationes resp. per $[\alpha\alpha]$, $[\alpha\beta]$, $[\alpha\gamma]$, etc., et addendo, obtinemus propter (IV)

$$(\kappa - \alpha)\alpha + (\kappa' - \alpha')\alpha' + (\kappa'' - \alpha'')\alpha'' + \text{etc.} = 0$$

sive quod idem est

$$\begin{aligned} & \kappa\kappa + \kappa'\kappa' + \kappa''\kappa'' + \text{etc.} \\ = \; & \alpha\alpha + \alpha'\alpha' + \alpha''\alpha'' + \text{etc.} + (\kappa - \alpha)^2 + (\kappa' - \alpha')^2 + (\kappa'' - \alpha'')^2 + \text{etc.} \end{aligned}$$

unde patet, aggregatum $\kappa\kappa + \kappa'\kappa' + \kappa''\kappa'' +$ etc. valorem minimum obtinere, si statuatur $\kappa = \alpha$, $\kappa' = \alpha'$, $\kappa'' = \alpha''$, etc. Q.E.I.

*)Ratio, cur ad denotandos coëfficientes e tali eliminatione prodeuntes, hos potissimum characteres elegerimus, infra elucebit.

etc., namely, $\rho$. Hence by elimination[7] we obtain the equation*)

$$x = A + [\alpha\alpha]\xi + [\alpha\beta]\eta + [\alpha\gamma]\zeta + \text{etc.}$$

If we substitute the values of $\xi$, $\eta$, $\zeta$, etc. from (III) into this equation,[8] we must get an identity. Hence if we set

$$\left.\begin{array}{l} a[\alpha\alpha] \ + b[\alpha\beta] \ + c[\alpha\gamma] \ + \text{etc.} = \alpha \\ a'[\alpha\alpha] \ + b'[\alpha\beta] \ + c'[\alpha\gamma] \ + \text{etc.} = \alpha' \\ a''[\alpha\alpha] + b''[\alpha\beta] + c''[\alpha\gamma] + \text{etc.} = \alpha'' \text{ etc.} \end{array}\right\} \text{(IV)}$$

we will have in general

$$\alpha v + \alpha' v' + \alpha'' v'' + \text{etc.} = x - A. \quad \text{(V)}$$

This equation shows that the coefficients $\kappa = \alpha$, $\kappa' = \alpha'$, $\kappa'' = \alpha''$, etc. form one of our systems of coefficients $\kappa$, $\kappa'$, $\kappa''$, etc. Moreover, for any such system the equation

$$(\kappa - \alpha)v + (\kappa' - \alpha')v' + (\kappa'' - \alpha'')v'' + \text{etc.} = A - k$$

holds generally. This equation implies that

$$\begin{aligned} (\kappa - \alpha)a + (\kappa' - \alpha')a' + (\kappa'' - \alpha'')a'' + \text{etc.} &= 0 \\ (\kappa - \alpha)b + (\kappa' - \alpha')b' + (\kappa'' - \alpha'')b'' + \text{etc.} &= 0 \\ (\kappa - \alpha)c + (\kappa' - \alpha')c' + (\kappa'' - \alpha'')c'' + \text{etc.} &= 0 \text{ etc.} \end{aligned}$$

Multiplying these equations in turn by $[\alpha\alpha]$, $[\alpha\beta]$, $[\alpha\gamma]$, etc. and adding, we get from (IV)

$$(\kappa - \alpha)\alpha + (\kappa' - \alpha')\alpha' + (\kappa'' - \alpha'')\alpha'' + \text{etc.} = 0.$$

This is the same as

$$\begin{aligned} & \kappa\kappa + \kappa'\kappa' + \kappa''\kappa'' + \text{etc.} \\ = \ & \alpha\alpha + \alpha'\alpha' + \alpha''\alpha'' + \text{etc.} + (\kappa - \alpha)^2 + (\kappa' - \alpha')^2 + (\kappa'' - \alpha'')^2 + \text{etc.} \end{aligned}$$

---

[7]Gauss distinguishes between two types of elimination: *eliminatio definita*, in which a system of equations is solved, and *eliminatio indefinita*, in which the system is conceptually inverted to give coefficients of the inverse system. The latter will be translated as "general elimination."

*)It will become clear later why we chose particular symbols above to denote the coefficients from the elimination.

[8]Equation (II) seems to be meant here.

Ceterum hic valor minimus ipse sequenti modo eruitur. Aequatio (V) docet, esse

$$\alpha a + \alpha' a' + \alpha'' a'' + \text{etc.} = 1$$
$$\alpha b + \alpha' b' + \alpha'' b'' + \text{etc.} = 0$$
$$\alpha c + \alpha' c' + \alpha'' c'' + \text{etc.} = 0 \text{ etc.}$$

Multiplicando has aequationes resp. per $[\alpha\alpha]$, $[\alpha\beta]$, $[\alpha\gamma]$, etc. et addendo, protinus habemus adiumento aequationum (IV)

$$\alpha\alpha + \alpha'\alpha' + \alpha''\alpha'' + \text{etc.} = [\alpha\alpha]$$

## 21.

Quum observationes suppeditaverint aequationes (proxime veras) $v = 0$, $v' = 0$, $v'' = 0$, etc., ad valorem incognitae $x$ eliciendum, combinatio allarum aequationum talis

$$\kappa v + \kappa' v' + \kappa'' v'' + \text{etc.} = 0$$

adhibenda est, quae ipsi $x$ coëfficientem 1 conciliet, incognitasque reliquas $y$, $z$, etc. eliminet; cui determinationi per art. 18 pondus

$$= \frac{1}{\kappa\kappa + \kappa'\kappa' + \kappa''\kappa'' + \text{etc.}}$$

tribuendum erit. Ex art. praec. itaque sequitur, determinationem maxime idoneam eam fore, ubi statuatur $\kappa = \alpha$, $\kappa' = \alpha'$, $\kappa'' = \alpha''$, etc. Hoc pacto $x$ obtinet valorem $A$, manifestoque idem valor etiam (absque cognitione multiplicatorum $\alpha$, $\alpha'$, $\alpha''$, etc.) protinus per eliminationem ex aequationibus $\xi = 0$, $\eta = 0$, $\zeta = 0$, etc. elici potest. Pondus huic determinationi tribuendum erit $= \frac{1}{[\alpha\alpha]}$, sive error medius in ipsa metuendus

$$m\sqrt{p[\alpha\alpha]} = m'\sqrt{p'[\alpha\alpha]} = m''\sqrt{p''[\alpha\alpha]} \text{ etc.}$$

Prorsus simili modo determinatio maxime idonea incognitarum reliquarum $y$, $z$, etc. eosdem valores ipsis conciliabit, qui per elimination ex iisdem aequatinibus $\xi = 0$, $\eta = 0$, $\zeta = 0$, etc. prodeunt.

Denotando aggregatum indefinitum $vv + v'v' + v''v''$ etc., sive quod idem est hoc

$$p(V - L)^2 + p'(V' - L')^2 + p''(V'' - L'')^2 + \text{etc.}$$

from which it is obvious that the sum $\kappa\kappa + \kappa'\kappa' + \kappa''\kappa'' + \text{etc.}$ attains its minimum when $\kappa = \alpha$, $\kappa' = \alpha'$, $\kappa'' = \alpha''$, etc., which is what we set out to establish.

In addition, the minimum value is found as follows. Equations (V) imply that

$$\begin{aligned} \alpha a + \alpha' a' + \alpha'' a'' + \text{etc.} &= 1 \\ \alpha b + \alpha' b' + \alpha'' b'' + \text{etc.} &= 0 \\ \alpha c + \alpha' c' + \alpha'' c'' + \text{etc.} &= 0 \text{ etc.} \end{aligned}$$

Multiplying these equations by $[\alpha\alpha]$, $[\alpha\beta]$, $[\alpha\gamma]$, etc. and adding, we get from (IV)

$$\alpha\alpha + \alpha'\alpha' + \alpha''\alpha'' + \text{etc.} = [\alpha\alpha].$$

## 21.

Whenever observations supply (approximately true) equations $v = 0$, $v' = 0$, $v'' = 0$, etc., we can use combinations of the form

$$\kappa v + \kappa' v' + \kappa'' v'' + \text{etc.} = 0$$

to recover a value of $x$ from them. Such a combination must give $x$ a coefficient of one and eliminate the remaining unknowns $y$, $z$, etc. By Art. 18, a weight of

$$\frac{1}{\kappa\kappa + \kappa'\kappa' + \kappa''\kappa'' + \text{etc.}}$$

must be assigned to this estimate. From the preceding article it follows that the best estimate will be the one for which $\kappa = \alpha$, $\kappa' = \alpha'$, $\kappa'' = \alpha''$, etc. In this case $x$ assumes the value $A$. Obviously the same value could be obtained (without our knowing the multipliers $\alpha$, $\alpha'$, $\alpha''$, etc.) by elimination in the equations $\xi = 0$, $\eta = 0$, $\zeta = 0$, etc. The weight of this estimate will be $\frac{1}{[\alpha\alpha]}$, and the mean error to be feared in it will be

$$m\sqrt{p[\alpha\alpha]} = m'\sqrt{p'[\alpha\alpha]} = m''\sqrt{p''[\alpha\alpha]} \text{ etc.}$$

Similarly, the best estimates of the remaining unknowns $y$, $z$, etc. assign them the values obtained by elimination in the equations $\xi = 0$, $\eta = 0$, $\zeta = 0$, etc.

Let $\Omega$ denote the general sum $vv + v'v' + v''v''$ etc., or equivalently

$$p(V - L)^2 + p'(V' - L')^2 + p''(V'' - L'')^2 + \text{etc.}$$

per $\Omega$, patet $2\xi$, $2\eta$, $2\zeta$, etc. esse quotientes differentiales partiales functionis $\Omega$, puta

$$2\xi = \frac{\mathrm{d}\Omega}{\mathrm{d}x}, \quad 2\eta = \frac{\mathrm{d}\Omega}{\mathrm{d}y}, \quad 2\zeta = \frac{\mathrm{d}\Omega}{\mathrm{d}z}, \quad \text{etc.}$$

Quapropter valores incognitarum ex observationum combinatione maxime idonea prodeuntes, quos *valores maxime plausibles* commode voacare possumus, identici erunt cum iis, per quos $\Omega$ valorem minimum obtinet. Iam $V-L$ indefinite exprimit differentiam inter valorem computatum et observatum. Valores itaque incognitarum maxime plausibles iidem erunt, qui summam quadratorum differentiarum inter quantitatum $V$, $V'$, $V''$, etc. valores observatos et computatos, per observationum ponder multiplicorum, minimam efficiunt, quod principium in *Theoria Motus Corporum Coelestium* longe alia via stabiliveramus. Et si insuper praecisio relativa singularum determinationum assignanda est, per elimination indefinitam ex aequationibus (III) ipsas $x$, $y$, $z$ etc. in tali forma exhibere oportet:

$$\left.\begin{array}{l} x = A + [\alpha\alpha]\xi + [\alpha\beta]\eta + [\alpha\gamma]\zeta + \text{etc.} \\ y = B + [\beta\alpha]\xi + [\beta\beta]\eta + [\beta\gamma]\zeta + \text{etc.} \\ z = C + [\gamma\alpha]\xi + [\gamma\beta]\eta + [\gamma\gamma]\zeta + \text{etc.} \\ \text{etc.} \end{array}\right\} \text{(VII)}$$

quo pacto valores maxime plausibles incognitarum $x$, $y$, $z$, etc. erunt resp. $A$, $B$, $C$, etc., atque pondera his determinationibus tribuenda $\frac{1}{[\alpha\alpha]}$, $\frac{1}{[\beta\beta]}$, $\frac{1}{[\gamma\gamma]}$, etc., sive errores medii in ipsis metuendi

$$\begin{array}{ll} \text{for } x \ldots\ldots\ldots & = m\sqrt{p[\alpha\alpha]} = m'\sqrt{p'[\alpha\alpha]} = m''\sqrt{p''[\alpha\alpha]} \text{ etc.,} \\ \text{for } y \ldots\ldots\ldots & = m\sqrt{p[\beta\beta]} = m'\sqrt{p'[\beta\beta]} = m''\sqrt{p''[\beta\beta]} \text{ etc.,} \\ \text{for } z \ldots\ldots\ldots & = m\sqrt{p[\gamma\gamma]} = m'\sqrt{p'[\gamma\gamma]} = m''\sqrt{p''[\gamma\gamma]} \text{ etc.,} \\ \text{etc.} & \end{array}$$

quod convenit cum iis, quae in *Theoria Motus Corporum Coelestium* docuimus.

## 22.

De casu omnium simplicissimo, simul vero frequentissimo, ubi unica incognita adest, atque $V = x$, $V' = x$, $V'' = x$, etc., paucis seorsim agere conveniet. Erit scilicet $a = \sqrt{p}$, $a' = \sqrt{p'}$, $a'' = \sqrt{p''}$, etc., $l = -L\sqrt{p}$, $l' = -L'\sqrt{p'}$, $l'' = -L''\sqrt{p''}$, It is obvious that $2\xi$, $2\eta$, $2\zeta$, etc. are the partial differential quotients

$$2\xi = \frac{d\Omega}{dx}, \quad 2\eta = \frac{d\Omega}{dy}, \quad 2\zeta = \frac{d\Omega}{dz}, \text{ etc.}$$

of the function $\Omega$. Hence the values of the unknowns produced by the best combinations of the observations — which we may rightly call the *most reliable values*[9] — are the same values that minimize $\Omega$. Now $V - L$ represents the general difference between the computed and the observed values. Thus the most reliable values of the unknowns are those that minimize the sum of the products of the weights of the observations with the squares of the differences between the observed and the computed values of the quantities $V$, $V'$, $V''$, etc. I established this principle by quite different means some time ago in the *Theoria Motus Corporum Coelestium*.

If, in addition, we must assign relative precisions to the individual estimates, we may use general elimination in (III) to express $x$, $y$, $z$, etc. in the form

$$\left.\begin{array}{l} x = A + [\alpha\alpha]\xi + [\alpha\beta]\eta + [\alpha\gamma]\zeta + \text{etc.} \\ y = B + [\beta\alpha]\xi + [\beta\beta]\eta + [\beta\gamma]\zeta + \text{etc.} \\ z = C + [\gamma\alpha]\xi + [\gamma\beta]\eta + [\gamma\gamma]\zeta + \text{etc.} \\ \text{etc.} \end{array}\right\} \text{(VII)}$$

and the most reliable values for the unknowns $x$, $y$, $z$, etc. will be $A$, $B$, $C$, etc. The weights of these estimates are $\frac{1}{[\alpha\alpha]}$, $\frac{1}{[\beta\beta]}$, $\frac{1}{[\gamma\gamma]}$, etc., and the mean errors to be feared in them are

$$\begin{array}{l} \text{for } x \ldots\ldots\ldots = m\sqrt{p[\alpha\alpha]} = m'\sqrt{p'[\alpha\alpha]} = m''\sqrt{p''[\alpha\alpha]} \text{ etc.,} \\ \text{for } y \ldots\ldots\ldots = m\sqrt{p[\beta\beta]} = m'\sqrt{p'[\beta\beta]} = m''\sqrt{p''[\beta\beta]} \text{ etc.,} \\ \text{for } z \ldots\ldots\ldots = m\sqrt{p[\gamma\gamma]} = m'\sqrt{p'[\gamma\gamma]} = m''\sqrt{p''[\gamma\gamma]} \text{ etc.,} \\ \text{etc.} \end{array}$$

All this agrees with what I showed in *Theoria Motus Corporum Coelestium*.

## 22.

It is appropriate to give a brief, separate treatment of the simplest and most frequent case of all. This occurs when there is a single unknown and $V = x$,

---

[9]The English word "plausible" has negative connotations which make it unsuitable to translate Gauss's *valores maxime plausibiles*. The phrase "most reliable values," which we use here, is a translation of *sicherste werte*, which Gauss uses in his *Anzeigen*.

etc., et proin

$$\xi = (p + p' + p'' + \text{etc.})x - (pL + p'L' + p''L'' + \text{etc.})$$

Hinc porro

$$[\alpha\alpha] = \frac{1}{p+p'+p''+\text{etc.}}$$
$$A \quad = \frac{pL+p'L'+p''L''+\text{etc.}}{p+p'+p''+\text{etc.}}$$

Si itaque e pluribus observationibus inaequali praecisione gaudentibus, et quarum ponder resp. sunt $p$, $p'$, $p''$, etc., valor eiusdem quantitatis inventus est e prima $= L$, e secunda $= L'$, e tertia $= L''$, etc., huius valor maxime plausibilis erit

$$\frac{pL+p'L'+p''L''+\text{etc.}}{p+p'+p''+\text{etc.}}$$

pondusque huius determinationis $= p+p'+p''+$etc. Si omnes observationes aequali praecisione gaudent, valor maxime plausibilis erit

$$\frac{L+L'+L''+\text{etc.}}{\pi}$$

i.e. aequalis medio arithmetico valorum observatorum, huiusque determinationis pondus $= \pi$, accepto pondere observationum pro unitate.

$V' = x$, $V'' = x$, etc. Here we have $a = \sqrt{p}$, $a' = \sqrt{p'}$, $a'' = \sqrt{p''}$, etc., $l = -L\sqrt{p}$, $l' = -L'\sqrt{p'}$, $l'' = -L''\sqrt{p''}$, etc., and hence

$$\xi = (p + p' + p'' + \text{etc.})x - (pL + p'L' + p''L'' + \text{etc.}).$$

Further

$$[\alpha\alpha] = \frac{1}{p+p'+p''+\text{etc.}},$$

$$A \quad = \frac{pL+p'L'+p''L''+\text{etc.}}{p+p'+p''+\text{etc.}}.$$

Thus if values of the same quantity have been found from observations $L$, $L'$, $L''$, etc. of unequal precision, whose weights are $p$, $p'$, $p''$, etc., then its most reliable value will be

$$\frac{pL+p'L'+p''L''+\text{etc.}}{p+p'+p''+\text{etc.}},$$

and the weight of this estimate will be $p + p' + p'' + \text{etc.}$ If all the observations have equal precision, the most reliable value will be

$$\frac{L+L'+L''+\text{etc.}}{\pi};$$

i.e., it will be equal to the arithmetic mean. If we take the weight of the observations to be one, the weight of this estimate will be $\pi$.

*This page intentionally left blank*

# Theoria
# Combinationis Observationum
# Erroribus Minimis Obnoxiae

*Pars Posterior*

# Theory of the
# Combination of Observations
# Least Subject to Errors

*Part Two*

# Theoria Combinationis Observationum Erroribus Minimis Obnoxiae

## Pars Posterior

### 23.

Plures adhunc supersunt disquisitiones, per quas theoria praecedens tum illustrabitur tum ampliabitur.

Ante omnia investigare oportet, num negotium eliminationis, cuius adiumento indeterminatae $x$, $y$, $z$, etc. per $\xi$, $\eta$, $\zeta$, etc. exprimendae sunt, semper sit possible. Quum multitudo illarum multitudini harum aequalis sit, e theoria eliminationis in aequationibus linearibus constat, illam eliminationem, si $\xi$, $\eta$, $\zeta$, etc. ab invicem independentes sint, certo possibilem fore; sin minus, impossibilem. Supponamus aliquantisper, $\xi$, $\eta$, $\zeta$, etc. non esse ab invicem independentes, sed exstare inter ipsas aequationem identicam

$$0 = F\xi + G\eta + H\zeta + \text{etc.} + K$$

Habebimus itaque

$$\begin{gathered} F\Sigma aa + G\Sigma ab + H\Sigma ac + \text{etc.} = 0 \\ F\Sigma ab + G\Sigma bb + H\Sigma bc + \text{etc.} = 0 \\ F\Sigma ac + G\Sigma bb + H\Sigma cc + \text{etc.} = 0 \end{gathered}$$

etc., nec non

$$F\Sigma al + G\Sigma bl + H\Sigma cl + \text{etc.} = -K$$

Statuendo porro

$$\left.\begin{aligned} aF + bG + cH + \text{etc.} &= \theta \\ a'F + b'G + c'H + \text{etc.} &= \theta' \\ a''F + b''G + c''H + \text{etc.} &= \theta'' \end{aligned}\right\} \text{(I)}$$

etc., eruitur

$$\begin{gathered} a\theta + a'\theta' + a''\theta'' + \text{etc.} = 0 \\ b\theta + b'\theta' + b''\theta'' + \text{etc.} = 0 \\ c\theta + c'\theta' + c''\theta'' + \text{etc.} = 0 \end{gathered}$$

etc., nec non

$$l\theta + l'\theta' + l''\theta'' + \text{etc.} = -K$$

# Theory of the Combination of Observations Least Subject to Errors

## Part Two

### 23.

A number of topics remain to be treated, topics that will illustrate and amplify the above theory.

First of all we must determine if it is always possible to carry out the elimination by means of which the unknowns $x$, $y$, $z$, etc. were represented in terms of $\xi$, $\eta$, $\zeta$, etc. Since the number of variables in these two groups is the same, the theory of elimination for linear equations tells us that the elimination will be possible if and only if $\xi$, $\eta$, $\zeta$, etc. are mutually independent.

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

$$0 = F\xi + G\eta + H\zeta + \text{etc.} + K.$$

Then we have

$$\begin{aligned}
F\Sigma aa + G\Sigma ab + H\Sigma ac + \text{etc.} &= 0\\
F\Sigma ab + G\Sigma bb + H\Sigma bc + \text{etc.} &= 0\\
F\Sigma ac + G\Sigma bb + H\Sigma cc + \text{etc.} &= 0
\end{aligned}$$

etc. and

$$F\Sigma al + G\Sigma bl + H\Sigma cl + \text{etc.} = -K.$$

If we set

$$\left.\begin{aligned}
aF + bG + cH + \text{etc.} &= \theta\\
a'F + b'G + c'H + \text{etc.} &= \theta'\\
a''F + b''G + c''H + \text{etc.} &= \theta''
\end{aligned}\right\} \text{(I)}$$

etc., we find that

$$\begin{aligned}
a\theta + a'\theta' + a''\theta'' + \text{etc.} &= 0\\
b\theta + b'\theta' + b''\theta'' + \text{etc.} &= 0\\
c\theta + c'\theta' + c''\theta'' + \text{etc.} &= 0
\end{aligned}$$

etc. and

$$l\theta + l'\theta' + l''\theta'' + \text{etc.} = -K.$$

Multiplicando itaque aequationes (I) resp. per $\theta$, $\theta'$, $\theta''$, etc., et addendo, obtinemus:

$$0 = \theta\theta + \theta'\theta' + \theta''\theta'' + \text{etc.}$$

quae aequatio manifesto consistere nequit, nisi simul fuerit $\theta = 0$, $\theta' = 0$, $\theta'' = 0$, etc. Hinc primo colligimus, necessario esse debere $K = 0$. Dein aequationes (I) docent, functiones $v$, $v'$, $v''$, etc. ita comparatas esse, ut ipsarum valores non mutentur, si valores quantitatum $x$, $y$, $z$, etc. capiant incrementa vel decrementa ipsis $F$, $G$, $H$, etc. resp. proportionalia, idemque manifesto de functionibus $V$, $V'$, $V''$, etc. valebit. Suppositio itaque consistere nequit, nisi in casu tali, ubi vel e valoribus exactis quantitatum $V$, $V'$, $V''$, etc. valores incognitarum $x$, $y$, $z$, etc. determinare impossible fuisset, i.e. ubi problema natura sua fuisset indeterminatum, quem casum a disquisitione nostra exclusimus.

24.

Denotemus per $\beta$, $\beta'$, $\beta''$, etc. multiplicatores, qui eandem relationem habent ad indeterminatam $y$, quam habent $\alpha$, $\alpha'$, $\alpha''$, etc. ad $x$, puta sit

$$\begin{aligned} a[\beta\alpha] + b[\beta\beta] + c[\beta\gamma] + \text{etc.} &= \beta \\ a'[\beta\alpha] + b'[\beta\beta] + c'[\beta\gamma] + \text{etc.} &= \beta' \\ a''[\beta\alpha] + b''[\beta\beta] + c''[\beta\gamma] + \text{etc.} &= \beta'' \end{aligned}$$

etc., ita ut fiat indefinite

$$\beta v + \beta' v' + \beta'' v'' + \text{etc.} = y - B$$

Perinde sint $\gamma$, $\gamma'$, $\gamma''$, etc. multiplicatores similes respectu indeterminatae $z$, puta

$$\begin{aligned} a[\gamma\alpha] + b[\gamma\beta] + c[\gamma\gamma] + \text{etc.} &= \gamma \\ a'[\gamma\alpha] + b'[\gamma\beta] + c'[\gamma\gamma] + \text{etc.} &= \gamma' \\ a''[\gamma\alpha] + b''[\gamma\beta] + c''[\gamma\gamma] + \text{etc.} &= \gamma'' \end{aligned}$$

etc., ita ut fiat indefinite

$$\gamma v + \gamma' v' + \gamma'' v'' + \text{etc.} = z - C$$

If we multiply the equations (I) by $\theta$, $\theta'$, $\theta''$, etc. and add, we get

$$0 = \theta\theta + \theta'\theta' + \theta''\theta'' + \text{etc.},$$

which can only hold if $\theta = 0$, $\theta' = 0$, $\theta'' = 0$, etc. From this we first conclude that $K = 0$. It then follows from (I) that the functions $v$, $v'$, $v''$, etc. do not change when $x$, $y$, $z$, etc. are increased or decreased by quantities proportional to $F$, $G$, $H$, etc. The same is obviously true of the functions $V$, $V'$, $V''$, etc. Thus our supposition can hold only if it is impossible to determine $x$, $y$, $z$, etc. from the exact values of $V$, $V'$, $V''$, etc., that is, when the problem is indeterminate, a case which we have excluded from our inquiry.

24.

Let $\beta$, $\beta'$, $\beta''$, etc. be the multipliers for $y$ corresponding to the multipliers $\alpha$, $\alpha'$, $\alpha''$, etc. for $x$; i.e.,

$$\begin{aligned}
a[\beta\alpha] + b[\beta\beta] + c[\beta\gamma] + \text{etc.} &= \beta \\
a'[\beta\alpha] + b'[\beta\beta] + c'[\beta\gamma] + \text{etc.} &= \beta' \\
a''[\beta\alpha] + b''[\beta\beta] + c''[\beta\gamma] + \text{etc.} &= \beta''
\end{aligned}$$

etc. Then we have the identity

$$\beta v + \beta' v' + \beta'' v'' + \text{etc.} = y - B.$$

Next, let $\gamma$, $\gamma'$, $\gamma''$, etc. be the multipliers for $z$; i.e.,

$$\begin{aligned}
a[\gamma\alpha] + b[\gamma\beta] + c[\gamma\gamma] + \text{etc.} &= \gamma \\
a'[\gamma\alpha] + b'[\gamma\beta] + c'[\gamma\gamma] + \text{etc.} &= \gamma' \\
a''[\gamma\alpha] + b''[\gamma\beta] + c''[\gamma\gamma] + \text{etc.} &= \gamma''
\end{aligned}$$

etc., so that we have the identity

$$\gamma v + \gamma' v' + \gamma'' v'' + \text{etc.} = z - C.$$

et sic porro. Hoc pacto, perinde ut iam in art. 20 inveneramus

$$\Sigma\alpha a = 1,\quad \Sigma\alpha b = 0,\quad \Sigma\alpha c = 0, \text{ etc.,}\qquad \text{nec non}\quad \Sigma\alpha l = -A$$

etiam habebimus

$$\Sigma\beta a = 0,\quad \Sigma\beta b = 1,\quad \Sigma\beta c = 0, \text{ etc.,}\quad \text{atque}\quad \Sigma\beta l = -B$$
$$\Sigma\gamma a = 0,\quad \Sigma\gamma b = 0,\quad \Sigma\gamma c = 1, \text{ etc.,}\quad \text{atque}\quad \Sigma\gamma l = -C$$

et sic porro. Nec minus, quemadmodum in art. 20 prodiit $\Sigma\alpha\alpha = [\alpha\alpha]$, etiam erit

$$\Sigma\beta\beta = [\beta\beta],\quad \Sigma\gamma\gamma = [\gamma\gamma], \text{etc.}$$

Multiplicando porro valores ipsorum $\alpha$, $\alpha'$, $\alpha''$, etc. (art. 20 IV) resp. per $\beta$, $\beta'$, $\beta''$, etc., et addendo, obtinemus

$$\alpha\beta + \alpha'\beta' + \alpha''\beta'' + \text{etc.} = [\alpha\beta],\ \text{ sive }\ \Sigma\alpha\beta = [\alpha\beta]$$

Multiplicando autem valores ipsorum $\beta$, $\beta'$, $\beta''$, etc. resp. per $\alpha$, $\alpha'$, $\alpha''$, etc., et addendo, perinde prodit

$$\alpha\beta + \alpha'\beta' + \alpha''\beta'' + \text{etc.} = [\beta\alpha],\ \text{ adeoque }\ [\alpha\beta] = [\beta\alpha]$$

Prorsus simili modo eruitur

$$[\alpha\gamma] = [\gamma\alpha] = \Sigma\alpha\gamma,\quad [\beta\gamma] = [\gamma\beta] = \Sigma\beta\gamma,\quad \text{etc.}$$

## 25.

Denotemus porro per $\lambda$, $\lambda'$, $\lambda''$, etc. valores functionum $v$, $v'$, $v''$, etc., qui prodeunt, dum pro $x$, $y$, $z$, etc. ipsarum valores maxime plausibile $A$, $B$, $C$, etc. substituuntur, puta

$$\begin{aligned} aA + bB + cC + \text{etc.} + l &= \lambda \\ a'A + b'B + c'C + \text{etc.} + l' &= \lambda' \\ a''A + b''B + c''C + \text{etc.} + l'' &= \lambda'' \end{aligned}$$

etc.; statuamus praeterae

$$\lambda\lambda + \lambda'\lambda' + \lambda''\lambda'' + \text{etc.} = M$$

Continue in this manner for the other unknowns. Then just as we found in Art. 20 that

$$\Sigma\alpha a = 1, \quad \Sigma\alpha b = 0, \quad \Sigma\alpha c = 0, \text{ etc.,} \quad \text{and} \quad \Sigma\alpha l = -A,$$

we now have

$$\begin{aligned} &\Sigma\beta a = 0, \quad \Sigma\beta b = 1, \quad \Sigma\beta c = 0, \text{ etc.,} \quad \text{and} \quad \Sigma\beta l = -B; \\ &\Sigma\gamma a = 0, \quad \Sigma\gamma b = 0, \quad \Sigma\gamma c = 1, \text{ etc.,} \quad \text{and} \quad \Sigma\gamma l = -C; \end{aligned}$$

and so on. Moreover, just as we had $\Sigma\alpha\alpha = [\alpha\alpha]$ in Art. 20, we now have

$$\Sigma\beta\beta = [\beta\beta], \quad \Sigma\gamma\gamma = [\gamma\gamma], \quad \text{etc.}$$

Multiplying $\alpha$, $\alpha'$, $\alpha''$, etc. in (IV) of Art. 20 by $\beta$, $\beta'$, $\beta''$, etc. and adding, we get

$$\alpha\beta + \alpha'\beta' + \alpha''\beta'' + \text{etc.} = [\alpha\beta] \ \text{ or } \ \Sigma\alpha\beta = [\alpha\beta].$$

Multiplying $\beta$, $\beta'$, $\beta''$, etc. by $\alpha$, $\alpha'$, $\alpha''$, etc. and adding, we get

$$\alpha\beta + \alpha'\beta' + \alpha''\beta'' + \text{etc.} = [\beta\alpha] \ \text{ whence } \ [\alpha\beta] = [\beta\alpha].$$

Similarly,

$$[\alpha\gamma] = [\gamma\alpha] = \Sigma\alpha\gamma, \quad [\beta\gamma] = [\gamma\beta] = \Sigma\beta\gamma, \quad \text{etc.}$$

## 25.

Let $\lambda$, $\lambda'$, $\lambda''$, etc. denote the values of the functions $v$, $v'$, $v''$, etc. at the most reliable values $A$, $B$, $C$, etc. of the unknowns $x$, $y$, $z$, etc.; that is,

$$\begin{aligned} aA + bB + cC + \text{etc.} + l &= \lambda \\ a'A + b'B + c'C + \text{etc.} + l' &= \lambda' \\ a''A + b''B + c''C + \text{etc.} + l'' &= \lambda'' \end{aligned}$$

etc. Let us also set

$$\lambda\lambda + \lambda'\lambda' + \lambda''\lambda'' + \text{etc.} = M.$$

ita ut sit $M$ valor functionis $\Omega$ valoribus maxime plausibilibus indeterminatarum respondens, adeoque per ea, quae in art. 20 demonstravimus, valor minimus huis functionis. Hinc erit $a\lambda + a'\lambda' + a''\lambda'' +$ etc. valor ipsius $\xi$, valoribus $x = A$, $y = B$, $z = C$, etc. respondens, adeoque $= 0$, i.e., habebimus

$$\Sigma a\lambda = 0$$

et perinde fiet

$$\Sigma b\lambda = 0, \quad \Sigma c\lambda = 0, \text{ etc.}; \quad \text{nec non} \quad \Sigma\alpha\lambda = 0, \quad \Sigma\beta\lambda = 0, \quad \Sigma\gamma\lambda = 0, \quad \text{etc.}$$

Denique multiplicando expressiones ipsarum $\lambda$, $\lambda'$, $\lambda''$, etc. per $\lambda$, $\lambda'$, $\lambda''$, etc. et addendo, obtinemum $l\lambda + l'\lambda' + l''\lambda'' +$ etc. $= \lambda\lambda + \lambda'\lambda' + \lambda''\lambda'' +$ etc., sive

$$\Sigma l\lambda = M$$

26.

Substituendo in aequatione $v = ax + by + cz +$ etc. $+ l$, pro $x$, $y$, $z$, etc. expressiones VII. art. 21, prodibit, adhibitis reductionibus ex praecedentibus obviis,

$$v = \alpha\xi + \beta\eta + \gamma\zeta + \text{etc.} + \lambda$$

et perinde erit indefinite

$$\begin{aligned} v' &= \alpha'\xi + \beta'\eta + \gamma'\zeta + \text{etc.} + \lambda' \\ v'' &= \alpha''\xi + \beta''\eta + \gamma''\zeta + \text{etc.} + \lambda'' \end{aligned}$$

etc. Multiplicando vel has aequationes, vel aequationes I art. 20 resp. per $\lambda$, $\lambda'$, $\lambda''$, etc., et addendo, discimus esse indefinite

$$\lambda v + \lambda' v' + \lambda'' v'' + \text{etc.} = M$$

27.

Function $\Omega$ indefinite in pluribus formis exhiberi potest, quas evolvere operae pretium erit. Ac primo quidem quadrando aequationes I art. 20 et addendo, statim fit

$$\begin{aligned} \Omega = {} & xx\Sigma aa + yy\Sigma bb + zz\Sigma cc + \text{etc.} \\ & + 2xy\Sigma ab + 2xz\Sigma ac + 2yz\Sigma bc + \text{etc.} \\ & + 2x\Sigma al + 2y\Sigma bl + 2z\Sigma cl + \text{etc.} \\ & + \Sigma ll \end{aligned}$$

Then $M$ is the value of the function $\Omega$ corresponding to the most reliable values of the unknowns, and as we showed in Art. 20, $M$ is the minimum value of $\Omega$. Thus $a\lambda + a'\lambda' + a''\lambda'' +$ etc. is the value of $\xi$ corresponding to $x = A$, $y = B$, $z = C$, etc., which is zero. Hence,

$$\Sigma a\lambda = 0.$$

Similarly,

$$\Sigma b\lambda = 0, \quad \Sigma c\lambda = 0, \text{ etc.}; \quad \text{and} \quad \Sigma \alpha\lambda = 0, \quad \Sigma\beta\lambda = 0, \quad \Sigma\gamma\lambda = 0, \text{ etc.}$$

Finally, if we multiply the above expressions for $\lambda$, $\lambda'$, $\lambda''$, etc. by $\lambda$, $\lambda'$, $\lambda''$, etc. and add, we get $l\lambda + l'\lambda' + l''\lambda'' + \text{etc.} = \lambda\lambda + \lambda'\lambda' + \lambda''\lambda'' + \text{etc.}$, or

$$\Sigma l\lambda = M.$$

## 26.

If we substitute the expressions for $x$, $y$, $z$, etc. from VII of Art. 21 into the equation $v = ax + by + cz + \text{etc.} + l$, then we obtain from the preceding reductions

$$v = \alpha\xi + \beta\eta + \gamma\zeta + \text{etc.} + \lambda.$$

Likewise, we have generally

$$\begin{aligned} v' &= \alpha'\xi + \beta'\eta + \gamma'\zeta + \text{etc.} + \lambda' \\ v'' &= \alpha''\xi + \beta''\eta + \gamma''\zeta + \text{etc.} + \lambda'' \end{aligned}$$

etc. Multiplying these equations, or the equations I in Art. 20, by $\lambda$, $\lambda'$, $\lambda''$, etc., we find that the relation

$$\lambda v + \lambda' v' + \lambda'' v'' + \text{etc.} = M$$

holds generally.

## 27.

The function $\Omega$ can be written in several forms, and it will be worthwhile to derive them. By squaring the equations I in Art. 20 and adding we have

$$\begin{aligned} \Omega = {} & xx\Sigma aa + yy\Sigma bb + zz\Sigma cc + \text{etc.} \\ & + 2xy\Sigma ab + 2xz\Sigma ac + 2yz\Sigma bc + \text{etc.} \\ & + 2x\Sigma al + 2y\Sigma bl + 2z\Sigma cl + \text{etc.} \\ & + \Sigma ll. \end{aligned}$$

quae est forma *prima.*

Multiplicando easdem aequationes resp. per $v$, $v'$, $v''$, etc., et addendo, obtinemus:

$$\Omega = \xi x + \eta y + \zeta z + \text{etc.} + lv + l'v' + l''v'' + \text{etc.}$$

atque hinc, substituendo pro $v$, $v'$, $v''$, etc. expressiones in art. praec. traditas,

$$\Omega = \xi x + \eta y + \zeta z + \text{etc.} - A\xi - B\eta - C\zeta - \text{etc.} + M$$

sive

$$\Omega = \xi(x - A) + \eta(y - B) + \zeta(z - C) + \text{etc.} + M$$

quae est forma *secunda.*

Substituendo in forma secunda pro $x - A$, $y - B$, $z - C$, etc. expressiones VII. art. 21, obtinemus formam *tertiam*:

$$\begin{aligned}\Omega = {} & [\alpha\alpha]\xi\xi + [\beta\beta]\eta\eta + [\gamma\gamma]\zeta\zeta + \text{etc.}\\ & + 2[\alpha\beta]\xi\eta + 2[\alpha\gamma]\xi\zeta + 2[\beta\gamma]\eta\zeta + \text{etc.} + M\end{aligned}$$

His adiungi potest forma *quarta,* ex forma tertia atque formulis art. praec. sponte demanans:

$$\begin{aligned}&\Omega = (v - \lambda)^2 + (v' - \lambda')^2 + (v'' - \lambda'')^2 + \text{etc.} + M, \quad \text{sive}\\ &\Omega = M + \Sigma(v - \lambda)^2\end{aligned}$$

quae forma conditionem minimi directe ob oculos sistit.

28.

Sint $e$, $e'$, $e''$, etc. errores in observationibus, quae dederunt $V = L$, $V' = L'$, $V'' = L''$, etc., commissi, i.e. sint valores veri functionem $V$, $V'$, $V''$, etc. resp. $L - e$, $L' - e'$, $L'' - e''$, etc. adeoque valores veri ipsarum $v$, $v'$, $v''$, etc. resp. $-e\sqrt{p}$, $-e'\sqrt{p'}$, $-e''\sqrt{p''}$, etc. Hinc valor verus ipsius $x$ erit

$$= A - \alpha e\sqrt{p} - \alpha' e'\sqrt{p'} - \alpha'' e''\sqrt{p''} - \text{etc.}$$

sive error valoris ipsius $x$, in determinatione maxime idonea commissus, quem per $Ex$ denotare convenit,

$$= \alpha e\sqrt{p} + \alpha' e'\sqrt{p'} + \alpha'' e''\sqrt{p''} + \text{etc.}$$

This is the *first* form.

If we multiply the same equations by $v$, $v'$, $v''$, etc. and add, we get

$$\Omega = \xi x + \eta y + \zeta z + \text{etc.} + lv + l'v' + l''v'' + \text{etc.}$$

Hence if we replace $v$, $v'$, $v''$, etc. by the expressions from the preceding article, we have

$$\Omega = \xi x + \eta y + \zeta z + \text{etc.} - A\xi - B\eta - C\zeta - \text{etc.} + M$$

or

$$\Omega = \xi(x - A) + \eta(y - B) + \zeta(z - C) + \text{etc.} + M.$$

This is the *second* form.

By substituting the expressions VII in Art. 21 for $x - A$, $y - B$, $z - C$, etc. in the second form, we get the *third* form:

$$\begin{aligned}\Omega = [\alpha\alpha]\xi\xi &+ [\beta\beta]\eta\eta + [\gamma\gamma]\zeta\zeta + \text{etc.}\\ &+ 2[\alpha\beta]\xi\eta + 2[\alpha\gamma]\xi\zeta + 2[\beta\gamma]\eta\zeta + \text{etc.} + M.\end{aligned}$$

In addition there is a *fourth* form, which follows directly from the third form and the formulas of the preceding article:

$$\begin{aligned}&\Omega = (v - \lambda)^2 + (v' - \lambda')^2 + (v'' - \lambda'')^2 + \text{etc.} + M, \quad \text{or}\\ &\Omega = M + \Sigma(v - \lambda)^2.\end{aligned}$$

This form makes the condition for a minimum obvious.

28.

Let $e$, $e'$, $e''$, etc. be the errors in the observations which gave the equations $V = L$, $V' = L'$, $V'' = L''$, etc.; i.e., let the true values of the functions $V$, $V'$, $V''$, etc. be $L - e$, $L' - e'$, $L'' - e''$, etc., so that the true values of $v$, $v'$, $v''$, etc. are $-e\sqrt{p}$, $-e'\sqrt{p'}$, $-e''\sqrt{p''}$, etc. Then the true value of $x$ is

$$A - \alpha e\sqrt{p} - \alpha' e'\sqrt{p'} - \alpha'' e''\sqrt{p''} - \text{etc.},$$

and the error $Ex$ in the most reliable estimate of $x$ is

$$\alpha e\sqrt{p} + \alpha' e'\sqrt{p'} + \alpha'' e''\sqrt{p''} + \text{etc.}$$

Perinde error valoris ipsius $y$ in determinatione maxime idonea commissus, quem per $Ey$ denotabimus, erit

$$= \beta e\sqrt{p} + \beta' e'\sqrt{p'} + \beta'' e''\sqrt{p''} + \text{etc.}$$

Valor medius quadrati $(Ex)^2$ invenitur

$$= mmp(\alpha\alpha + \alpha'\alpha' + \alpha''\alpha'' + \text{etc.}) = mmp[\alpha\alpha]$$

valor medius quadrati $(Ey)^2$ perinde $= mmp[\beta\beta]$ etc., ut iam supra docuimus. Iam vero etiam valorem medium *producti* $Ex.Ey$ assignare licet, quippe qui invenitur

$$mmp(\alpha\beta + \alpha'\beta' + \alpha'\beta' + \text{etc.}) = mmp[\alpha\beta]$$

Concinne haec ita quoque exprimi possunt. Valores medii quadratorum$(Ex)^2$, $(Ey)^2$, etc. resp. aequales sunt productis ex $\frac{1}{2}mmp$ in quotientes differentialium partialium secundi ordinis

$$\frac{\mathrm{dd}\Omega}{\mathrm{d}\xi^2}, \quad \frac{\mathrm{dd}\Omega}{\mathrm{d}\eta^2}, \quad \text{etc.}$$

valorque medius producti talis, ut $Ex.Ey$, aequalis est producto ex $\frac{1}{2}mmp$ in quotientem differentialem $\frac{\mathrm{dd}\Omega}{\mathrm{d}\xi.\mathrm{d}\eta}$, quatenus quidem $\Omega$ tamquam functio indeterminatarum $\xi$, $\eta$, $\zeta$, etc. consideratur.

29.

Designet $t$ functionem datam linearem quantitatum $x$, $y$, $z$, etc. puta sit

$$t = fx + gy + hz + \text{etc.} + k$$

Valor ipsius $t$, e valoribus maxime plausibilibus ipsarum $x$, $y$, $z$, etc. prodiens hinc erit $= fA + gB + hC + \text{etc.} + k$, quem per $K$ denotabimus. Qui si tamquam valor verus ipsius $t$ adoptatur, error committitur, qui erit

$$= fEx + gEy + hEz + \text{etc.}$$

atque per $Et$ denotabitur. Manifest valor medius huius erroris fit $= 0$, sive error a parte constante liber erit. At valor medius quadrati $(Et)^2$, sive valor medius aggretati

$$\begin{aligned} ff(Ex)^2 &+ 2fgEx.Ey + 2fhEx.Ez + \text{etc.} \\ &+ gg(Ey)^2 \quad + 2ghEy.Ez + \text{etc.} \\ &\qquad\qquad\quad + hh(Ez)^2 \quad + \text{etc. etc.} \end{aligned}$$

Similarly, the error $Ey$ in the most reliable estimate of $y$ is

$$\beta e\sqrt{p} + \beta' e'\sqrt{p'} + \beta'' e''\sqrt{p''} + \text{etc.}$$

As we showed above, the mean value of $(Ex)^2$ is

$$mmp(\alpha\alpha + \alpha'\alpha' + \alpha''\alpha'' + \text{etc.}) = mmp[\alpha\alpha],$$

and similarly the mean value of $(Ey)^2$ is $mmp[\beta\beta]$. But now we can also determine the mean value of the *product* $Ex.Ey$, which is obviously

$$mmp(\alpha\beta + \alpha'\beta' + \alpha'\beta' + \text{etc.}) = mmp[\alpha\beta].$$

These mean values may also be elegantly represented as follows. When $\Omega$ is regarded as a function of the unknowns $\xi$, $\eta$, $\zeta$, etc., the mean values of the squares $(Ex)^2$, $(Ey)^2$, etc. are equal to the product of $\frac{1}{2}mmp$ with the second-order partial differential quotients

$$\frac{\mathrm{dd}\Omega}{\mathrm{d}\xi^2}, \quad \frac{\mathrm{dd}\Omega}{\mathrm{d}\eta^2}, \quad \text{etc.},$$

and the mean value of a product such as $Ex.Ey$, is equal to the product of $\frac{1}{2}mmp$ with the differential quotient $\frac{\mathrm{dd}\Omega}{\mathrm{d}\xi.\mathrm{d}\eta}$.

29.

Let $t$ be a linear function of $x$, $y$, $z$, etc., say

$$t = fx + gy + hz + \text{etc.} + k.$$

The value $K$ of $t$ at the most reliable values of $x$, $y$, $z$, etc. is $fA+gB+hC+\text{etc.}+k$. If we accept $K$ in place of the true value of $t$, we will make an error $Et$ whose value is

$$fEx + gEy + hEz + \text{etc.}$$

Clearly the mean value of this error is zero; i.e., the error has no constant part. The mean value of the square $(Et)^2$ is the mean value of the sum

$$\begin{aligned} ff(Ex)^2 + 2fgEx.Ey &+ 2fhEx.Ez + \text{etc.} \\ + gg(Ey)^2 \quad &+ 2ghEy.Ez + \text{etc.} \\ &+ hh(Ez)^2 \quad + \text{etc. etc.,} \end{aligned}$$

per ea, quae in art. praec. exposuimus, aequalis fit producto ex $mmp$ in aggregatum

$$\begin{aligned} ff[\alpha\alpha] + 2fg[\alpha\beta] &+ 2fh[\alpha\gamma] + \text{etc.} \\ + gg[\beta\beta] &+ 2gh[\beta\gamma] + \text{etc.} \\ &+ hh[\gamma\gamma] \quad + \text{etc. etc.} \end{aligned}$$

sive product ex $mmp$ in valorem functionis $\Omega - M$, qui prodit per substitutiones

$$\xi = f, \quad \eta = g, \quad \zeta = h \text{ etc.}$$

Denotando igitur hunc valorem determinatum functionis $\Omega - M$ per $\omega$, error medius metuendus, dum determinationi $t = K$ adhaeremus, erit $= m\sqrt{p\omega}$, sive pondus huius determinationis $= \frac{1}{\omega}$.

Quum indefinite habeatur

$$\Omega - M = (x - A)\xi + (y - B)\eta + (z - C)\zeta + \text{etc.}$$

patet, $\omega$ quoque aequalem esse valori determinato expressionis

$$(x - A)f + (y - B)g + (z - C)h + \text{etc.}$$

sive valori determinato ipsius $t - K$, qui prodit, si indeterminatis $x$, $y$, $z$, etc. trubuuntur valores ii, qui respondent valoribus ipsarum $\xi$, $\eta$, $\zeta$, etc. his $f$, $g$, $h$, etc.

Denique observamus, si $t$ indefinite in formam functionis ipsarum $\xi$, $\eta$, $\zeta$, etc. redigatur, ipsius partem constantem necessario fieri $= K$. Quodsi igiter indefinite fit

$$t = F\xi + G\eta + H\xi + \text{etc.} + K, \quad \text{erit} \quad \omega = fF + gG + hH + \text{etc.}$$

## 30.

Functio $\Omega$ valorem suum *absolute minimum* $M$, ut supra vidimus, nanciscitur, faciendo $x = A$, $y = B$, $z = C$, etc., sive $\xi = 0$, $\eta = 0$, $\zeta = 0$, etc. Si vero alicui illarum quantitatum valor *alius* iam tributus est, e.g. $x = A + \Delta$ variantibus reliquis $\Omega$ assequi potest valorem relative minimum, qui manifesto obtinetur adiumento aequationum

$$x = A + \Delta, \quad \tfrac{\mathrm{d}\Omega}{\mathrm{d}y} = 0, \quad \tfrac{\mathrm{d}\Omega}{\mathrm{d}z} = 0, \quad \text{etc.}$$

which, by the results of the preceding article, is equal to the product of $mmp$ with the sum

$$\begin{aligned} ff[\alpha\alpha] + 2fg[\alpha\beta] + 2fh[\alpha\gamma] &+ \text{etc.} \\ + gg[\beta\beta] \quad + 2gh[\beta\gamma] &+ \text{etc.} \\ + hh[\gamma\gamma] \quad &+ \text{etc. etc.} \end{aligned}$$

This is the same as the product of $mmp$ with the value $\omega$ of the function $\Omega - M$ at

$$\xi = f, \quad \eta = g, \quad \zeta = h, \quad \text{etc.}$$

Hence the mean error to be feared in the estimate $t = K$ is $m\sqrt{p\omega}$, and its weight is $\frac{1}{\omega}$.

Since the equation

$$\Omega - M = (x - A)\xi + (y - B)\eta + (z - C)\zeta + \text{etc.}$$

holds generally, $\omega$ is also equal to

$$(x - A)f + (y - B)g + (z - C)h + \text{etc.}$$

In other words, $\omega$ is equal to the value of $t - K$ which results from assigning $x$, $y$, $z$, etc. values corresponding to values $f$, $g$, $h$, etc. of $\xi$, $\eta$, $\zeta$, etc.

Finally, we observe that if $t$ is written as a function of $\xi$, $\eta$, $\zeta$, etc., then its constant term must be $K$. Therefore, if $t$ has the general form

$$t = F\xi + G\eta + H\xi + \text{etc.} + K, \quad \text{then} \quad \omega = fF + gG + hH + \text{etc.}$$

## 30.

We have seen that the function $\Omega$ attains its *absolute minimum* when $x = A$, $y = B$, $z = C$, etc. or when $\xi = 0$, $\eta = 0$, $\zeta = 0$, etc. But if a *different* value is assigned to any of these quantities, say $x = A + \Delta$, and the others are allowed to vary, then $\Omega$ assumes a relative minimum, which may be obtained from the equations

$$x = A + \Delta, \quad \frac{\mathrm{d}\Omega}{\mathrm{d}y} = 0, \quad \frac{\mathrm{d}\Omega}{\mathrm{d}z} = 0, \quad \text{etc.}$$

Fieri debet itaque $\eta = 0$, $\zeta = 0$, etc., adeoque, quoniam

$$x = A + [\alpha\alpha]\xi + [\alpha\beta]\eta + [\alpha\gamma]\zeta + \text{etc.}, \quad \xi = \tfrac{\Delta}{[\alpha\alpha]}$$

Simul habebitur

$$y = B + \tfrac{[\alpha\beta]\Delta}{[\alpha\alpha]}, \quad z = C + \tfrac{[\alpha\gamma]\Delta}{[\alpha\alpha]}, \quad \text{etc.}$$

Valor relative minimus ipsius $\Omega$ autem fit $= [\alpha\alpha]\xi\xi + M = M + \frac{\Delta\Delta}{[\alpha\alpha]}$. Vice versa hinc colligimus, si valor ipsius $\Omega$ limitem praescriptum $M + \mu\mu$ non superare debet, valorem ipsius $x$ necessario inter limites $A - \mu\sqrt{[\alpha\alpha]}$ et $A + \mu\sqrt{[\alpha\alpha]}$ contentum esse debere. Notari meretur, $\mu\sqrt{[\alpha\alpha]}$ aequalem fieri errori medio in valore maxime plausibili ipsius $x$ metuendo, si statuatur $\mu = m\sqrt{p}$, i.e. si $\mu$ aequalis sit errori medio observationum talium, quibus pondus $= 1$ tribuitur.

Generalius investigemus valorem minimum ipsius $\Omega$, qui pro valore dato ipsius $t$ locum habere potest, denotante $t$ ut in art. praec. functionem linearem $t = fx + gy + hz + \text{etc.} + k$, et cuius valor maxime plausibilis $= K$: valor praescriptus ipsius $t$ denotetur per $K + \kappa$. E theoria maximorum et minimorum constat, problematis solutionem petendam esse ex aequationibus

$$\begin{aligned} \tfrac{d\Omega}{dx} &= \theta\tfrac{dt}{dx} \\ \tfrac{d\Omega}{dy} &= \theta\tfrac{dt}{dy} \\ \tfrac{d\Omega}{dz} &= \theta\tfrac{dt}{dz} \text{ etc.,} \end{aligned}$$

sive $\xi = \theta f$, $\eta = \theta g$, $\zeta = \theta h$, etc., designante $\theta$ multiplicatorem adhuc indeterminatum. Quare si, ut in art. praec., statuimus, esse *indefinite*

$$t = F\xi + G\eta + H\zeta + \text{etc.} + K,$$

habebimus

$$\begin{aligned} K + \kappa &= \theta(fF + gG + hH + \text{etc.}) + K, \quad \text{sivi} \\ \theta &= \tfrac{\kappa}{\omega} \end{aligned}$$

accipiendo $\omega$ in eadem significatione ut in art. praec. Et quum $\Omega - M$, indefinite, sit functio homogenea secundi ordinis indeterminatarum $\xi$, $\eta$, $\zeta$, etc., sponte patet, eius valorem pro $\xi = \theta f$, $\eta = \theta g$, $\zeta = \theta h$, etc. fieri $= \theta\theta\omega$, et proin valorem minimum, quem $\Omega$ pro $t = K - \kappa$ obtinere potest, fieri $= M + \theta\theta\omega = M + \frac{\kappa\kappa}{\omega}$. Vice versa, si $\Omega$ debet valorem aliquem praescriptum $M + \mu\mu$ non superare, valor ipsius $t$ necessario inter limites $K - \mu\sqrt{\omega}$ et $K + \mu\sqrt{\omega}$ contentus esse debet, ubi $\mu\sqrt{\omega}$ aequalis fit errori medio in determinatione maxime plausibili ipsius $t$ metuendo, si pro $\mu$ accipitur error medius observationum, quibus pondus $= 1$ tribuitur.

Thus $\eta = 0$, $\zeta = 0$, etc., and since

$$x = A + [\alpha\alpha]\xi + [\alpha\beta]\eta + [\alpha\gamma]\zeta + \text{etc.}, \quad \xi = \tfrac{\Delta}{[\alpha\alpha]}.$$

Similarly,

$$y = B + \tfrac{[\alpha\beta]\Delta}{[\alpha\alpha]}, \quad z = C + \tfrac{[\alpha\gamma]\Delta}{[\alpha\alpha]}, \quad \text{etc.}$$

The value of $\Omega$ at its relative minimum is $[\alpha\alpha]\xi\xi + M = M + \frac{\Delta\Delta}{[\alpha\alpha]}$. Conversely, if the value of $\Omega$ is not to exceed a prescribed limit $M + \mu\mu$, then $x$ must lie between $A - \mu\sqrt{[\alpha\alpha]}$ and $A + \mu\sqrt{[\alpha\alpha]}$. It is worth noting that $\mu\sqrt{[\alpha\alpha]}$ is the mean error to be feared in the most reliable value of $x$ provided $\mu = m\sqrt{p}$; i.e., provided $\mu$ is equal to the mean error of the observations whose weights are one.

More generally, let us determine the minimum value $\Omega$ attains when the linear function $t = fx + gy + hz + \text{etc.} + k$ has a prescribed value. As in the preceding article, let $K$ be the most reliable value of $t$, and let the prescribed value of $t$ be $K + \kappa$. By the theory of maxima and minima, the solution of our problem may be found from the equations

$$\begin{aligned} \tfrac{d\Omega}{dx} &= \theta\tfrac{dt}{dx} \\ \tfrac{d\Omega}{dy} &= \theta\tfrac{dt}{dy} \\ \tfrac{d\Omega}{dz} &= \theta\tfrac{dt}{dz} \quad \text{etc.,} \end{aligned}$$

where $\theta$ denotes a multiplier to be determined. Equivalently, $\xi = \theta f$, $\eta = \theta g$, $\zeta = \theta h$, etc. It follows that if as above we write $t$ in the *general* form

$$t = F\xi + G\eta + H\zeta + \text{etc.} + K,$$

then

$$\begin{aligned} K + \kappa &= \theta(fF + gG + hH + \text{etc.}) + K, \quad \text{or} \\ \theta &= \tfrac{\kappa}{\omega}, \end{aligned}$$

where $\omega$ is defined as in the preceding article. Since $\Omega - M$ is a homogeneous function of the second degree in $\xi$, $\eta$, $\zeta$, etc., its value for $\xi = \theta f$, $\eta = \theta g$, $\zeta = \theta h$, etc. is clearly $\theta\theta\omega$. Hence the minimum value $\Omega$ can attain when $t = K + \kappa$ is $M + \theta\theta\omega = M + \frac{\kappa\kappa}{\omega}$. Conversely, if $\Omega$ is not to exceed some prescribed value $M + \mu\mu$, then $t$ must lie between $K - \mu\sqrt{\omega}$ and $K + \mu\sqrt{\omega}$. Here $\mu\sqrt{\omega}$ is the mean error to be feared in the most reliable estimate of $t$, provided $\mu$ is the mean error in the observations whose weights are one.

## 31.

Quoties multitudo quantitatum $x$, $y$, $z$, etc. paullo maior est, determinatio numerica valorem $A$, $B$, $C$, etc. ex aequationibus $\xi = 0$, $\eta = 0$, $\zeta = 0$, etc. per eliminationem vulgarem satis molesta evadit. Propterea in Theoria Motus Corporum Coelestium art. 182 algorithmum peculiarem addigitavimus, atque in *Disquistione de elementis ellipticis Palladis* (*Comm. recent. Soc. Gotting. Vol.* I) copiose explicavimus, per quem labor ille ad tantam quantam quidem res fert simplicitatem evehitur. Reducenda scilicet est functio $\Omega$ sub formam talem:

$$\frac{u^0u^0}{\mathcal{A}^0} + \frac{u'u'}{\mathcal{B}'} + \frac{u''u''}{\mathcal{C}''} + \frac{u'''u'''}{\mathcal{D}'''} + \text{etc.} + M$$

ubi divisores $\mathcal{A}^0$, $\mathcal{B}'$, $\mathcal{C}''$, $\mathcal{C}'''$, etc. sunt quantitates determinatae; $u^0$, $u'$, $u''$, $u'''$, etc. autem functiones lineares ipsarum $x$, $y$, $z$, etc., quarum tamen secunda $u'$ libera est ab $x$, tertia $u''$ libera ab $x$ et $y$, quarta $u'''$ libera $x$, $y$, et $z$, et sic porro, ita ut ultima $u^{(\pi-1)}$ solam ultimam indeterminatarum $x$, $y$, $z$, etc. implicet, denique coëfficientes, per quos $x$, $y$, $z$, etc. resp. multiplicatae sunt in $u^0$, $u'$, $u''$, etc. resp. aequales sunt ipsis $\mathcal{A}^0$, $\mathcal{B}'$, $\mathcal{C}''$, etc. Quibus ita factis statuendum est $u^0 = 0$, $u' = 0$, $u'' = 0$, $u''' = 0$, etc., unde valores incognitarum $x$, $y$, $z$, etc. inverso ordine commodissime elicientur. Haud opus videtur, algorithmum ipsum, per quem haec transformatio functionis $\Omega$ absolvitur, hic denuo repetere.

Sed multo adhuc magis prolixum calculum requirit eliminatio indefinita, cuius adiumento illarum determinationum pondera invenire oportet. Pondus quidem determinationis incognitae ultimae (quae sola ultimam $u^{(\pi-1)}$) per ea, quae in Theoria Motus Corporum Coelestium demonstrata sunt, facile invenitur aequale termino ultimo in serie divisorum $\mathcal{A}^0$, $\mathcal{B}'$, $\mathcal{C}''$, etc.; quapropter plures calculatores, ut eliminationem illam molestam evitarent, deficientibus aliis subsidiis, ita sibi consuluerunt, ut algorithmum, de quo diximus pluries, mutato quantitatum $x$, $y$, $z$, etc., ordine, repeterent, singulis deinceps ultimum locum occupantibus. Gratum itaque geometris fore speramus, si modum novum pondera determinationum calculandi, e penitiori argumenti perscrutatione haustum, hic exponamus, qui nihil amplius desiderandum relinquere videtur.

31.

Whenever the number of the quantities $x$, $y$, $z$, etc. is somewhat too large, the calculation of $A$, $B$, $C$, etc. from the equations $\xi = 0$, $\eta = 0$, $\zeta = 0$, etc. by the usual method of elimination becomes quite difficult. For this reason, in Art. 182 of the *Theoria Motus Corporum Coelestium* I sketched a special algorithm that simplified the calculations as far as possible, and later in the *Disquisitione de elementis ellipticis Palladis* (*Comm. recent. Soc. Gotting. Vol.* I) I set it out it in detail. Specifically, the function $\Omega$ can be reduced to the form

$$\frac{u^0u^0}{\mathcal{A}^0} + \frac{u'u'}{\mathcal{B}'} + \frac{u''u''}{\mathcal{C}''} + \frac{u'''u'''}{\mathcal{D}'''} + \text{etc.} + M,$$

in which the divisors $\mathcal{A}^0$, $\mathcal{B}'$, $\mathcal{C}''$, $\mathcal{C}'''$, etc. are constants and $u^0$, $u'$, $u''$, $u'''$, etc. are linear functions of $x$, $y$, $z$, etc. However, the second function $u'$ is independent of $x$; the third $u''$ is independent of $x$ and $y$; the fourth $u'''$ is independent of $x$, $y$, and $z$, and so on. The last function $u^{(\pi-1)}$ depends only on the last of the unknowns $x$, $y$, $z$, etc. Moreover, the coefficients $\mathcal{A}^0$, $\mathcal{B}'$, $\mathcal{C}''$, etc. multiply $x$, $y$, $z$, etc. in $u^0$, $u'$, $u''$, etc. respectively. Given this reduction, we may easily find $x$, $y$, $z$, etc. in reverse order after setting $u^0 = 0$, $u' = 0$, $u'' = 0$, $u''' = 0$, etc. There is no need to repeat here the description of the algorithm by which the function $\Omega$ is transformed.

However, the general elimination by which the weights of the estimates are found requires a great deal more work. Now the weight of the estimate of the last unknown (the only unknown in $u^{(\pi-1)}$) is the last term in the series of divisors $\mathcal{A}^0$, $\mathcal{B}'$, $\mathcal{C}''$, etc.—a fact which follows easily from results established in *Theoria Motus Corporum Coelestium*. For this reason many calculators, seeing no other way of avoiding the trouble of a general elimination, have adopted the strategy of repeating my algorithm with the order of the quantities $x$, $y$, $z$, etc. changed so that each in turn occupies the last position. I therefore hope that mathematicians will be grateful if I set forth a new method for calculating the weights of the estimates, a method based on a deep examination of the subject, which seems to leave nothing more to be desired.

32.

Statuamus itaque esse (I)

$$\begin{aligned}
u^0 &= \mathcal{A}^0 x + \mathcal{B}^0 y + \mathcal{C}^0 z + \text{etc.} + \mathcal{L}^0\\
u' &= \qquad\;\; \mathcal{B}' y + \mathcal{C}' z + \text{etc.} + \mathcal{L}'\\
u'' &= \qquad\qquad\quad\;\; \mathcal{C}'' z + \text{etc.} + \mathcal{L}''\\
&\text{etc.}
\end{aligned}$$

Hinc erit indefinite

$$\begin{aligned}
\tfrac{1}{2}\mathrm{d}\Omega &= \xi \mathrm{d}x + \eta \mathrm{d}y + \zeta \mathrm{d}z + \text{etc.}\\
&= \tfrac{u^0 \mathrm{d}u^0}{\mathcal{A}^0} + \tfrac{u' \mathrm{d}u'}{\mathcal{B}'} + \tfrac{u'' \mathrm{d}u''}{\mathcal{C}''} + \text{etc.}\\
&= u^0(\mathrm{d}x + \tfrac{\mathcal{B}^0}{\mathcal{A}^0}\mathrm{d}y + \tfrac{\mathcal{C}^0}{\mathcal{A}^0}\mathrm{d}z + \text{etc.})\\
&\quad + u'(\mathrm{d}y + \tfrac{\mathcal{C}'}{\mathcal{B}'}\mathrm{d}z + \text{etc.}) + u''(\mathrm{d}z + \text{etc.}) + \text{etc.}
\end{aligned}$$

unde colligimus (II)

$$\begin{aligned}
\xi &= u^0\\
\eta &= \tfrac{\mathcal{B}^0}{\mathcal{A}^0}u^0 + u'\\
\zeta &= \tfrac{\mathcal{C}^0}{\mathcal{A}^0}u^0 + \tfrac{\mathcal{C}'}{\mathcal{B}'}u' + u''\\
&\text{etc.}
\end{aligned}$$

Supponamus, hinc derivari formulas sequentes (III)

$$\begin{aligned}
u^0 &= \xi\\
u' &= A'\xi + \eta\\
u'' &= A''\xi + B''\eta + \zeta\\
&\text{etc.}
\end{aligned}$$

Iam e differentiali complete aequationis

$$\Omega = \xi(x - A) + \eta(y - B) + \zeta(z - C) + \text{etc.} + M$$

subracta aequatione

$$\tfrac{1}{2}\mathrm{d}\Omega = \xi \mathrm{d}x + \eta \mathrm{d}y + \zeta \mathrm{d}z + \text{etc.}$$

sequitur

$$\tfrac{1}{2}\mathrm{d}\Omega = (x - A)\mathrm{d}\xi + (y - B)\mathrm{d}\eta + (z - C)\mathrm{d}\zeta + \text{etc.}$$

32.

To derive the method set (I)

$$\begin{aligned} u^0 &= \mathcal{A}^0 x + \mathcal{B}^0 y + \mathcal{C}^0 z + \text{etc.} + \mathcal{L}^0 \\ u' &= \qquad\quad \mathcal{B}' y + \mathcal{C}' z + \text{etc.} + \mathcal{L}' \\ u'' &= \qquad\qquad\qquad \mathcal{C}'' z + \text{etc.} + \mathcal{L}'' \\ &\text{etc.} \end{aligned}$$

From this we have

$$\begin{aligned} \tfrac{1}{2}\mathrm{d}\Omega &= \xi \mathrm{d}x + \eta \mathrm{d}y + \zeta \mathrm{d}z + \text{etc.} \\ &= \tfrac{u^0 \mathrm{d}u^0}{\mathcal{A}^0} + \tfrac{u' \mathrm{d}u'}{\mathcal{B}'} + \tfrac{u'' \mathrm{d}u''}{\mathcal{C}''} + \text{etc.} \\ &= u^0(\mathrm{d}x + \tfrac{\mathcal{B}^0}{\mathcal{A}^0}\mathrm{d}y + \tfrac{\mathcal{C}^0}{\mathcal{A}^0}\mathrm{d}z + \text{etc.}) \\ &\quad + u'(\mathrm{d}y + \tfrac{\mathcal{C}'}{\mathcal{B}'}\mathrm{d}z + \text{etc.}) + u''(\mathrm{d}z + \text{etc.}) + \text{etc.} \end{aligned}$$

and hence (II)

$$\begin{aligned} \xi &= u^0 \\ \eta &= \tfrac{\mathcal{B}^0}{\mathcal{A}^0} u^0 + u' \\ \zeta &= \tfrac{\mathcal{C}^0}{\mathcal{A}^0} u^0 + \tfrac{\mathcal{C}'}{\mathcal{B}'} u' + u'' \\ &\text{etc.} \end{aligned}$$

Assume that from these we have derived the following formulas (III):

$$\begin{aligned} u^0 &= \xi \\ u' &= A'\xi + \eta \\ u'' &= A''\xi + B''\eta + \zeta \\ &\text{etc.} \end{aligned}$$

Now by subtracting the equation

$$\tfrac{1}{2}\mathrm{d}\Omega = \xi \mathrm{d}x + \eta \mathrm{d}y + \zeta \mathrm{d}z + \text{etc.}$$

from the total differential of the equation

$$\Omega = \xi(x - A) + \eta(y - B) + \zeta(z - C) + \text{etc.} + M$$

we get

$$\tfrac{1}{2}\mathrm{d}\Omega = (x - A)\mathrm{d}\xi + (y - B)\mathrm{d}\eta + (z - C)\mathrm{d}\zeta + \text{etc.}$$

quae expressio identica esse debet cum hac ex III demanante:

$$\frac{u^0}{A^0}\cdot \mathrm{d}\xi + \frac{u'}{B'}(A'\mathrm{d}\xi + \mathrm{d}\eta) + \frac{u''}{C''}(A''\mathrm{d}\xi + B''\mathrm{d}\eta + \mathrm{d}\zeta) + \text{etc.}$$

Hinc colligimus (IV)

$$\begin{aligned}
x &= \tfrac{u^0}{A^0} + A'\cdot\tfrac{u'}{B'} + A''\cdot\tfrac{u''}{C''} + \text{etc.} + A\\
y &= \phantom{\tfrac{u^0}{A^0} + A'\cdot{}} \tfrac{u'}{B'} + B''\cdot\tfrac{u''}{C''} + \text{etc.} + B\\
z &= \phantom{\tfrac{u^0}{A^0} + A'\cdot\tfrac{u'}{B'} + B''\cdot{}} \tfrac{u''}{C''} + \text{etc.} + C\\
&\text{etc.}
\end{aligned}$$

Substituendo in his expressionibus pro $u^0$, $u'$, $u''$, etc. valores earum ex III depromtos eliminatio indefinita absoluta erit. Et quidem ad pondera determinanda habebimus (V)

$$\begin{aligned}
[\alpha\alpha] &= \tfrac{1}{A^0} + \tfrac{A'A'}{B'} + \tfrac{A''A''}{C''} + \tfrac{A'''A'''}{D'''} + \text{etc.}\\
[\beta\beta] &= \tfrac{1}{B'} + \tfrac{B''B''}{C''} + \tfrac{B'''B'''}{D'''} + \text{etc.}\\
[\gamma\gamma] &= \tfrac{1}{C''} + \tfrac{C'''C'''}{D'''} + \text{etc.}\\
&\text{etc.}
\end{aligned}$$

quarum formularum simplicitas nihil desiderandum relinquit. Ceterum etiam pro coëfficientibus reliquis $[\alpha\beta]$, $[\alpha\gamma]$, $[\beta\gamma]$, etc. formulae aeque simplices prodeunt, quas tamen, quum ilorum usus sit rarior, hic apponere supersedemus.

33.

Propter rei gravitatem, et ut omnia ad calculum parata sint, etiam formulas explicitas ad determinationem coëfficientium $A$, $A'$, $A''$, etc. $B$, $B'$, $B''$, etc., etc. hic adscribere visum est. Duplici modo hic calculus adornari potest, quum aequationes identicae prodire debeant, tum si valores ipsarum $u^0$, $u'$, $u''$, etc. ex III depromti in II substituuntur, tum ex substitutione valorum ipsarum $\xi$, $\eta$, $\zeta$, etc. ex II in III. Prior modus haec formularum systemata subministrat:

$$\begin{aligned}
&\tfrac{B^0}{A^0} + A' = 0\\
&\tfrac{C^0}{A^0} + \tfrac{C'}{B'}\cdot A' + A'' = 0\\
&\tfrac{D^0}{A^0} + \tfrac{D'}{B'}\cdot A' + \tfrac{D''}{C''}\cdot A'' + A''' = 0
\end{aligned}$$

This expression must be the same as the expression from III:

$$\frac{u^0}{A^0}\cdot \mathrm{d}\xi + \frac{u'}{B'}(A'\mathrm{d}\xi + \mathrm{d}\eta) + \frac{u''}{C''}(A''\mathrm{d}\xi + B''\mathrm{d}\eta + \mathrm{d}\zeta) + \text{etc.}$$

From this we find that (IV)

$$\begin{aligned} x &= \frac{u^0}{A^0} + A'\cdot\frac{u'}{B'} + A''\cdot\frac{u''}{C''} + \text{etc.} + A \\ y &= \frac{u'}{B'} + B''\cdot\frac{u''}{C''} + \text{etc.} + B \\ z &= \frac{u''}{C''} + \text{etc.} + C \\ &\text{etc.} \end{aligned}$$

We may now complete the general elimination by replacing $u^0$, $u'$, $u''$, etc. in these formulas with the values in III. For the weights, we get (V)

$$\begin{aligned} [\alpha\alpha] &= \frac{1}{A^0} + \frac{A'A'}{B'} + \frac{A''A''}{C''} + \frac{A'''A'''}{D'''} + \text{etc.} \\ [\beta\beta] &= \frac{1}{B'} + \frac{B''B''}{C''} + \frac{B'''B'''}{D'''} + \text{etc.} \\ [\gamma\gamma] &= \frac{1}{C''} + \frac{C'''C'''}{D'''} + \text{etc.} \\ &\text{etc.} \end{aligned}$$

The simplicity of these formulas leaves nothing to be desired. In addition, there are equally simple formulas for the other coefficients $[\alpha\beta]$, $[\alpha\gamma]$, $[\beta\gamma]$, etc. But they are seldom used, and we will omit them.

## 33.

Since the determination of weights is very important, we will write down formulas for the coefficients $A$, $A'$, $A''$, etc., $B$, $B'$, $B''$, etc., etc., so that everything needed for the calculation will be available. This can be done in two ways. We can either replace $u^0$, $u'$, $u''$, etc. in II by their values from III, or we can replace the values $\xi$, $\eta$, $\zeta$, etc. in III by their values in II.

The first way gives the following collection of formulas. First

$$\begin{aligned} &\frac{B^0}{A^0} + A' = 0 \\ &\frac{C^0}{A^0} + \frac{C'}{B'}\cdot A' + A'' = 0 \\ &\frac{D^0}{A^0} + \frac{D'}{B'}\cdot A' + \frac{D''}{C''}\cdot A'' + A''' = 0 \end{aligned}$$

etc. unde inveniuntur $A'$, $A''$, $A'''$, etc.

$$\frac{\mathcal{C}'}{\mathcal{B}'} + B'' = 0$$
$$\frac{\mathcal{D}'}{\mathcal{B}'} + \frac{\mathcal{D}''}{\mathcal{C}''} \cdot B'' + B''' = 0$$

etc. unde inveniuntur $B''$, $B'''$, etc.

$$\frac{\mathcal{D}''}{\mathcal{C}''} + C''' = 0$$

etc. unde inveniuntur $C'''$ etc. Et sic porro.

Alter modus has formulas suggerit:

$$\mathcal{A}^0 A' e+ \mathcal{B}^0 = 0$$

unde habetur $A'$:

$$\begin{aligned} \mathcal{A}^0 A'' + \mathcal{B}^0 B'' + \mathcal{C}^0 &= 0 \\ \mathcal{B}' B'' + \mathcal{C}' &= 0 \end{aligned}$$

unde inveniuntur $B''$ et $A''$:

$$\begin{aligned} \mathcal{A}^0 A''' + \mathcal{B}^0 B''' + \mathcal{C}^0 C''' + \mathcal{D}^0 &= 0 \\ \mathcal{B}' B''' + \mathcal{C}' C''' + \mathcal{D}' &= 0 \\ \mathcal{C}'' C''' + \mathcal{D}'' &= 0 \end{aligned}$$

unde inveniuntur $C'''$, $C'''$, $A'''$. Et sic porro.

Uterque modus aeque fere commodus est, si pondera determinationum cunctarum $x$, $y$, $z$, etc. desiderantur; quoties vero e quantitatibus $[\alpha\alpha]$, $[\beta\beta]$, $[\gamma\gamma]$, etc. una tantum vel altera requiritur, manifesto systema prius longe praeferendum erit.

Ceterum combinatio aequationum I cum IV ad easdem formulas perducit, insuperque calculum duplicem ad eruendos valores maxime plausibles $A$, $B$, $C$, etc. ipsos suppeditat, puto *primo*

$$\begin{aligned} A &= -\frac{\mathcal{L}^0}{\mathcal{A}^0} - A'\frac{\mathcal{L}'}{\mathcal{B}'} - A''\frac{\mathcal{L}''}{\mathcal{C}''} - A'''\frac{\mathcal{L}'''}{\mathcal{D}'''} - \text{etc.} \\ B &= -\frac{\mathcal{L}'}{\mathcal{B}'} - B''\frac{\mathcal{L}''}{\mathcal{C}''} - B'''\frac{\mathcal{L}'''}{\mathcal{D}'''} - \text{etc.} \\ C &= -\frac{\mathcal{L}''}{\mathcal{C}''} - C'''\frac{\mathcal{L}'''}{\mathcal{D}'''} - \text{etc.} \\ &\text{etc.} \end{aligned}$$

Calculus alter identicus est cum vulgari, ubi statuitur $u^0 = 0$, $u' = 0$, $u'' = 0$, etc.

etc. from which we may determine $A'$, $A''$, $A'''$, etc. Next

$$\frac{\mathcal{C}'}{\mathcal{B}'} + B'' = 0$$

$$\frac{\mathcal{D}'}{\mathcal{B}'} + \frac{\mathcal{D}''}{\mathcal{C}''} \cdot B'' + B''' = 0$$

etc., from which we may determine $B''$, $B'''$, etc. Then

$$\frac{\mathcal{D}''}{\mathcal{C}''} + C''' = 0$$

etc., from which we may determine $C'''$, etc., and so on.

The second approach yields the following formulas:

$$\mathcal{A}^0 A' e + \mathcal{B}^0 = 0$$

from which we get $A'$;

$$\begin{aligned} \mathcal{A}^0 A'' + \mathcal{B}^0 B'' + \mathcal{C}^0 &= 0 \\ \mathcal{B}' B'' + \mathcal{C}' &= 0 \end{aligned}$$

from which we get $B''$ and $A''$;

$$\begin{aligned} \mathcal{A}^0 A''' + \mathcal{B}^0 B''' + \mathcal{C}^0 C''' + \mathcal{D}^0 &= 0 \\ \mathcal{B}' B''' + \mathcal{C}' C''' + \mathcal{D}' &= 0 \\ \mathcal{C}'' C''' + \mathcal{D}'' &= 0 \end{aligned}$$

from which we get $C'''$, $C'''$, $A'''$; and so on.

When all the weights of the estimates $x$, $y$, $z$, etc. are needed, the two methods are roughly equivalent. When only the one or two of the quantities $[\alpha\alpha]$, $[\beta\beta]$, $[\gamma\gamma]$, etc. are required, the first method is clearly preferable.

A combination of equations I and IV leads to the same formulas. In addition, it yields two methods for computing the most reliable values $A$, $B$, $C$, etc. The *first* is

$$\begin{aligned} A &= -\frac{\mathcal{L}^0}{\mathcal{A}^0} - A'\frac{\mathcal{L}'}{\mathcal{B}'} - A''\frac{\mathcal{L}''}{\mathcal{C}''} - A'''\frac{\mathcal{L}'''}{\mathcal{D}'''} - \text{etc.} \\ B &= -\frac{\mathcal{L}'}{\mathcal{B}'} - B''\frac{\mathcal{L}''}{\mathcal{C}''} - B'''\frac{\mathcal{L}'''}{\mathcal{D}'''} - \text{etc.} \\ C &= -\frac{\mathcal{L}''}{\mathcal{C}''} - C'''\frac{\mathcal{L}'''}{\mathcal{D}'''} - \text{etc.} \end{aligned}$$

etc.

The second method is the usual one, in which we set $u^0 = 0$, $u' = 0$, $u'' = 0$, etc.

## 34.

Quae in art. 32 exposuimus, sunt tantummodo casus speciales theorematis generalioris, quod ita se habet:

THEOREM. *Designet $t$ functionem linearem indeterminatarum $x$, $y$, $z$, etc. hanc*

$$t = fx + gy + hz + \text{etc.} + k$$

*quae transmutata in functionem indeterminatarum $u^0 = 0$, $u' = 0$, $u'' = 0$, etc. fiat*

$$t = k^0u^0 + k'u' + k''u'' + \text{etc.} + K$$

*Quibus ita se habentibus erit $K$ valor maxime plausibilis ipsius $t$, atque pondus huius determinationis*

$$= \frac{1}{\mathcal{A}^0k^0k^0 + \mathcal{B}'k'k' + \mathcal{C}''k''k'' + \text{etc.}}$$

*Dem.* Pars prior theorematis inde patet, quod valor maxime plausibilis ipsius $t$ valoribus $u^0 = 0$, $u' = 0$, $u'' = 0$, etc. respondere debet. Ad posteriorem demonstrandum observamus, quoniam $\frac{1}{2}\mathrm{d}\Omega = \xi\mathrm{d}x + \eta\mathrm{d}y + \zeta\mathrm{d}z + \text{etc.}$ atque $\mathrm{d}t = f\mathrm{d}x + g\mathrm{d}y + h\mathrm{d}z + \text{etc.}$, esse, pro $\xi = f$, $\eta = g$, $\zeta = h$ etc., independenter a valoribus differentialium $\mathrm{d}x$, $\mathrm{d}y$, $\mathrm{d}z$, etc.

$$\mathrm{d}\Omega = 2\mathrm{d}t$$

Hinc vero sequitur, pro iisdem valoribus $\xi = f$, $\eta = g$, $\zeta = h$, etc., fieri

$$\frac{u^0}{\mathcal{A}^0}\mathrm{d}u^0 + \frac{u'}{\mathcal{B}'}\mathrm{d}u' + \frac{u''}{\mathcal{C}''}\mathrm{d}u'' + \text{etc.} = k^0\mathrm{d}u^0 + k'\mathrm{d}u' + k''\mathrm{d}u'' + \text{etc.}$$

Iam facile perspicitur, si $\mathrm{d}x$, $\mathrm{d}y$, $\mathrm{d}z$, etc. sint ab invicem independentes, etiam $\mathrm{d}u^0$, $\mathrm{d}u'$, $\mathrm{d}u''$, etc. ab invicem independentes esse; unde colligimus, pro $\xi = f$, $\eta = g$, $\zeta = h$, etc. esse

$$u^0 = \mathcal{A}^0k^0 \quad u' = \mathcal{B}'k' \quad u'' = \mathcal{C}''k'' \quad \text{etc.}$$

Quamobrem valor ipsius $\Omega$, iisdem valoribus respondens erit

$$= \mathcal{A}^0k^0k^0 + \mathcal{B}'k'k' + \mathcal{C}''k''k'' + \text{etc.} + M$$

unde per art. 29 theorematis nostri veritas protinus demanat.

## 34.

The results we derived in Art. 32 are just special cases of the following general theorem.

THEOREM. *Let the linear function*

$$t = fx + gy + hz + \text{etc.} + k$$

*of $x$, $y$, $z$, etc. be written as a linear function of the unknowns $u^0 = 0$, $u' = 0$, $u'' = 0$, etc. as follows:*

$$t = k^0u^0 + k'u' + k''u'' + \text{etc.} + K.$$

*Then the most reliable value of $t$ is $K$, and its weight is*

$$\frac{1}{\mathcal{A}^0k^0k^0 + \mathcal{B}'k'k' + \mathcal{C}''k''k'' + \text{etc.}}.$$

*Proof.* The first part of the theorem is obvious, since the most reliable value of $t$ corresponds to the values $u^0 = 0$, $u' = 0$, $u'' = 0$, etc.

To prove the second part note that $\frac{1}{2}\mathrm{d}\Omega = \xi\mathrm{d}x + \eta\mathrm{d}y + \zeta\mathrm{d}z + \text{etc.}$ and $\mathrm{d}t = f\mathrm{d}x + g\mathrm{d}y + h\mathrm{d}z + \text{etc.}$ Hence if $\xi = f$, $\eta = g$, $\zeta = h$, etc. then

$$\mathrm{d}\Omega = 2\mathrm{d}t,$$

whatever the differentials $\mathrm{d}x$, $\mathrm{d}y$, $\mathrm{d}z$, etc. It follows that for $\xi = f$, $\eta = g$, $\zeta = h$, etc. we have

$$\frac{u^0}{\mathcal{A}^0}\mathrm{d}u^0 + \frac{u'}{\mathcal{B}'}\mathrm{d}u' + \frac{u''}{\mathcal{C}''}\mathrm{d}u'' + \text{etc.} = k^0\mathrm{d}u^0 + k'\mathrm{d}u' + k''\mathrm{d}u'' + \text{etc.}$$

Now it is easy to see that if $\mathrm{d}x$, $\mathrm{d}y$, $\mathrm{d}z$, etc. are mutually independent so are $\mathrm{d}u^0$, $\mathrm{d}u'$, $\mathrm{d}u''$, etc. Hence for $\xi = f$, $\eta = g$, $\zeta = h$, etc. we have

$$u^0 = \mathcal{A}^0k^0, \quad u' = \mathcal{B}'k', \quad u'' = \mathcal{C}''k'', \quad \text{etc.}$$

The corresponding value of $\Omega$ is

$$\mathcal{A}^0k^0k^0 + \mathcal{B}'k'k' + \mathcal{C}''k''k'' + \text{etc.} + M,$$

and our theorem follows immediately from Art. 29.

Ceterum si transformationem functionis $t$ immediate, i.e. absque cognitione substitutionum IV. art. 32, perficere cupimus, praesto sunt formulae:

$$\begin{aligned} f &= \mathcal{B}^0 k^0 \\ g &= \mathcal{B}^0 k^0 + \mathcal{B}' k' \\ h &= \mathcal{C}^0 k^0 + \mathcal{C}' k' + \mathcal{C}'' k'' \ \text{ etc.}, \end{aligned}$$

unde coëfficientes $k^0$, $k'$, $k''$, etc. deinceps determinabuntur, tandemque habebitur

$$K = k - \mathcal{L}^0 k^0 - \mathcal{L}' k' - \mathcal{L}'' k'' - \text{etc.}$$

35.

Tractatione peculiari dignum est problema sequens, tum propter utilitatem practicam, tum propter solutionis concinnitatem.

*Invenire mutationes valorum maxime plausiblium incognitarum ab accessione aequationis novae productas, nec non pondera novarum determinationum.*

Retinebimus designationes in praecedentibus adhibitas, ita ut aequationes primitivae, ad pondus $= 1$ reductae, sint hae $v = 0$, $v' = 0$, $v'' = 0$, etc.; aggregatum indefinitum $vv + v'v' + v''v'' + \text{etc.} = \Omega$; porro ut $\xi$, $\eta$, $\zeta$, etc. sint quotientes differentiales partiales

$$\frac{\mathrm{d}\Omega}{2\mathrm{d}x}, \quad \frac{\mathrm{d}\Omega}{2\mathrm{d}y}, \quad \frac{\mathrm{d}\Omega}{2\mathrm{d}z}, \quad \text{etc.}$$

denique ut ex eliminatione indefinita sequatur

$$\left.\begin{aligned} x &= A + [\alpha\alpha]\xi + [\alpha\beta]\eta + [\alpha\gamma]\zeta + \text{etc.} \\ y &= B + [\beta\alpha]\xi + [\beta\beta]\eta + [\beta\gamma]\zeta + \text{etc.} \\ z &= C + [\gamma\alpha]\xi + [\gamma\beta]\eta + [\gamma\gamma]\zeta + \text{etc.} \end{aligned}\right\} \text{(I)}$$

Iam supposamus, accedere aequationem novam $v^* = 0$ (proxime veram, et cuius pondus $= 1$), et inquiramus, quantas mutationes hinc nacturi sint tum valores incognitarum maxime plausibiles $A$, $B$, $C$, etc., tum coëfficientes $[\alpha\alpha]$, $[\beta\beta]$, etc.

Statuamus $\Omega + v^*v^* = \Omega^*$, $\frac{\mathrm{d}\Omega^*}{2\mathrm{d}x} = \xi^*$, $\frac{\mathrm{d}\Omega^*}{2\mathrm{d}y} = \eta^*$, $\frac{\mathrm{d}\Omega^*}{2\mathrm{d}z} = \zeta^*$, etc. supponamusque, hinc per eliminationem sequi

$$x = A^* + [\alpha\alpha^*]\xi^* + [\alpha\beta^*]\eta^* + [\alpha\gamma^*]\zeta^* \ \text{ etc.}$$

Denique sit

$$v^* = fx + gy + hz + \text{etc.} + k$$

In addition, if we want to transform $t$ directly, i.e., without knowing the transformations IV of Art. 32, we may determine $k^0$, $k'$, $k''$, etc. in succession from the formulas

$$\begin{aligned} f &= \mathcal{B}^0 k^0 \\ g &= \mathcal{B}^0 k^0 + \mathcal{B}' k' \\ h &= \mathcal{C}^0 k^0 + \mathcal{C}' k' + \mathcal{C}'' k'' \ \text{ etc.}, \end{aligned}$$

Finally, we have

$$K = k - \mathcal{L}^0 k^0 - \mathcal{L}' k' - \mathcal{L}'' k'' - \text{etc.}$$

35.

The following problem is worth special treatment because of its practical utility and its elegant solution.

*Determine the changes in the most reliable values of the unknowns and in their weights when a new equation is appended.*

We will retain the notation used earlier, so that the basic equations, reduced to weight one, are $v = 0$, $v' = 0$, $v'' = 0$, etc. The general sum $vv + v'v' + v''v'' +$ etc. is $\Omega$, and $\xi$, $\eta$, $\zeta$, etc. are its partial differential quotients

$$\frac{d\Omega}{2dx}, \quad \frac{d\Omega}{2dy}, \quad \frac{d\Omega}{2dz}, \quad \text{etc.}$$

Finally,

$$\left.\begin{aligned} x &= A + [\alpha\alpha]\xi + [\alpha\beta]\eta + [\alpha\gamma]\zeta + \text{etc.} \\ y &= B + [\beta\alpha]\xi + [\beta\beta]\eta + [\beta\gamma]\zeta + \text{etc.} \\ z &= C + [\gamma\alpha]\xi + [\gamma\beta]\eta + [\gamma\gamma]\zeta + \text{etc.} \end{aligned}\right\} \text{(I)}$$

which follows from a general elimination. Now suppose that we have a new equation $v^* = 0$ (which is approximately true and has weight one), and let us ask how it changes the most reliable values $A$, $B$, $C$, etc. of the unknowns and the coefficients $[\alpha\alpha]$, $[\beta\beta]$, etc.

Set $\Omega + v^*v^* = \Omega^*$, $\frac{d\Omega^*}{2dx} = \xi^*$, $\frac{d\Omega^*}{2dy} = \eta^*$, $\frac{d\Omega^*}{2dz} = \zeta^*$, etc. and suppose that by elimination we have

$$x = A^* + [\alpha\alpha^*]\xi^* + [\alpha\beta^*]\eta^* + [\alpha\gamma^*]\zeta^* \ \text{ etc.}$$

Next let

$$v^* = fx + gy + hz + \text{etc.} + k,$$

prodeat inde, substitutis pro $x$, $y$, $z$, etc. valoribus ex (I),

$$v^* = F\xi + G\eta + H\zeta + \text{etc.} + K$$

statuaturque

$$Ff + Gg + Hh + \text{etc.} = \omega$$

Manifesto $K$ erit valor maxime plausiblis functionis $v^*$, quatenus ex aequationibus primitivis sequitur, sine respectu valoris 0, quem observatio accessoria praebuit, atque $\frac{1}{\omega}$ pondus istius determinationis.

Iam habemus

$$\xi^* = \xi + fv^*, \quad \eta^* = \eta + gv^*, \quad \zeta^* = \zeta^* + hv^*, \quad \text{etc.}$$

adeoque

$$F\xi^* + G\eta^* + H\zeta^* + \text{etc.} + K = v^*(1 + Ff + Gg + Hh + \text{etc.})$$

sive

$$v^* = \tfrac{F\xi^* + G\eta^* + H\zeta^* + \text{etc.} + K}{1+\omega}$$

Perinde fit

$$\begin{aligned} x &= A + [\alpha\alpha]\xi^* + [\alpha\beta]\eta^* + [\alpha\gamma]\zeta^* + \text{etc.} - v^*(f[\alpha\alpha] + g[\alpha\beta] + h[\alpha\gamma] + \text{etc.}) \\ &= A + [\alpha\alpha]\xi^* + [\alpha\beta]\eta^* + [\alpha\gamma]\zeta^* + \text{etc.} - Fv^* \\ &= A + [\alpha\alpha]\xi^* + [\alpha\beta]\eta^* + [\alpha\gamma]\zeta^* + \text{etc.} - \tfrac{F}{1+\omega}(F\xi^* + G\eta^* + H\zeta^* + \text{etc.} + K) \end{aligned}$$

Hinc itaque colligimus

$$A^* = A - \tfrac{FK}{1+\omega}$$

qui erit valor maxime plausibilis ipsius $x$ ex *omnibus* observationibus;

$$[\alpha\alpha^*] = [\alpha\alpha] - \tfrac{FF}{1+\omega}$$

adeoque pondus istius determinationis

$$= \frac{1}{[\alpha\alpha] - \frac{FF}{1+\omega}}$$

Prorsus eodem modo invenitur valor maxime plausibilis ipsius $y$, *omnibus* observationibus superstructus

$$B^* = B - \tfrac{GK}{1+\omega}$$

and suppose that the equation

$$v^* = F\xi + G\eta + H\zeta + \text{etc.} + K$$

results from replacing $x$, $y$, $z$, etc. by their values in (I). Finally set

$$Ff + Gg + Hh + \text{etc.} = \omega.$$

Clearly $K$ is the most reliable value of $v^*$ arising from the original observations when no account is taken of the value 0 which is furnished by the new observation. The weight of this estimate is $\frac{1}{\omega}$.

Now

$$\xi^* = \xi + fv^*, \quad \eta^* = \eta + gv^*, \quad \zeta^* = \zeta^* + hv^*, \quad \text{etc.},$$

and

$$F\xi^* + G\eta^* + H\zeta^* + \text{etc.} + K = v^*(1 + Ff + Gg + Hh + \text{etc.})$$

or

$$v^* = \frac{F\xi^* + G\eta^* + H\zeta^* + \text{etc.} + K}{1+\omega}.$$

Hence

$$\begin{aligned} x &= A + [\alpha\alpha]\xi^* + [\alpha\beta]\eta^* + [\alpha\gamma]\zeta^* + \text{etc.} - v^*(f[\alpha\alpha] + g[\alpha\beta] + h[\alpha\gamma] + \text{etc.}) \\ &= A + [\alpha\alpha]\xi^* + [\alpha\beta]\eta^* + [\alpha\gamma]\zeta^* + \text{etc.} - Fv^* \\ &= A + [\alpha\alpha]\xi^* + [\alpha\beta]\eta^* + [\alpha\gamma]\zeta^* + \text{etc.} - \tfrac{F}{1+\omega}(F\xi^* + G\eta^* + H\zeta^* + \text{etc.} + K). \end{aligned}$$

From this we find that the most reliable value of $x$ from *all* the observations is

$$A^* = A - \frac{FK}{1+\omega}.$$

Moreover,

$$[\alpha\alpha^*] = [\alpha\alpha] - \frac{FF}{1+\omega}$$

so that the weight of this estimate is

$$\frac{1}{[\alpha\alpha] - \frac{FF}{1+\omega}}.$$

In the same way we find that the most reliable value of $y$ based on *all* the observations is

$$B^* = B - \frac{GK}{1+\omega}$$

atque pondus huius determinationis

$$= \frac{1}{[\beta\beta] - \frac{GG}{1+\omega}}.$$

et sic porro. Q. E. I.

Liceat huic solutioni quasdam annotationes adiicere.

I. Substitutis his novis valoribus $A^*$, $B^*$, $C^*$, etc., functio $v^*$ obtinet valorem maxime plausibilem

$$K - \tfrac{K}{1+\omega}(Ff + Gg + Hh + \text{etc.}) = \tfrac{K}{1+\omega}$$

Et quum indefinite sit

$$v^* = \tfrac{F}{1+\omega}.\xi^* + \tfrac{G}{1+\omega}.\eta^* + \tfrac{H}{1+\omega}.\zeta^* + \text{etc.} + \tfrac{K}{1+\omega}$$

pondus istius determinationis per principia art. 29 eruitur

$$\frac{1+\omega}{Ff+Gg+Hh+\text{etc.}} = \frac{1}{\omega} + 1$$

Eadem immediate resultant ex applicatione regulae in fine art. 21 traditae; scilicet complexus aequationum primitivarum praebuerat determinationem $v^* = K$ cum pondere $= \frac{1}{\omega}$, dein observatio nova dedit determinininationem aliam, ab illa independentem, $v^* = 0$, cum pondere $= 1$, quibus combinatis prodit determinatio $v^* = \frac{K}{1+\omega}$ cum pondere $= \frac{1}{\omega} + 1$.

II. Hinc porro sequitur, quum pro $x = A^*$, $y = B^*$, $z = C^*$, etc. esse debeat $\xi^* = 0$, $\eta^* = 0$, $\zeta^* = 0$, etc., pro iisdem valoribus fieri

$$\xi = -\tfrac{fK}{1+\omega}, \quad \eta = -\tfrac{gK}{1+\omega}, \quad \zeta = -\tfrac{hK}{1+\omega}, \quad \text{etc.}$$

nec non, quoniam indefinite $\Omega = \xi(x - A) + \eta(y - B) + \zeta(z - C) + \text{etc.} + M$,

$$\Omega = \tfrac{KK}{(1+\omega)^2}(Ff + Gg + Hh + \text{etc.}) + M = M + \tfrac{\omega KK}{(1+\omega)^2}$$

denique, quoniam indefinite $\Omega^* = \Omega + v^*v^*$,

$$\Omega^* = M + \tfrac{\omega KK}{(1+\omega)^2} + \tfrac{KK}{(1+\omega)^2} = M + \tfrac{KK}{1+\omega}$$

III. Comparando haec cum iis, quae in art. 30 docuimus, animadvertimus, functionem $\Omega$ hic valorem minimum obtinere, quem pro valore determinato functionis $\frac{K}{1+\omega}$ accipere potest.

and the weight of this estimate is

$$\frac{1}{[\beta\beta]-\frac{GG}{1+\omega}},$$

and so on. This solves the problem.

We may add some comments to this solution.

I. If we evaluate $v^*$ at the new values $A^*$, $B^*$, $C^*$, etc., we obtain the most reliable value

$$K - \tfrac{K}{1+\omega}(Ff + Gg + Hh + \text{etc.}) = \tfrac{K}{1+\omega}.$$

Since

$$v^* = \tfrac{F}{1+\omega}.\xi^* + \tfrac{G}{1+\omega}.\eta^* + \tfrac{H}{1+\omega}.\zeta^* + \text{etc.} + \tfrac{K}{1+\omega}$$

holds generally, according to the principles of Art. 29 the weight of this estimate is

$$\frac{1+\omega}{Ff+Gg+Hh+\text{etc.}} = \frac{1}{\omega} + 1.$$

The same results follow from a direct application of the rule treated in Art. 21. Specifically, the original system of equations gives an estimate $v^* = K$ with weight $\frac{1}{\omega}$, while the new observation gives an independent estimate $v^* = 0$ with weight one. The combination produces an estimate $v^* = \frac{K}{1+\omega}$ with weight $\frac{1}{\omega} + 1$.

II. Since $\xi^* = 0$, $\eta^* = 0$, $\zeta^* = 0$, etc. when $x = A^*$, $y = B^*$, $z = C^*$, etc., for these values

$$\xi = -\tfrac{fK}{1+\omega}, \quad \eta = -\tfrac{gK}{1+\omega}, \quad \zeta = -\tfrac{hK}{1+\omega}, \quad \text{etc.}$$

Since $\Omega = \xi(x-A) + \eta(y-B) + \zeta(z-C) + \text{etc.} + M$ holds generally, we have

$$\Omega = \tfrac{KK}{(1+\omega)^2}(Ff + Gg + Hh + \text{etc.}) + M = M + \tfrac{\omega KK}{(1+\omega)^2}.$$

Finally since $\Omega^* = \Omega + v^*v^*$, it follows from the first comment that

$$\Omega^* = M + \tfrac{\omega KK}{(1+\omega)^2} + \tfrac{KK}{(1+\omega)^2} = M + \tfrac{KK}{1+\omega}.$$

III. Comparing this with what we showed in Art. 30, we see that $\Omega$ attains the smallest value it can when the function $v^*$ assumes the fixed value $\frac{K}{1+\omega}$.

36.

Problematis alius, praecedenti affinis, puta

*Investigare mutationes valorum maxime plausibilium incognitarum, a mutato pondere unius ex observationibus primitivis oriundas, nec non pondera novarum determinationum.*

solutionem tantummodo hic adscribemus, demonstrationem, quae ad instar art. praec. facile absolvitur, brevatatis caussa supprimentes.

Supponamus, peracto demum calculo animadverti, alicui observationum pondus seu nimis parvum, seu nimis magnum tributum esse, e.g. observationi primae, quae dedit $V = L$, loco ponderis $p$ in calculo adhibiti rectius tribui pondus $= p^*$. Tunc haud opus erit calculum integrum repetere, sed commodius correctiones per formulas sequentes computare licebit.

Valores incognitarum maxime plausibles correcti erunt hi:

$$\begin{aligned} x &= A - \frac{(p^*-p)\alpha\lambda}{p+(p^*-p)(a\alpha+b\beta+c\gamma+\text{etc.})} \\ y &= B - \frac{(p^*-p)\beta\lambda}{p+(p^*-p)(a\alpha+b\beta+c\gamma+\text{etc.})} \\ z &= C - \frac{(p^*-p)\gamma\lambda}{p+(p^*-p)(a\alpha+b\beta+c\gamma+\text{etc.})} \end{aligned}$$

etc. ponderaque harum determinationum invenientur, dividendo unitatem resp. per

$$\begin{aligned} &[\alpha\alpha] - \frac{(p^*-p)\alpha\alpha}{p+(p^*-p)(a\alpha+b\beta+c\gamma+\text{etc.})} \\ &[\beta\beta] - \frac{(p^*-p)\beta\beta}{p+(p^*-p)(a\alpha+b\beta+c\gamma+\text{etc.})} \\ &[\gamma\gamma] - \frac{(p^*-p)\gamma\gamma}{p+(p^*-p)(a\alpha+b\beta+c\gamma+\text{etc.})} \text{ etc.} \end{aligned}$$

Haec solutio simul complectitur casum, ubi peracto calculo percipitur, unam ex observationibus omnino reiici debuisse, quum hoc idem sit ac si facias $p^* = 0$; et perinde valor $p^* = \infty$ refertur ad casum eum, ubi aequatio $V = L$, quae in calculo tamquam approximata tractata erat, revera praecisione absoluta gaudet.

Ceterum quoties vel aequationibus, quibus calculus superstructus erat, *plures* novae accedunt, vel *pluribus* ex illis pondera erronea tributa esse percipitur, computus correctionum nimis complicatus evaderet; quocirca in tali casu calculum ab integro reficere praestabit.

37.

In artt. 15. 16 methodum explicavimus, observationum praecisionem proxi-

36.

Another problem, related to the preceding, is the following.

*Determine the changes in the most reliable values of the unknowns and in their weights when the weight of one of the original observations is changed.*

For the sake of brevity we only present the solution and omit its proof, which is analogous to the proof in the preceding article.

Suppose that after the calculation is finished we notice that some observation has the wrong weight; for example, suppose the first observation, which gives $V = L$, should have had weight $p^*$ instead of $p$. In this case there is no need to repeat the whole calculation. Instead corrections may be computed directly from the following formulas.

The corrected most reliable values of the unknowns are

$$x = A - \frac{(p^*-p)\alpha\lambda}{p+(p^*-p)(a\alpha+b\beta+c\gamma+\text{etc.})}$$
$$y = B - \frac{(p^*-p)\beta\lambda}{p+(p^*-p)(a\alpha+b\beta+c\gamma+\text{etc.})}$$
$$z = C - \frac{(p^*-p)\gamma\lambda}{p+(p^*-p)(a\alpha+b\beta+c\gamma+\text{etc.})}$$

etc. The weights of these estimates are the reciprocals of

$$[\alpha\alpha] - \frac{(p^*-p)\alpha\alpha}{p+(p^*-p)(a\alpha+b\beta+c\gamma+\text{etc.})}$$
$$[\beta\beta] - \frac{(p^*-p)\beta\beta}{p+(p^*-p)(a\alpha+b\beta+c\gamma+\text{etc.})}$$
$$[\gamma\gamma] - \frac{(p^*-p)\gamma\gamma}{p+(p^*-p)(a\alpha+b\beta+c\gamma+\text{etc.})} \quad \text{etc.}$$

This solution includes the case where we find after the fact that one of the observations ought to have been thrown out, which is equivalent to setting $p^* = 0$. It also includes the case $p^* = \infty$, in which an exact value was treated as approximate in the calculation.

When *several* equations are added to the original ones or *several* weights are in error, it becomes too complicated to calculate the corrections. In this case, it is better to redo the calculations from the beginning.

37.

In Arts. 15 and 16 we gave a method for approximating the precision of me determinandi.*) Sed haec methodus supponit, errores, qui revera occurrerint, satis multos exacte cognitos esse, quae conditio, stricte loquendo, rarissime, ne dicam numquam, locum habebit. Quodsi quidem quantitates, quarum valores approximati per observationes innotuerunt, secundum legem cognitum, ab una pluribusve quantitatibus incognitis pendent, harum valores maxime plausibles per methodum quadratorum minimorum eruere licebit, ac dein valores quantitatum, quae observationum obiecta fuerant, illinc computati perparum a valoribus veris discrepare censebuntur, ita ut ipsorum differentias a valoribus observatis eo maiore iure tamquam observationum errores veros adoptare liceat, quo maior fuerit harum multitudo. Hanc praxin sequuti sunt omnes calculatores, qui observationum pracisionem in casibus concretis a posteriore aestimare susceperunt: sed manifesto illa theoretice erronea est, et quamquam in casibus multis ad usus practicos sufficere possit, tamen in aliis enormiter peccare potest. Summopere itaque hoc argumentum dignum est, quod accuratius enodetur.

Retinebimus in hac disquisitione designationes inde ab art. 19 adhibitas. Praxis ea, de qua diximus, quantitates $A$, $B$, $C$, etc. tamquam valores veros ipsarum $x$, $y$, $z$, etc. considerat, et proin ipsas $\lambda$, $\lambda'$, $\lambda''$, etc. tamquam valores veros ipsarum $v$, $v'$, $v''$, etc. Si omnes observationes aequali pracisione gaudent, ipsarumque pondus $p = p' = p''$, etc. pro unitate acceptum est, eadem quantitates, signis mutatis, in illa suppositione observationum errores exhibent, unde praecepta art. 15 praebent observationum errorem medium $m$

$$= \sqrt{\frac{\lambda\lambda + \lambda'\lambda' + \lambda''\lambda'' + \text{etc.}}{\pi}} = \sqrt{\frac{M}{\pi}}$$

Si observationum praecisio non est eadem, quantitates $-\lambda$, $-\lambda'$, $-\lambda''$, etc. exhiberent observationum errores per radices quatratas e ponderibus multiplicatos, praeceptaque art. 16 ad eandem formulam $\sqrt{\frac{M}{\pi}}$ perducerent, iam errorem medium talium observationum, quibus pondus $= 1$ tribuitur, denotantem. Sed manifesto calculus exactus requireret, ut loco quantitatum $\lambda$, $\lambda'$, $\lambda''$, etc. valores functionum $v$, $v'$, $v''$, etc. e valoribus veris ipsarum $x$, $y$, $z$, etc. prodeuntes adhiberentur, i.e. loco ipsius $M$, valor functionis $\Omega$ valoribus veris ipsarum $x$, $y$, $z$, etc. respondens. Qui quamquam assignari nequeat, tamen certi sumus, eum esse maiorem quam $M$

*)Disquisitio de eodem argumento, quam in commentatione anteriore (*Bestimmung der Genauigkeit der Beobachtungen. Zeitschrift für Astronomie und verwandte Wissenschaften* Vol. I, p. 185) tradideramus eidem hypothesi circa indolem functinis probabilitatem errorum exprimentis innixa erat, cui in Theoria motus corporum coelestium methodum quadratorum minimorum superstruxeramus (vid. art. 9, III).

observations.*) But this method presupposes that a sufficient number of the errors themselves are known exactly, a condition which seldom, if ever, holds in practice. However, if some observed quantities depend on one or more unknowns according to a known law, we may find the most reliable values of the unknowns by the method of least squares. If the values of the observed quantities are then computed from them, they will be felt to differ very little from the true values, so that the greater the number of observations the more surely we may take the differences as the true observation errors.

This procedure is used in actual problems by all calculators who try to estimate precision a posteriori. But it is theoretically unsound; and although it is good enough for practical purposes in many cases, in others it can fail spectacularly. Therefore, it is well worthwhile to examine the problem more closely.

In our investigation we will use the same notation we used in Art. 19 and the following articles. The procedure under consideration takes the quantities $A$, $B$, $C$, etc. for the true values $x$, $y$, $z$, etc. and similarly $\lambda$, $\lambda'$, $\lambda''$, etc. for the true values of the functions $v$, $v'$, $v''$, etc. If all the observations are equally precise — say their weights $p = p' = p''$ etc. are one — then by the above reasoning $\lambda$, $\lambda'$, $\lambda''$, etc. with their signs changed represent the observation errors, and by the principles of Art. 15 they yield a mean error in the observations of

$$m = \sqrt{\frac{\lambda\lambda + \lambda'\lambda' + \lambda''\lambda'' + \text{etc.}}{\pi}} = \sqrt{\frac{M}{\pi}}.$$

If the observations are not equally precise, the quantities $-\lambda$, $-\lambda'$, $-\lambda''$, etc. represent the observation errors multiplied by the square roots of their weights, and according to the principles of Art. 16 they lead to the same formula $\sqrt{\frac{M}{\pi}}$, which now gives the weight of the observation error with weight one.

Now it is clear that an exact calculation requires that instead of $\lambda$, $\lambda'$, $\lambda''$, etc. we use the values of the functions $v$, $v'$, $v''$, etc. at the true values of $x$, $y$, $z$, etc. That is, in place of $M$ we should use the value of the function $\Omega$ corresponding to the true values of $x$, $y$, $z$, etc. Although we cannot determine this value, we know that except for the infinitely improbable case where the most reliable values of the

*) An inquiry into the same problem, which I published in an earlier memoir (*Bestimmung der Genauigkeit der Beobachtungen. Zeitschrift für Astronomie und vervandte Wissenschaften* Vol. I, p. 185) was based on the same hypothesis about the probability function of the experimental error that I used to construct the method of least squares in the *Theoria Motus Corporum Coelestium.*

(quippe qui est minimus possibilis), excipiendo casum infinite parum probabilem, ubi incognitarum valores maxime plausibiles exacte cum veris quadrant. In genere itaque affirmare possumus, praxin vulgarem errorem medium iusto minorem producere, sive observationibus praecisionem nimis magnam tribuere. Videamus iam, quid doceat theoria rigorosa.

38.

Ante omnia investigare oportet, quonam modo $M$ ab observationum erroribus veris pendeat. Denotemus hos, ut in art. 28, per $e$, $e'$, $e''$, etc., statuamusque ad maiorem simplicitatem

$$\begin{gathered} e\sqrt{p} = \epsilon, \quad e'\sqrt{p'} = \epsilon', \quad e''\sqrt{p''} = \epsilon'', \text{ etc.,} \quad \text{nec non} \\ m\sqrt{p} = m'\sqrt{p'} = m''\sqrt{p''} = \text{etc.} = \mu \end{gathered}$$

Porro sint valores veri ipsarum $x$, $y$, $z$, etc., resp. $A-x^0$, $B-y^0$, $C-z^0$, etc., quibus respondeant valores ipsarum $\xi$, $\eta$, $\zeta$, etc. hi $-\xi^0$, $-\eta^0$, $-\zeta^0$, etc. Manifesto iisdem respondebunt valores ipsarum $v$, $v'$, $v''$, etc. hi $-\epsilon'$, $-\epsilon''$, etc., ita ut habeatur

$$\begin{aligned} \xi^0 &= a\epsilon + a'\epsilon' + a''\epsilon'' + \text{etc.} \\ \eta^0 &= b\epsilon + b'\epsilon' + b''\epsilon'' + \text{etc.} \\ \zeta^0 &= c\epsilon + c'\epsilon' + c''\epsilon'' + \text{etc.} \end{aligned}$$

etc. nec non

$$\begin{aligned} x^0 &= \alpha\epsilon + \alpha'\epsilon' + \alpha''\epsilon'' + \text{etc.} \\ y^0 &= \beta\epsilon + \beta'\epsilon' + \beta''\epsilon'' + \text{etc.} \\ z^0 &= \gamma\epsilon + \gamma'\epsilon' + \gamma''\epsilon'' + \text{etc.} \end{aligned}$$

Denique statuemus

$$\Omega^0 = \epsilon\epsilon + \epsilon'\epsilon' + \epsilon''\epsilon'' + \text{etc.}$$

ita ut sit $\Omega^0$ aequalis valori functionis $\Omega$, valoribus veris ipsarum $x$, $y$, $z$, etc. respondenti. Hinc quum habeatur indefinite

$$\Omega = M + (x-A)\xi + (y-B)\eta + (z-C)\zeta + \text{etc.}$$

erit etiam

$$M = \Omega^0 - x^0\xi^0 - y^0\eta^0 - z^0\zeta^0 - \text{etc.}$$

Hinc manifestum est, $M$, evolutione facta esse functionem homogeneam secundi ordinis errorum $e$, $e'$, $e''$, etc., quae, pro diversis errorum valoribus maior minorve evadere poterit. Sed quantenus errorum magnitudo nobis incognita manet, unknowns coincide with the true values it is certainly greater than $M$ (which is the smallest possible value). In general we can say that the usual practice produces a mean error that is always too small and assigns the observations too large a weight. Let us now see what rigorous theory shows.

38.

First of all we must determine how $M$ depends on the true observation errors. As in Art. 28, we will denote these errors by $e$, $e'$, $e''$, etc., and for simplicity we will set

$$e\sqrt{p} = \epsilon, \quad e'\sqrt{p'} = \epsilon', \quad e''\sqrt{p''} = \epsilon'', \text{ etc., and}$$
$$m\sqrt{p} = m'\sqrt{p'} = m''\sqrt{p''} = \text{etc.} = \mu.$$

In addition let the true values of $x$, $y$, $z$, etc. be $A - x^0$, $B - y^0$, $C - z^0$, etc., corresponding to the values $-\xi^0$, $-\eta^0$, $-\zeta^0$, etc. of $\xi$, $\eta$, $\zeta$, etc. Obviously these values correspond to the values $-\epsilon$, $-\epsilon'$, $-\epsilon''$, etc. of $v$, $v'$, $v''$, etc., so that we have

$$\xi^0 = a\epsilon + a'\epsilon' + a''\epsilon'' + \text{etc.}$$
$$\eta^0 = b\epsilon + b'\epsilon' + b''\epsilon'' + \text{etc.}$$
$$\zeta^0 = c\epsilon + c'\epsilon' + c''\epsilon'' + \text{etc.}$$

etc. and

$$x^0 = \alpha\epsilon + \alpha'\epsilon' + \alpha''\epsilon'' + \text{etc.}$$
$$y^0 = \beta\epsilon + \beta'\epsilon' + \beta''\epsilon'' + \text{etc.}$$
$$z^0 = \gamma\epsilon + \gamma'\epsilon' + \gamma''\epsilon'' + \text{etc.}$$

Finally let

$$\Omega^0 = \epsilon\epsilon + \epsilon'\epsilon' + \epsilon''\epsilon'' + \text{etc.},$$

so that $\Omega^0$ is equal to the value of the function $\Omega$ corresponding to the true values of $x$, $y$, $z$, etc.

Since in general

$$\Omega = M + (x - A)\xi + (y - B)\eta + (z - C)\zeta + \text{etc.}$$

we have

$$M = \Omega^0 - x^0\xi^0 - y^0\eta^0 - z^0\zeta^0 - \text{etc.}$$

From this it is obvious on expansion that $M$ is a homogeneous function of the second order in $e$, $e'$, $e''$, etc., which will vary depending on the values of the errors. Since we cannot know the sizes of the errors, it is reasonable to regard $M$

functionem hanc indefinite considerare, imprimisque secundum principia calculi probabilitatis eius valorem medium assignare conveniet. Quem inveniemus, si loco quadratorum $ee$, $e'e'$, $e''e''$, etc. resp. scribimus $mm$, $m'm'$, $m''m''$, etc., producta vero $ee'$, $ee''$, $e'e''$, etc. omnino omittimus, vel quod idem est, si loco cuiusvis quadrati $\epsilon\epsilon$, $\epsilon'\epsilon'$, $\epsilon''\epsilon''$, etc. scribimus $\mu\mu$, productis $\epsilon\epsilon'$, $\epsilon\epsilon''$, $\epsilon'\epsilon''$, etc. prorsus neglectis. Hoc modo e termino $\Omega^0$ manifesto provenit $\pi\mu\mu$; terminus $-x^0\xi^0$ producet

$$-(a\alpha + a'\alpha' - a''\alpha'' - \text{etc.})\mu\mu = -\mu\mu$$

et similiter singulae partes reliquae praebebunt $-\mu\mu$, ita ut valor medius fiat $= (\pi - \rho)\mu\mu$, denotante $\pi$ multitudinem observationum, $\rho$ multitudinem incognitarum. Valor verus quidem ipsius $M$, prout fors errores obtulit, maior minorve medio fieri potest, set discrepantia eo minoris momenti erit, quo maior fuerit observationum multitudo, ita ut pro valore approximato ipsius $\mu$ accipere liceat

$$\sqrt{\tfrac{M}{\pi-\rho}}$$

Valor itaque ipsius $\mu$, ex praxi erronea, de qua in art. praec. loquuti sumus, prodiens, augeri debet in ratione quantitatis $\sqrt{\pi - \rho}$ ad $\sqrt{\pi}$.

## 39.

Quo clarius eluceat, quanto iure valorem fortitum ipsius $M$ medio aequiparare liceat, adhuc investigare oportet errorem medium metuendum, dum statuimus $\frac{M}{\pi-\rho} = \mu\mu$. Iste error medius aequalis est radici quadratae e valore medio quantitatis

$$\left(\tfrac{\Omega^0 - x^0\xi^0 - y^0\eta^0 - z^0\zeta^0 - \text{etc.} - (\pi-\rho)\mu\mu}{\pi-\rho}\right)^2$$

quam ita exhibebimus

$$\left(\tfrac{\Omega^0 - x^0\xi^0 - y^0\eta^0 - z^0\zeta^0 - \text{etc.}}{\pi-\rho}\right)^2 - \tfrac{2\mu\mu}{\pi-\rho}(\Omega^0 - x^0\xi^0 - y^0\eta^0 - z^0\zeta^0 - \text{etc.} - (\pi-\rho)\ \mu\mu) - \mu^4$$

et quum manifesto valor medius termini secundi fiat $= 0$, res in eo vertitur, ut indagemus valorem medium functionis

$$\Psi = (\Omega^0 - x^0\xi^0 - y^0\eta^0 - z^0\zeta^0 - \text{etc.})^2$$

quo invento et per $N$ designato, error medius quaesitus erit

$$\sqrt{\tfrac{N}{(\pi-\rho)^2} - \mu^4}$$

as a general function of the errors and determine its mean value by the calculus of probabilities. If we write $mm$, $m'm'$, $m''m''$, etc. for the squares $ee$, $e'e'$, $e''e''$, etc. of the errors and omit the products $ee'$, $ee''$, $e'e''$, etc., the result will be the mean value. This is the same as writing $\mu\mu$ for any of the squares $\epsilon\epsilon$, $\epsilon'\epsilon'$, $\epsilon''\epsilon''$, etc. and neglecting the products $\epsilon\epsilon'$, $\epsilon\epsilon''$, $\epsilon'\epsilon''$, etc. In this way the term $\Omega^0$ gives $\pi\mu\mu$. The term $-x^0\xi^0$ gives

$$-(a\alpha + a'\alpha' - a''\alpha'' - \text{etc.})\mu\mu = -\mu\mu,$$

and similarly each of the remaining terms yields $-\mu\mu$. Hence the total mean value is $(\pi - \rho)\mu\mu$, where $\pi$ is the number of observations and $\rho$ is the number of unknowns.

Depending on how the errors fall out, the true value of $M$ may be larger or smaller than its mean; but the greater the number of observations, the less important the discrepancy, so that we may take

$$\sqrt{\tfrac{M}{\pi-\rho}}$$

as an approximation to $\mu$. Thus the value from the erroneous practice discussed in the preceding article should be increased as the ratio of $\sqrt{\pi}$ to $\sqrt{\pi - \rho}$.

39.

To see how well the random value of $M$ compares with its mean, we must now determine the error to be feared in setting $\frac{M}{\pi-\rho} = \mu\mu$. This mean error is equal to the square root of the mean value of

$$\left(\tfrac{\Omega^0 - x^0\xi^0 - y^0\eta^0 - z^0\zeta^0 - \text{etc.} - (\pi-\rho)\mu\mu}{\pi-\rho}\right)^2,$$

which can be written

$$\left(\tfrac{\Omega^0 - x^0\xi^0 - y^0\eta^0 - z^0\zeta^0 - \text{etc.}}{\pi-\rho}\right)^2 - \tfrac{2\mu\mu}{\pi-\rho}(\Omega^0 - x^0\xi^0 - y^0\eta^0 - z^0\zeta^0 - \text{etc.} - (\pi-\rho)\mu\mu) - \mu^4.$$

Since the mean value of the second term is obviously zero, the problem becomes one of working out the mean value of the function

$$\Psi = (\Omega^0 - x^0\xi^0 - y^0\eta^0 - z^0\zeta^0 - \text{etc.})^2.$$

If this mean value is denoted by $N$, then our mean error will be

$$\sqrt{\tfrac{N}{(\pi-\rho)^2} - \mu^4}.$$

Expressio $\Psi$ evoluta manifest est functio homogenea sive errorum $e$, $e'$, $e''$, etc., sive quantitatum $\epsilon$, $\epsilon'$, $\epsilon''$, etc., eiusque valor medius invenietur, si

1° pro biquadratis $e^4$, $e'^4$, $e''^4$, etc. substituuntur eorum valores medii

2° pro singulis productis e binis quadratis ut $eee'e'$, $eee''e''$, $e'e'e''e''$, etc. producta ex ipsorum valoribus mediis, puta $mmm'm'$, $mmm''m''$, $m'm'm''m''$, etc.

3° partes vero reliquae, quae implicabunt vel factorem talem $e^3e'$ , vel talem $eee'e''$, omnino omittuntur. Valores medios biquadratorum $e^4$, $e'^4$, $e''^4$, etc. ipsis biquadratis $m^4$, $m'^4$, $m''^4$, etc. proportionales supponemus (vid. art. 16), ita ut illi sint ad heac ut $\nu^4$ ad $\mu^4$, adeoque $\nu^4$ denotet valorem medium biquadratorum observationum talium quarum pondus $= 1$. Hinc pracepta praecedentia ita quoque exprimi poterunt: Loco singulorum biquadratorum $\epsilon^4$, $\epsilon'^4$, $\epsilon''^4$, etc. scribendum erit $\nu^4$, loco singulorum productorum e binis quadratis ut $\epsilon\epsilon\epsilon'\epsilon'$, $\epsilon\epsilon\epsilon''\epsilon''$, $\epsilon'\epsilon'\epsilon''\epsilon''$, etc., scribendum erit $\mu^4$, omnesque reliqui termini, qui implicabunt factores tales ut $\epsilon^3\epsilon'$, $\epsilon\epsilon\epsilon'\epsilon''$, vel $\epsilon\epsilon'\epsilon''\epsilon'''$ erunt supprimendi.

His probe intellectis facile patebit.

I. Valorem medium quadrati $\Omega^0\Omega^0$ esse $\pi\nu^4 + (\pi\pi - \pi)\mu^4$

II. Valor medius producti $\epsilon\epsilon x^0\xi^0$ fit $= a\alpha\nu^4 + (a'\alpha' + a''\alpha'' + \text{etc.})\mu^4$, sive quoniam $a\alpha + a'\alpha' + a''\alpha'' + \text{etc.} = 1$,

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

Et quum perinde valor medius producti $\epsilon'\epsilon'x^0\xi^0$ fiat $= a'\alpha'(\nu^4 - \mu^4) + \mu^4$, valor medius producti $\epsilon''\epsilon''x^0\xi^0$ autem $= a''\alpha''(\nu^4 - \mu^4) + \mu^4$ et sic porro, patet, valorem medium producti $(\epsilon\epsilon + \epsilon'\epsilon' + \epsilon''\epsilon'' + \text{etc.})x^0\xi^0$ sivi $\Omega^0x^0\xi^0$ esse

$$= \nu^4 - \mu^4 + \pi\mu^4$$

Eundem valorem medium habebunt producta $\Omega^0y^0\eta^0$, $\Omega^0z^0\zeta^0$, etc. Quapropter valor medius producti $\Omega^0(x^0\xi^0 + y^0\eta^0 + z^0\zeta^0 + \text{etc.})$ fit

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

III. Ne evolutiones reliquae nimis prolixae evadant, idonea denotatio introducenda erit. Utemur itaque characteristica $\Sigma$ sensu aliquantum latiore quam supra passim factum est, ita ut denotet aggregatum termini, cui praefixa est, cum omnibus similibus sed non identicis inde per omnes observationum permutationes

When it is expanded, $\Psi$ is obviously a homogeneous function of the errors $e$, $e'$, $e''$, etc. or the quantities $\epsilon$, $\epsilon'$, $\epsilon''$, etc. Its mean value may be found by

1° replacing the quartics $e^4$, $e'^4$, $e''^4$, etc. by their mean values;

2° replacing the products of two quadratics like $eee'e'$, $eee''e''$, $e'e'e''e''$, etc. by the products of their mean values, namely, $mmm'm'$, $mmm''m''$, $m'm'm''m''$, etc.;

3° dropping all the remaining terms, which are associated with terms like $e^3e'$ or $eee'e''$.

We will assume that the mean values of the quartics $e^4$, $e'^4$, $e''^4$, etc. are proportional to the quartics $m^4$, $m'^4$, $m''^4$, etc. (see Art. 16) so that they are to the latter as $\nu^4$ is to $\mu^4$, where $\nu^4$ denotes the mean value of the fourth power of a observation whose weight is one. Hence the above rules may be expressed as follows: write $\nu^4$ for each quartic $\epsilon^4$, $\epsilon'^4$, $\epsilon''^4$, etc.; write $\mu^4$ for each product of two quadratics like $\epsilon\epsilon\epsilon'\epsilon'$, $\epsilon\epsilon\epsilon''\epsilon''$, $\epsilon'\epsilon'\epsilon''\epsilon''$, etc.; ignore all remaining terms, which are associated with factors like $\epsilon^3\epsilon'$, $\epsilon\epsilon\epsilon'\epsilon''$, or $\epsilon\epsilon'\epsilon''\epsilon'''$.

From these rules it is obvious that

I. The mean value of $\Omega^0\Omega^0$ is $\pi\nu^4 + (\pi\pi - \pi)\mu^4$.

II. The mean value of the product $\epsilon\epsilon x^0\xi^0$ is $a\alpha\nu^4 + (a'\alpha' + a''\alpha'' + \text{etc.})\mu^4$, or

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

since $a\alpha + a'\alpha' + a''\alpha'' + \text{etc.} = 1$. Similarly the mean value of the product $\epsilon'\epsilon' x^0\xi^0$ is $a'\alpha'(\nu^4 - \mu^4) + \mu^4$, the mean value of $\epsilon''\epsilon'' x^0\xi^0$ is $a''\alpha''(\nu^4 - \mu^4) + \mu^4$, and so on. Hence the mean error of the product $(\epsilon\epsilon + \epsilon'\epsilon' + \epsilon''\epsilon'' + \text{etc.})x^0\xi^0$ or $\Omega^0 x^0\xi^0$ is

$$\nu^4 - \mu^4 + \pi\mu^4.$$

The products $\Omega^0 y^0\eta^0$, $\Omega^0 z^0\zeta^0$, etc. have the same value. Hence the mean value of the product $\Omega^0(x^0\xi^0 + y^0\eta^0 + z^0\zeta^0 + \text{etc.})$ is

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

III. To keep the remaining expansions from becoming too complicated, we must introduce appropriate notation. Specifically we will use the symbol $\Sigma$ in a somewhat wider sense than previously to denote the sum of the term to which it is prefixed with all similar but not identical terms arising from all permutations of the observations. For example, with this convention we have $x^0 = \Sigma\alpha\epsilon$ and $x^0x^0 = \Sigma\alpha\alpha\epsilon\epsilon + 2\Sigma\alpha\alpha'\epsilon\epsilon'$.

oriundis. Hoc pacto e.g. habemus $x^0 = \Sigma\alpha\epsilon$, $x^0x^0 = \Sigma\alpha\alpha\epsilon\epsilon + 2\Sigma\alpha\alpha'\epsilon\epsilon'$. Coligendo itaque valorem medium producti $x^0x^0\xi^0\xi^0$ per partes, habemus primo valorem medium producti $\alpha\alpha\epsilon\epsilon\xi^0\xi^0$

$$\begin{aligned}&= aa\alpha\alpha\nu^4 + \alpha\alpha(a'a' + a''a'' + \text{etc.})\mu^4\\ &= aa\alpha\alpha(\nu^4 - \mu^4) + \alpha\alpha\mu^4\Sigma aa\end{aligned}$$

Perinde valor medius producti $\alpha'\alpha'\epsilon'\epsilon'\xi^0\xi^0$ fit $= a'a'\alpha'\alpha'(\nu^4 - \mu^4) + \alpha'\alpha'\mu^4\Sigma aa$, et sic porro, adeoque valor medius producti $\xi^0\xi^0\Sigma\alpha\alpha\epsilon\epsilon$

$$= (\nu^4 - \mu^4)\Sigma aa\alpha\alpha + \mu^4\Sigma aa.\Sigma\alpha\alpha$$

Porro valor medius producti $\alpha\alpha'\epsilon\epsilon'\xi^0\xi^0$ fit $2\alpha\alpha'aa'\mu^4$, valor medius producti $\alpha\alpha'' \times \epsilon\epsilon''\xi^0\xi^0$ perinde $= 2\alpha\alpha''aa''\mu^4$, etc., unde facile concluditur, valorem medium producti $\xi^0\xi^0\Sigma\alpha\alpha'\epsilon\epsilon'$ fieri

$$= 2\mu^4\Sigma\alpha a\alpha'a' = \mu^4\left((\Sigma a\alpha)^2 - \Sigma aa\alpha\alpha\right) = \mu^4(1 - \Sigma aa\alpha\alpha)$$

His collectis habemus valorem medium producti $x^0x^0\xi^0\xi^0$

$$= (\nu^4 - 3\mu^4)\Sigma aa\alpha\alpha + 2\mu^4 + \mu^4\Sigma aa.\Sigma\alpha\alpha$$

IV. Haud absimili modo invenitur valor medius producti $x^0y^0\xi^0\eta^0$

$$= \nu^4\Sigma ab\alpha\beta + \mu^4\Sigma a\alpha b'\beta' + \mu^4\Sigma ab\alpha'\beta' + \mu^4\Sigma a\beta b'\alpha'$$

Sed habetur

$$\begin{aligned}\Sigma a\alpha b'\beta' &= \Sigma a\alpha.\Sigma b\beta - \Sigma a\alpha b\beta,\\ \Sigma ab\alpha'\beta' &= \Sigma ab.\Sigma\alpha\beta - \Sigma ab\alpha\beta,\\ \Sigma a\beta b'\alpha' &= \Sigma a\beta.\Sigma b\alpha - \Sigma a\beta b\alpha,\end{aligned}$$

unde valor ille medius fit, propter $\Sigma a\alpha = 1$, $\Sigma b\beta = 1$, $\Sigma a\beta = 0$, $\Sigma b\alpha = 0$,

$$= (\nu^4 - 3\mu^4)\Sigma ab\alpha\beta + \mu^4(1 + \Sigma ab.\Sigma\alpha\beta)$$

V. Quum prorsus eodem modo valor medius producti $x^0z^0\xi^0\zeta^0$ fiat

$$= (\nu^4 - 3\mu^4)\Sigma ac\alpha\gamma + \mu^4(1 + \Sigma ac.\Sigma\alpha\gamma)$$

et sic porro, additio valorem medium producti $x^0\xi^0(x^0\xi^0 + y^0\eta^0 + z^0\zeta^0 + \text{etc.})$ suppeditat

$$\begin{aligned}&= (\nu^4 - 3\mu^4)\Sigma(a\alpha(a\alpha + b\beta + c\gamma + \text{etc.})) + (\rho + 1)\mu^4\\ &\qquad + \mu^4(\Sigma aa.\Sigma\alpha\alpha + \Sigma ab.\Sigma\alpha\beta + \Sigma ac.\Sigma\alpha\gamma + \text{etc.})\\ &= (\nu^4 - 3\mu^4)\Sigma(a\alpha(a\alpha + b\beta + c\gamma + \text{etc.})) + (\rho + 2)\mu^4\end{aligned}$$

We will now assemble pieces of the mean value of the product $x^0x^0\xi^0\xi^0$. First, the mean value of the product $\alpha\alpha\epsilon\epsilon\xi^0\xi^0$ is

$$\begin{aligned} &aa\alpha\alpha\nu^4 + \alpha\alpha(a'a' + a''a'' + \text{etc.})\mu^4 \\ &= aa\alpha\alpha(\nu^4 - \mu^4) + \alpha\alpha\mu^4\Sigma aa. \end{aligned}$$

Similarly the mean value of the product $\alpha'\alpha'\epsilon'\epsilon'\xi^0\xi^0$ is $a'a'\alpha'\alpha'(\nu^4-\mu^4)+\alpha'\alpha'\mu^4\Sigma aa$ and so on. Thus the mean value of the product $\xi^0\xi^0\Sigma\alpha\alpha\epsilon\epsilon$ is

$$(\nu^4 - \mu^4)\Sigma aa\alpha\alpha + \mu^4\Sigma aa.\Sigma\alpha\alpha.$$

Moreover, the mean value of the product $\alpha\alpha'\epsilon\epsilon'\xi^0\xi^0$ is $2\alpha\alpha'aa'\mu^4$, the mean value of the product $\alpha\alpha''\epsilon\epsilon''\xi^0\xi^0$ is likewise $2\alpha\alpha''aa''\mu^4$, etc. From this we easily conclude that the mean value of the product $\xi^0\xi^0\Sigma\alpha\alpha'\epsilon\epsilon'$ is

$$2\mu^4\Sigma\alpha a\alpha' a' = \mu^4\left((\Sigma a\alpha)^2 - \Sigma aa\alpha\alpha\right) = \mu^4(1 - \Sigma aa\alpha\alpha).$$

Collecting these pieces, we find that the mean value of the product $x^0x^0\xi^0\xi^0$ is

$$(\nu^4 - 3\mu^4)\Sigma aa\alpha\alpha + 2\mu^4 + \mu^4\Sigma aa.\Sigma\alpha\alpha.$$

IV. In much the same way, we find that the mean value of the product $x^0y^0\xi^0\eta^0$ is

$$\nu^4\Sigma ab\alpha\beta + \mu^4\Sigma a\alpha b'\beta' + \mu^4\Sigma ab\alpha'\beta' + \mu^4\Sigma a\beta b'\alpha'.$$

But

$$\begin{aligned} \Sigma a\alpha b'\beta' &= \Sigma a\alpha.\Sigma b\beta - \Sigma a\alpha b\beta, \\ \Sigma ab\alpha'\beta' &= \Sigma ab.\Sigma\alpha\beta - \Sigma ab\alpha\beta, \\ \Sigma a\beta b'\alpha' &= \Sigma a\beta.\Sigma b\alpha - \Sigma a\beta b\alpha, \end{aligned}$$

from which we get the mean value

$$(\nu^4 - 3\mu^4)\Sigma ab\alpha\beta + \mu^4(1 + \Sigma ab.\Sigma\alpha\beta),$$

since $\Sigma a\alpha = 1$, $\Sigma b\beta = 1$, $\Sigma a\beta = 0$, $\Sigma b\alpha = 0$.

V. In the same way the mean value of the product $x^0z^0\xi^0\zeta^0$ is

$$(\nu^4 - 3\mu^4)\Sigma ac\alpha\gamma + \mu^4(1 + \Sigma ac.\Sigma\alpha\gamma),$$

and so on. Their sum is the mean of the product $x^0\xi^0(x^0\xi^0 + y^0\eta^0 + z^0\zeta^0 + \text{etc.})$: namely,

$$\begin{aligned} &(\nu^4 - 3\mu^4)\Sigma(a\alpha(a\alpha + b\beta + c\gamma + \text{etc.})) + (\rho + 1)\mu^4 \\ &+\mu^4(\Sigma aa.\Sigma\alpha\alpha + \Sigma ab.\Sigma\alpha\beta + \Sigma ac.\Sigma\alpha\gamma + \text{etc.}) \\ &= (\nu^4 - 3\mu^4)\Sigma(a\alpha(a\alpha + b\beta + c\gamma + \text{etc.})) + (\rho + 2)\mu^4. \end{aligned}$$

VI. Prorsus eodem modo valor medius producti $y^0\eta^0(x^0\xi^0+y^0\eta^0+z^0\zeta^0+\text{etc.})$ eruitur

$$= (\nu^4-3\mu^4)\Sigma(b\beta(a\alpha+b\beta+c\gamma+\text{etc.}))+(\rho+2)\mu^4$$

dein valor medius producti $z^0\zeta^0(x^0\xi^0+y^0\eta^0+z^0\zeta^0+\text{etc.})$

$$= (\nu^4-3\mu^4)\Sigma(c\gamma(a\alpha+b\beta+c\gamma+\text{etc.}))+(\rho+2)\mu^4$$

et sic porro. Hinc per additionem prodit valor medius quadrati $(x^0\xi^0+y^0\eta^0+z^0\zeta^0+\text{etc.})^2$

$$= (\nu^4-3\mu^4)\Sigma((a\alpha+b\beta+c\gamma+\text{etc.})^2)+(\rho\rho+2\rho)\mu^4$$

VII. Omnibus tandem rite collectis eruitur

$$\begin{aligned} N &= (\pi-2\rho)\nu^4+(\pi\pi-\pi-2\pi\rho+4\rho+\rho\rho)\mu^4 \\ &\quad +(\nu^4-3\mu^4)\Sigma((a\alpha+b\beta+c\gamma+\text{etc.})^2) \\ &= (\pi-\rho)(\nu^4\rho^4)+(\pi-\rho)^2\mu^4-(\nu^4-3\mu^4)[\rho-\Sigma((a\alpha+b\beta+c\gamma+\text{etc.})^2)] \end{aligned}$$

Error itaque medius in determinatione ipsius $\mu\mu$ per formulam

$$\mu\mu = \tfrac{M}{\pi-\rho}$$

metuendus erit

$$= \sqrt{\tfrac{\nu^4-\mu^4}{\pi-\rho} - \tfrac{\nu^4-3\mu^4}{(\pi-\rho)^2}.[\rho-\Sigma((a\alpha+b\beta+c\gamma+\text{etc.})^2)]}$$

## 40.

Quantitas $\Sigma((a\alpha+b\beta+c\gamma+\text{etc.})^2)$, quae in expressionem modo inventam ingreditur, generaliter quidem ad formam simplicorem reduci nequit: nihilominus duo limites assignari possunt, inter quos ipsius valor necessario iacere debet. *Primo* scilicet e relationibus supra evolutis facile demonstratur, esse

$$\begin{gathered}(a\alpha+b\beta+c\gamma+\text{etc.})^2+(a\alpha'+b\beta'+c\gamma'+\text{etc.})^2+(a\alpha''+b\beta''+c\gamma''+\text{etc.})^2 \\ +\ \text{etc.} = a\alpha+b\beta+c\gamma+\text{etc.}\end{gathered}$$

unde concludimus, $a\alpha+b\beta+c\gamma+\text{etc.}$ esse quantitatem positivam unitate minorem (saltem non maiorem). Idem valet de quantatitate $a'\alpha'+b'\beta'+c'\gamma'+\text{etc.}$, quippe cui aggregatum

$$(a'\alpha+b'\beta+c'\gamma+\text{etc.})^2+(a'\alpha'+b'\beta'+c'\gamma'+\text{etc.})^2+(a'\alpha''+b'\beta''+c'\gamma''+\text{etc.})^2+\text{etc.}$$

VI. In the same way the mean of the product $y^0\eta^0(x^0\xi^0 + y^0\eta^0 + z^0\zeta^0 + \text{etc.})$ is found to be

$$(\nu^4 - 3\mu^4)\Sigma(b\beta(a\alpha + b\beta + c\gamma + \text{etc.})) + (\rho + 2)\mu^4,$$

and the mean of the product $z^0\zeta^0(x^0\xi^0 + y^0\eta^0 + z^0\zeta^0 + \text{etc.})$ is

$$(\nu^4 - 3\mu^4)\Sigma(c\gamma(a\alpha + b\beta + c\gamma + \text{etc.})) + (\rho + 2)\mu^4,$$

and so on. Hence by summing we get the mean of the square $(x^0\xi^0 + y^0\eta^0 + z^0\zeta^0 + \text{etc.})^2$:

$$(\nu^4 - 3\mu^4)\Sigma((a\alpha + b\beta + c\gamma + \text{etc.})^2) + (\rho\rho + 2\rho)\mu^4.$$

VII. Finally, bringing everything together, we find that

$$\begin{aligned} N &= (\pi - 2\rho)\nu^4 + (\pi\pi - \pi - 2\pi\rho + 4\rho + \rho\rho)\mu^4 \\ &\quad + (\nu^4 - 3\mu^4)\Sigma((a\alpha + b\beta + c\gamma + \text{etc.})^2) \\ &= (\pi - \rho)(\nu^4\rho^4) + (\pi - \rho)^2\mu^4 - (\nu^4 - 3\mu^4)[\rho - \Sigma((a\alpha + b\beta + c\gamma + \text{etc.})^2)]. \end{aligned}$$

Hence the mean error to be feared in estimating $\mu\mu$ by the formula

$$\mu\mu = \tfrac{M}{\pi - \rho}$$

is

$$\sqrt{\tfrac{\nu^4 - \mu^4}{\pi - \rho} - \tfrac{\nu^4 - 3\mu^4}{(\pi - \rho)^2}.[\rho - \Sigma((a\alpha + b\beta + c\gamma + \text{etc.})^2)]}.$$

40.

The quantity $\Sigma((a\alpha + b\beta + c\gamma + \text{etc.})^2)$ in the expression we just derived cannot in general be simplified. Nonetheless, we can establish two limits within which it must lie. First, it is easy to show from the above relations that

$$\begin{gathered} (a\alpha + b\beta + c\gamma + \text{etc.})^2 + (a\alpha' + b\beta' + c\gamma' + \text{etc.})^2 \ + \ (a\alpha'' + b\beta'' + c\gamma'' + \text{etc.})^2 \\ + \ \text{etc.} = a\alpha + b\beta + c\gamma + \text{etc.}, \end{gathered}$$

from which we see that $a\alpha + b\beta + c\gamma + \text{etc.}$ is a positive number less than or equal to one. The same is true of $a'\alpha' + b'\beta' + c'\gamma' + \text{etc.}$, which is equal to

$$(a'\alpha + b'\beta + c'\gamma + \text{etc.})^2 + (a'\alpha' + b'\beta' + c'\gamma' + \text{etc.})^2 + (a'\alpha'' + b'\beta'' + c'\gamma'' + \text{etc.})^2 + \text{etc.}$$

aequale invenitur; ac perinde $a''\alpha'' + b''\beta'' + c''\gamma'' +$ etc. unitate minor erit, et sic porro. Hinc $\Sigma((a\alpha + b\beta + c\gamma + \text{etc.})^2)$ necessario est minor quam $\pi$. *Secundo* habetur $\Sigma(a\alpha + b\beta + c\gamma + \text{etc.}) = \rho$, quoniam fit $\Sigma a\alpha = 1$, $\Sigma b\beta = 1$, $\Sigma c\gamma = 1$, etc.; unde facile deducitur, summam quadratorum $\Sigma((a\alpha + b\beta + c\gamma + \text{etc.})^2)$ esse maiorem quam $\frac{\rho\rho}{\pi}$, vel saltem non minorem. Hinc terminus

$$\tfrac{\nu^4-3\mu^4}{(\pi-\rho)^2}.[\rho - \Sigma((a\alpha + b\beta + c\gamma + \text{etc.})^2)]$$

necessario iacet inter limites $-\frac{\nu^4-3\mu^4}{\pi-\rho}$ et $\frac{\nu^4-3\mu^4}{\pi-\rho}.\frac{\rho}{\pi}$ vel, si latiores praeferimus, inter hos $-\frac{\nu^4-3\mu^4}{\pi-\rho}$ et $+\frac{\mu^4-3\nu^4}{\pi-\rho}$, et proin erroris medii in valore ipsius $\mu\mu = \frac{M}{\pi-\rho}$ metuendi qudratum inter limites $\frac{2\nu^4-4\mu^4}{\pi-\rho}$ et $\frac{2\mu^4}{\pi-\rho}$, ita ut praecisionem quantamvis assequi liceat, si modo observationum multitudo fuerit satis magna.

Valde memoribile est, in hypothesi ea (art. 9, III), cui theoria quadratorum minimorum olim superstructa fuerat, illum terminum omnino excidere, et sicuti, ad cruendum valorem approximatum erroris medii observationum $\mu$, in ombibus casibus aggregatum $\lambda\lambda + \lambda'\lambda' + \lambda''\lambda'' +$ etc. $= M$ ita tractare oportet, ac si esset aggregatum $\pi - \rho$ errorum fortuitorum, ita in illa hypothesi etiam praecisionem ipsam huius determinationis aequalem fieri ei, quam determinationi ex $\pi - \rho$ erroribus veris tribuendam esse in art. 15 invenimus.

Similarly, $a''\alpha'' + b''\beta'' + c''\gamma'' + \text{etc.}$ is less than one, and so on. Hence $\Sigma((a\alpha + b\beta + c\gamma + \text{etc.})^2)$ must be less than $\pi$.

Second, since $\Sigma a\alpha = 1$, $\Sigma b\beta = 1$, $\Sigma c\gamma = 1$, etc., we have $\Sigma(a\alpha{+}b\beta{+}c\gamma{+}\text{etc.}) = \rho$. From this it is easy to see that the sum of squares $\Sigma((a\alpha + b\beta + c\gamma + \text{etc.})^2)$ is greater than or equal to $\frac{\rho\rho}{\pi}$.

Thus the term

$$\tfrac{\nu^4-3\mu^4}{(\pi-\rho)^2}.[\rho - \Sigma((a\alpha + b\beta + c\gamma + \text{etc.})^2)]$$

must lie between the limits $-\frac{\nu^4-3\mu^4}{\pi-\rho}$ and $\frac{\nu^4-3\mu^4}{\pi-\rho}.\frac{\rho}{\pi}$, or between $-\frac{\nu^4-3\mu^4}{\pi-\rho}$ and $+\frac{\mu^4-3\nu^4}{\pi-\rho}$ if we prefer looser bounds. Hence the square of the mean error to be feared in $\mu\mu = \frac{M}{\pi-\rho}$ lies between $\frac{2\nu^4-4\mu^4}{\pi-\rho}$ and $\frac{2\mu^4}{\pi-\rho}$, so that arbitrarily high precision can be ascribed to the estimate, provided that the number of observations is sufficiently large.

Under the hypothesis on which the theory of least squares was originally based (Art. 9, III) the term just treated vanishes. In this case the precision of our estimate is equal to the precision of an estimate from $\pi - \rho$ true errors, determined as in Art. 15. This is in striking analogy with the treatment of the sum of squares $\lambda\lambda + \lambda'\lambda' + \lambda''\lambda'' + \text{etc.} = M$ to estimate $\mu$ as if it were the sum of $\pi - \rho$ random errors.

*This page intentionally left blank*

# Theoria
# Combinationis Observationum
# Erroribus Minimis Obnoxiae

*Supplementum*

# Theory of the
# Combination of Observations
# Least Subject to Errors

*Supplement*

# Theoriae Combinationis Observationum Erroribus Minimis Obnoxiae

## Supplementum

1.

In tractatione theoriae combinationis observationum Volumini V Commentionum Recentiorum inserta supposuimus, quantatates eas, quarum valores per observationes praecisione absoluta non gaudentes propositi sunt, a certis elementis incognitis ita pendere, ut in forma functionum datarum horum elementorum exhibitae sint, reique cardinem in eo verti, ut haec elementa quam exactissime ex observationibus deriventur.

In plerisque quidem casibus suppositio ista immediate locum habet. In aliis vero casibus problematis conditio paullo aliter se offert, ita ut primo aspectu dubium videatur, quonam pacto ad formam requisitam reduci posset. Haud raro scilicet accidit, ut quantitates eae, ad quas referuntur observationes, nondum exhibitae sint in forma functionum certorum elementorum, neque etiam ad talem formam reducibles videantur, saltem non commode vel sine ambagibus: dum, ex altera parte, rei indoles quasdam conditiones suppeditat, quibus valores veri quantitatum observatarum exacte et necessario satisfacere debent.

Attamen, re propius considerata, facile perspicitur, hunc casum ab altero revera essentialiter haud differre, sed ad eundem reduci posse. Designando scilicet multitudinem quantitatum observatarum per $\pi$, multitudinem aequationum conditionalium autem per $\sigma$, eligendoque e prioribus $\pi - \sigma$ ad lubitum, nihil impedit, quominus has ipsas pro elementis accipiamus, reliquasque, quarum multitudo erit $\sigma$, adiumento aequationum conditionalium tamquam functiones illarum consideremus, quo pacto res ad suppositionem nostram reducta erit.

Verum enim vero etiamsi haec via in permultis casibus satis commode ad finem propositum perducat, tamen negari non potest, eam minus genuinam, operaeque adeo pretium esse, problema in ista altera forma seorsim tractare, tantoque magis, quod solutionem perelegantem admittit. Quin adeo, quum haec solutio nova ad calculos expeditiores perducat, quam solutio problematis in statu priore, quoties $\sigma$ est minor quam $\frac{1}{2}\pi$, sive quod idem est, quoties multitudo elementorum in commentatione priore per $\rho$ denotata maior est, quam $\frac{1}{2}\pi$, solutionem novam,

# Theory of the Combination of Observations Least Subject to Errors

## Supplement

1.

In a treatment of combinations of observations, which appeared in Vol. V of *Commentationes Recentiores*, it was assumed that certain quantities, whose values were given by imprecise observations, were functions of certain elements. The problem was to derive the elements from the observations as precisely as possible.

In very many cases the above assumption holds directly. In other cases, however, the problem is posed differently, and at first sight it is not obvious how to reduce it to the required form. In particular, it sometimes happens that the observed quantities are not represented as functions of given elements and do not seem to be reducible to such a form — at least not in a natural, straightforward way; yet at the same time the problem itself provides conditions which must be satisfied exactly by the true values of the observed quantities.

Now from a proper vantage it is easy to see that this case is not really different from the first but can be reduced to it. Specifically, let $\pi$ denote the number of observations and $\sigma$ the number of conditional equations. Then we may choose any $\pi - \sigma$ of the observations and regard them as elements. With the help of the conditional equations, the remaining $\sigma$ observations may be taken to be functions of these unknowns. This reduces the problem to one satisfying our assumption.

In many cases, this approach leads to a satisfactory solution of the problem. However, the approach is certainly not natural, and it is worthwhile to treat the problem on its own terms, especially since it has an elegant solution. Moreover, the new solution is easier to compute than the one just given, provided $\sigma$ is less than $\frac{1}{2}\pi$; that is, when the number of elements, denoted by $\rho$ in the preceding quam in commentatione praesenti explicabimus, in tali casu praeferre conveniet priori, siquidem aequationes conditionales e problematis indole absque ambagibus depromere licet.

2.

Designemus per $v$, $v'$, $v''$, etc. quantitates, multitudine $\pi$, quarum valores per observationem innotescunt, pendeatque quantitas incognita ab illis tali modo, ut per functionem datam illarum, puta $u$, exhibeatur: sint porro $l$, $l'$, $l''$, etc. valores quotientium differentialum

$$\frac{du}{dv}, \quad \frac{du}{dv'}, \quad \frac{du}{dv''}, \quad \text{etc.}$$

valoribus veris quantitatum $v$, $v'$, $v''$, etc. respondentes. Quemadmodem igitur per substitutionem horum valorum verorum in functione $u$ huius valor verus prodit, ita, si pro $v$, $v'$, $v''$, etc. valores erroribus $e$, $e'$, $e''$, etc. resp. discrepantes substituuntur, obtinebitur valor erroneus incognitae, cuius error statui potest

$$= le + l'e' + l''e'' + \text{etc.}$$

siquidem, quod semper supponemus, errores $e$, $e'$, $e''$, etc. tam exigui sunt, ut (pro functione $u$ non lineari) quadrata et producta negligere liceat. Et quamquam magnitudo errorum $e$, $e'$, $e''$, etc. incerta maneat, tamen incertitudinem tali incognitae determinationi inhaerentem generaliter aestimare licet, et quidem per errorem medium in tali determinatione metuendum, qui per principia commentationis prioris fit

$$= \sqrt{llmm + l'l'm'm' + l''l''m''m'' + \text{etc.}}$$

denotantibus $m$, $m'$, $m''$, etc. errores medios observationum, aut si singulae observationes aequali incertitudini obnoxiae sunt,

$$= m\sqrt{ll + l'l' + l''l'' + \text{etc.}}$$

Manifesto in hoc calculo pro $l$, $l'$, $l''$, etc. aequali iure etiam eos valores quotientium differentialium adoptare licebit, qui valoribus observatis quantitatum $v$, $v'$, $v''$, etc. respondent.

memoir, is greater than $\frac{1}{2}\pi$. In such cases we should prefer the new solution to the old whenever conditional equations can be determined straightforwardly from the problem itself.

2.

Let $v$, $v'$, $v''$, etc. denote $\pi$ quantities whose values are known by observation, and let an unknown quantity depend on these quantities in such a way that it can be written as a function $u$ of them. Let $l$, $l'$, $l''$, etc. be the values of the differential quotients

$$\frac{du}{dv}, \quad \frac{du}{dv'}, \quad \frac{du}{dv''}, \quad \text{etc.}$$

at the true values of $v$, $v'$, $v''$, etc. Substituting the true values into the function $u$ gives us the true value of the unknown. On the other hand, substituting values that have errors $e$, $e'$, $e''$, etc. gives us a value of the unknown whose error is

$$le + l'e' + l''e'' + \text{etc.}$$

Here, as always, we assume the errors $e$, $e'$, $e''$, etc. are so small that (for $u$ nonlinear) we may ignore their squares and products.

Although the sizes of the errors $e$, $e'$, $e''$, etc. may be uncertain, we can nonetheless estimate the uncertainty inherent in this kind of estimate. In fact, by the principles of the proceeding memoir, the mean error to be feared in this estimate is

$$\sqrt{llmm + l'l'm'm' + l''l''m''m'' + \text{etc.}},$$

where $m$, $m'$, $m''$, etc. are the mean errors in the observations. If each observation is of the same uncertainty, then the mean error is

$$m\sqrt{ll + l'l' + l''l'' + \text{etc.}}$$

Clearly we may replace $l$, $l'$, $l''$, etc. by the values of the differential quotients at the observed quantities $v$, $v'$, $v''$, etc.

3.

Quoties quantitates $v$, $v'$, $v''$, etc. penitus inter se sunt independentes, incognita unico tantum modo per illas determinari poterit: quamobrem tunc illam incertitudinem nullo modo nec evitare neque diminuere licet, et circa valorem incognitae ex observationibus deducendum nihil arbitrio relinquitur.

At longe secus se habet res, quoties inter quantitates $v$, $v'$, $v''$, etc. mutua dependentia intercedit, quam per $\sigma$ aequationes conditionales

$$X = 0, \quad Y = 0, \quad Z = 0, \quad \text{etc.}$$

exprimi supponemus, denotantibus $X$, $Y$, $Z$, etc. functiones datas indeterminatarum $v$, $v'$, $v''$, etc. In hoc casu incognitam nostram infinitis modis diversis per combinationes quantitatum $v$, $v'$, $v''$, etc. determinare licet, quum manifesto loco functionis $u$ adoptari possit quaecunque alia $U$ ita comparata, ut $U - u$ indefinite evanescat, statuendo $X = 0$, $Y = 0$, $Z = 0$, etc.

In applicatione ad casum determinatum nulla quidem hinc prodiret differentia respectu valoris incognitae, si observationes absoluta praecisione gauderent: sed quatenus hae erroribus obnoxiae manent, manifesto in genere alia combinatio alium valorem incognitae afferet. Puta, loco erroris

$$le + l'e' + l''e'' + \text{etc.}$$

quem functionem $u$ commiserat, iam habebimus

$$Le + L'e' + L''e'' + \text{etc.}$$

si functionem $U$ adoptamus, atque valores quotientium differentialium $\frac{\mathrm{d}U}{\mathrm{d}v}$, $\frac{\mathrm{d}U}{\mathrm{d}v'}$, $\frac{\mathrm{d}U}{\mathrm{d}v''}$, etc. resp. per $L$, $L'$, $L''$, etc. denotamus. Et quamquam errores ipsos assignare nequeamus, tamen errores medios in diversis observationum combinationibus metuendos inter se comparare licebit: optimaque combinatio ea erit, in qua hic error medius quam minimus evadit. Qui quum fiat

$$\sqrt{LLmm + L'L'm'm' + L''L''m''m'' + \text{etc.}}$$

in id erit incumbendum, ut aggregatum $LLmm + L'L'm'm' + L''L''m''m'' + \text{etc.}$ nanciscatur valorem minimum.

3.

As long as the quantities $v$, $v'$, $v''$, etc. are independent, the unknown can be estimated from them in only one way. Hence there is no way to eliminate or diminish the uncertainty, and there is no freedom in the value of the unknown derived from the observations.

However, things are quite different when there is a dependency among $v$, $v'$, $v''$, etc. We will represent this dependency by $\sigma$ conditional equations

$$X = 0, \quad Y = 0, \quad Z = 0, \quad \text{etc.},$$

in which $X$, $Y$, $Z$, etc. are known functions of $v$, $v'$, $v''$, etc. In this case, our unknown may be estimated by infinitely many combinations of $v$, $v'$, $v''$, etc., since in place of $u$ we may use any function $U$ for which $U - u$ vanishes whenever $X = 0$, $Y = 0$, $Z = 0$, etc.

In any given problem, this change would make no difference in the value of the unknown provided the observations were absolutely precise. But if the observations are in error, different combinations will generally yield different values of the unknown. Specifically, if we use the function $U$ and write $L$, $L'$, $L''$, etc. for its differential quotients $\frac{dU}{dv}$, $\frac{dU}{dv'}$, $\frac{dU}{dv''}$, etc., instead of the error

$$le + l'e' + l''e'' + \text{etc.},$$

which we get from $u$, we will make an error

$$Le + L'e' + L''e'' + \text{etc.}$$

Although we cannot determine the errors themselves, we may compare the mean errors to be feared in various combinations of the observations. The best combination will be the one for which the mean error is smallest. Since the mean error is

$$\sqrt{LLmm + L'L'm'm' + L''L''m''m'' + \text{etc.}},$$

the problem is to make the sum $LLmm + L'L'm'm' + L''L''m''m'' + \text{etc.}$ as small as possible.

4.

Quum varietas infinita functionum $U$, quae secundum contitionem in art. praec. enunciatam ipsius $u$ vice fungi possunt, eatenus tantume hic consideranda veniat, quatenus diversa systemata valorum coëfficientium $L$, $L'$, $L''$, etc. inde sequuntur, indagare oportebit ante omnia nexum, qui inter cuncta systemata admissibilia locum habere debet. Designemus valores determinatos quotientium differentialium partialium

$$\begin{array}{llll} \frac{\mathrm{d}X}{\mathrm{d}v}, & \frac{\mathrm{d}X}{\mathrm{d}v'}, & \frac{\mathrm{d}X}{\mathrm{d}v''} & \text{etc.} \\ \frac{\mathrm{d}Y}{\mathrm{d}v}, & \frac{\mathrm{d}Y}{\mathrm{d}v'}, & \frac{\mathrm{d}Y}{\mathrm{d}v''} & \text{etc.} \\ \frac{\mathrm{d}Z}{\mathrm{d}v}, & \frac{\mathrm{d}Z}{\mathrm{d}v'}, & \frac{\mathrm{d}Z}{\mathrm{d}v''} & \text{etc. etc.} \end{array}$$

quos obtinent, si ipsis $v$, $v'$, $v''$, etc. valores veri tribuuntur, resp. per

$$\begin{array}{llll} a, & a', & a'' & \text{etc.} \\ b, & b', & b'' & \text{etc.} \\ c, & c', & c'' & \text{etc. etc.} \end{array}$$

patetque, si ipsis $v$, $v'$, $v''$, etc. accedere concipiantur talia incrementa $dv$, $dv'$, $dv''$, etc., per quae $X$, $Y$, $Z$, etc. non mutentur, adeoque singulae maneant $= 0$, i.e. satisfacientia aequationibus

$$\begin{aligned} 0 &= a\mathrm{d}v + a'\mathrm{d}v' + a''\mathrm{d}v'' + \text{etc.} \\ 0 &= b\mathrm{d}v + b'\mathrm{d}v' + b''\mathrm{d}v'' + \text{etc.} \\ 0 &= c\mathrm{d}v + c'\mathrm{d}v' + c''\mathrm{d}v'' + \text{etc. etc.} \end{aligned}$$

etiam $u - U$ non mutari debere, adeoque fieri

$$0 = (l - L)\mathrm{d}v + (l' - L')\mathrm{d}v' + (l'' - L'')\mathrm{d}v'' + \text{etc.}$$

Hinc facile concluditur, coëfficientes $L$, $L'$, $L''$, etc. contentos esse debere sub formulis talibus

$$\begin{aligned} L &= l + ax + by + cz + \text{etc.} \\ L' &= l' + a'x + b'y + c'z + \text{etc.} \\ L'' &= l'' + a''x + b''y + c''z + \text{etc.} \end{aligned}$$

etc., denotantibus $x$, $y$, $z$, etc. multiplicatores determinatos. Vice versa patet, si systema multiplicatorum determinatorum $x$, $y$, $z$, etc. ad lubitum assumatur, semper assignari posse functionem $U$ talem, cui valores ipsorum $L$, $L'$, $L''$, etc.

4.

According to the condition stated in the preceding article, there are an infinite number of functions $U$ which can be used in place of $u$. Since these come into consideration only through their coefficients $L$, $L'$, $L''$, etc. we should first of all investigate which systems of coefficients are admissible. Let the values of the partial differential quotients

$$\begin{array}{llll} \frac{dX}{dv}, & \frac{dX}{dv'}, & \frac{dX}{dv''} & \text{etc.} \\ \frac{dY}{dv}, & \frac{dY}{dv'}, & \frac{dY}{dv''} & \text{etc.} \\ \frac{dZ}{dv}, & \frac{dZ}{dv'}, & \frac{dZ}{dv''} & \text{etc. etc.} \end{array}$$

at the true values of $v$, $v'$, $v''$, etc. be denoted by

$$\begin{array}{llll} a, & a', & a'' & \text{etc.} \\ b, & b', & b'' & \text{etc.} \\ c, & c', & c'' & \text{etc. etc.} \end{array}$$

Obviously if $v$, $v'$, $v''$, etc. are given increments $dv$, $dv'$, $dv''$, etc. for which $X$, $Y$, $Z$, etc. do not change, that is, increments satisfying the equations

$$\begin{aligned} 0 &= a\mathrm{d}v + a'\mathrm{d}v' + a''\mathrm{d}v'' + \text{etc.} \\ 0 &= b\mathrm{d}v + b'\mathrm{d}v' + b''\mathrm{d}v'' + \text{etc.} \\ 0 &= c\mathrm{d}v + c'\mathrm{d}v' + c''\mathrm{d}v'' + \text{etc. etc.} \end{aligned}$$

then $u - U$ must not change, so that

$$0 = (l - L)\mathrm{d}v + (l' - L')\mathrm{d}v' + (l'' - L'')\mathrm{d}v'' + \text{etc.}$$

From this we easily see that $L$, $L'$, $L''$, etc. must have the form

$$\begin{aligned} L &= l + ax + by + cz + \text{etc.} \\ L' &= l' + a'x + b'y + c'z + \text{etc.} \\ L'' &= l'' + a''x + b''y + c''z + \text{etc.} \end{aligned}$$

etc., where $x$, $y$, $z$, etc. denote constant multipliers.[10]

Conversely, if we are given a system of constant multipliers $x$, $y$, $z$, etc., we will always be able to find a function $U$, whose values $L$, $L'$, $L''$, etc. satisfy the

[10]The conclusion is not all that easy. Gauss's assertion is the same as saying that any vector orthogonal to the null space of a matrix is in the column space of the transpose.

his aequationibus conformes respondeant, et quae pro conditione in art. praec. enunciata ipsius $u$ vice funge possit: quin adeo hoc infinitis modis diversis effici posse. Modus simplicissimus erit statuere $U = u + xX + yY + zZ + \text{etc.}$; generalius statuere licet $U = u + xX + yY + zZ + \text{etc.} + u'$, denotante $u'$ talem functionem indeterminatarum $v$, $v'$, $v''$, etc., quae semper evanescit pro $X = 0$, $Y = 0$, $Z = 0$, etc., et cuius valor in casu determinato de quo agitur fit maximus vel minimus. Sed ad instutum nostrum nulla hinc oritur differentia.

## 5.

Facile iam erit, multiplicatoribus $x$, $y$, $z$, etc. valores tales tribuere, ut aggregatum

$$LLmm + L'L'm'm' + L''L''m''m'' + \text{etc.}$$

assequatur valorem minimum. Manifesto ad hunc finem haud opus est congnitione errorum mediorum $m$, $m'$, $m''$, etc. absoluta, sed sufficit ratio, quam inter se tenent. Introducemus itaque ipsorum loco pondera observationum $p$, $p'$, $p''$, etc., i.e. numeros quadratis $mm$, $m'm'$, $m''m''$, etc. reciproce proportionales, pondere alicuius observationis ad lubitum pro unitate accepto. Quantitates $x$, $y$, $z$, etc. itaque sic determinare debebunt ut polynomium indefinitum

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

nanciscatur valorem minimum, quod fieri supponemus per valores *determinatos* $x^0$, $y^0$, $z^0$, etc.

Introducendo denotationes sequentes

$$\begin{aligned}
&\tfrac{aa}{p} + \tfrac{a'a'}{p'} + \tfrac{a''a''}{p''} + \text{etc.} = [aa]\\
&\tfrac{ab}{p} + \tfrac{a'b'}{p'} + \tfrac{a''b''}{p''} + \text{etc.} = [ab]\\
&\tfrac{ac}{p} + \tfrac{a'c'}{p'} + \tfrac{a''c''}{p''} + \text{etc.} = [ac]\\
&\tfrac{bb}{p} + \tfrac{b'b'}{p'} + \tfrac{b''b''}{p''} + \text{etc.} = [bb]\\
&\tfrac{bc}{p} + \tfrac{b'c'}{p'} + \tfrac{b''c''}{p''} + \text{etc.} = [bc]\\
&\tfrac{cc}{p} + \tfrac{c'c'}{p'} + \tfrac{c''c''}{p''} + \text{etc.} = [cc]\\
&\text{etc.}
\end{aligned}$$

above equations and which can be used in place of $u$ according to the conditions of the preceding article. In fact, this can be done in infinitely many ways. The simplest is to set $U = u + xX + yY + zZ + \text{etc.}$ More generally, we may set $U = u + xX + yY + zZ + \text{etc.} + u'$, where $u'$ is a function of $v$, $v'$, $v''$, etc. that vanishes whenever $X = 0$, $Y = 0$, $Z = 0$, etc. and whose value for the problem in question is either maximal or minimal.[11] But this makes no difference for our purposes.

5.

It is now easy to assign values to the multipliers $x$, $y$, $z$, etc. that minimize

$$LLmm + L'L'm'm' + L''L''m''m'' + \text{etc.}$$

Obviously we do not need to know the actual values of the mean errors $m$, $m'$, $m''$, etc.—their ratios are sufficient. We will therefore introduce the weights $p$, $p'$, $p''$, etc., i.e., numbers proportional to the reciprocals of the squares $mm$, $m'm'$, $m''m''$, etc., the weight of some arbitrarily chosen observation being taken as one. The quantities $x$, $y$, $z$, etc. should be determined so that the polynomial

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

is minimized, which we will assumes happens for the constants $x^0$, $y^0$, $z^0$, etc.

If we set

$$\begin{aligned}
\tfrac{aa}{p} + \tfrac{a'a'}{p'} + \tfrac{a''a''}{p''} + \text{etc.} &= [aa] \\
\tfrac{ab}{p} + \tfrac{a'b'}{p'} + \tfrac{a''b''}{p''} + \text{etc.} &= [ab] \\
\tfrac{ac}{p} + \tfrac{a'c'}{p'} + \tfrac{a''c''}{p''} + \text{etc.} &= [ac] \\
\tfrac{bb}{p} + \tfrac{b'b'}{p'} + \tfrac{b''b''}{p''} + \text{etc.} &= [bb] \\
\tfrac{bc}{p} + \tfrac{b'c'}{p'} + \tfrac{b''c''}{p''} + \text{etc.} &= [bc] \\
\tfrac{cc}{p} + \tfrac{c'c'}{p'} + \tfrac{c''c''}{p''} + \text{etc.} &= [cc] \\
\text{etc.}
\end{aligned}$$

[11]That is, its derivatives vanish at the true values of $v$, $v'$, $v''$, ....

etc. nec non

$$\begin{aligned}
&\tfrac{al}{p} + \tfrac{a'l'}{p'} + \tfrac{a''l''}{p''} + \text{etc.} = [al]\\
&\tfrac{bl}{p} + \tfrac{b'l'}{p'} + \tfrac{b''l''}{p''} + \text{etc.} = [bl]\\
&\tfrac{cl}{p} + \tfrac{c'l'}{p'} + \tfrac{c''l''}{p''} + \text{etc.} = [cl]\\
&\text{etc.}
\end{aligned}$$

manifesto conditio minimi requirit, ut fiat

$$\left.\begin{aligned}
&0 = [aa]x^0 + [ab]y^0 + [ac]z^0 + \text{etc.} + [al]\\
&0 = [ab]x^0 + [bb]y^0 + [bc]z^0 + \text{etc.} + [bl]\\
&0 = [ac]x^0 + [bc]y^0 + [cc]z^0 + \text{etc.} + [cl]\\
&\text{etc.}
\end{aligned}\right\}(1)$$

Postquam quantitates $x^0$, $y^0$, $z^0$, etc. per eliminationem hinc derivatae sunt, statuetur

$$\left.\begin{aligned}
&ax^0 + by^0 + cz^0 + \text{etc.} + l = L\\
&a'x^0 + b'y^0 + c'z^0 + \text{etc.} + l' = L'\\
&a''x^0 + b''y^0 + c''z^0 + \text{etc.} + l'' = L''\\
&\text{etc.}
\end{aligned}\right\}(2)$$

His ita factis, functio quantitatum $v$, $v'$, $v''$, etc. ea ad determinationem incognitae nostrae maxime idonea minimaeque incertitudini obnoxia erit, cuius quotientes differentiales partiales in casu determanato de quo agitur habent valores $L$, $L'$, $L''$, etc. resp., pondusque huius determinationis, quod per $P$ denotabimus, erit

$$\frac{1}{\frac{LL}{p} + \frac{L'L'}{p'} + \frac{L''L''}{p'} + \text{etc.}} \qquad (3)$$

sive $\frac{1}{P}$ erit valor polynomii supra allati pro eo systemate valorum quantitatum $x$, $y$, $z$, etc., per quod aequationibus (1) satisfit.

## 6.

In art. praec. eam functionem $U$ dignoscere docuimus, quae determinationi maxime idoneae incognitae nostrae inservit: videamus iam, quemnam *valorem* incongnita hoc modo assequatur. Designetur hic valor per $K$, qui itaque oritur, si in $U$ valores observati quantitatum $v$, $v'$, $v''$, etc. substituuntur; per eandem substitutionem obtineat functio $u$ valorem $k$; denique sit $\kappa$ valor verus incognitae, and

$$\begin{aligned}
&\tfrac{al}{p} + \tfrac{a'l'}{p'} + \tfrac{a''l''}{p''} + \text{etc.} = [al]\\
&\tfrac{bl}{p} + \tfrac{b'l'}{p'} + \tfrac{b''l''}{p''} + \text{etc.} = [bl]\\
&\tfrac{cl}{p} + \tfrac{c'l'}{p'} + \tfrac{c''l''}{p''} + \text{etc.} = [cl]\\
&\text{etc.},
\end{aligned}$$

then the conditions for a minimum obviously require that

$$\left.\begin{aligned}
0 &= [aa]x^0 + [ab]y^0 + [ac]z^0 + \text{etc.} + [al]\\
0 &= [ab]x^0 + [bb]y^0 + [bc]z^0 + \text{etc.} + [bl]\\
0 &= [ac]x^0 + [bc]y^0 + [cc]z^0 + \text{etc.} + [cl]\\
&\text{etc.}
\end{aligned}\right\} (1)$$

Having found $x^0$, $y^0$, $z^0$, etc. by elimination in the above equations, we may set

$$\left.\begin{aligned}
&ax^0 + by^0 + cz^0 + \text{etc.} + l = L\\
&a'x^0 + b'y^0 + c'z^0 + \text{etc.} + l' = L'\\
&a''x^0 + b''y^0 + c''z^0 + \text{etc.} + l'' = L''\\
&\text{etc.}
\end{aligned}\right\} (2)$$

Then the function of $v$, $v'$, $v''$, etc. that is the best estimate of the unknown and is least subject to error is the one whose partial differential quotients for the problem in question are $L$, $L'$, $L''$, etc. The weight $P$ of this estimate is

$$\frac{1}{\frac{LL}{p} + \frac{L'L'}{p'} + \frac{L''L''}{p'} + \text{etc.}}. \quad (3)$$

In other words, $\frac{1}{P}$ is the value of the above polynomial at the values of $x$, $y$, $z$, etc. that satisfy (1).

6.

In the preceding article we showed how to select the function $U$ that best estimates our unknown. Let us now see what the *value* of this estimate is. Let $K$ denote the value of $U$ at the observed values of the quantities $v$, $v'$, $v''$, etc., and let $k$ denote the corresponding value of $u$. Let $\kappa$ be the true value of the unknown, which the true values of $v$, $v'$, $v''$, etc. would give if we could substitute them into qui proin e valoribus veris quantitatum $v$, $v'$, $v''$, etc. proditurus esset, si hos vel in $U$ vel in $u$ substituere possemus. Hinc itaque erit

$$\begin{aligned} k &= \kappa + le + l'e' + l''e'' + \text{etc.} \\ K &= \kappa + Le + L'e' + L''e'' + \text{etc.} \end{aligned}$$

adeoque

$$K = k + (L-l)e + (L'-l')e' + (L''-l'')e'' + \text{etc.}$$

Substituendo in hac aequatione pro $L-l$, $L'-l'$, $L''-l''$, etc. valores ex (2), statuendoque

$$\left.\begin{aligned} ae + a'e' + a''e'' + \text{etc.} &= \mathcal{A} \\ be + b'e' + b''e'' + \text{etc.} &= \mathcal{B} \\ ce + c'e' + c''e'' + \text{etc.} &= \mathcal{C} \end{aligned}\right\} (4)$$

etc., habebimus

$$K = k + \mathcal{A}x^0 + \mathcal{B}y^0 + \mathcal{C}z^0 + \text{etc.} \qquad (5)$$

Valores quantitatum $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, etc. per formulas (4) quidem calculare non possumus, quum errores $e$, $e'$, $e''$, etc. maneant incogniti; at sponte manifestum est, illos nihil aliud esse, nisi valores functionum $X$, $Y$, $Z$, etc., qui prodeunt, si pro $v$, $v'$, $v''$, etc. valores observati substitutuuntur. Hoc modo systema aequationum (1), (3), (5) completam problematis nostri solutionem exhibet, quum ea, quae in fine art. 2. de computo quantaitatum $l$, $l'$, $l''$, etc., valoribus observatis quantitatum $v$, $v'$, $v''$, etc. superstruendo monimus, manifesto aequali iure ad computum quantitatum $a$, $a'$, $a''$, etc. $b$, $b'$, $b''$, etc. etc. extendere liceat.

7.

Loco formula (3), pondus determinationis maxime plausibilis exprimentis, plures aliae exhiberi possunt, quas evolvere operae pretium erit.

Primo observamus, si aequationes (2) resp. per $\frac{a}{p}$, $\frac{a'}{p'}$, $\frac{a''}{p''}$, etc. multiplicentur et addantur, prodire

$$[aa]x^0 + [ab]y^0 + [ac]z^0 + \text{etc.} = \frac{aL}{p} + \frac{a'L'}{p'} + \frac{a''L''}{p''} + \text{etc.}$$

Pars ad laevam fit $= 0$, partem ad dextram iuxta analogiam per $[aL]$ denotamus: habebmus itaque

$$[aL] = 0$$

either $U$ or $u$. Then

$$\begin{aligned} k &= \kappa + le + l'e' + l''e'' + \text{etc.},\\ K &= \kappa + Le + L'e' + L''e'' + \text{etc.}, \end{aligned}$$

so that

$$K = k + (L-l)e + (L'-l')e' + (L''-l'')e'' + \text{etc.}$$

If we replace $L-l$, $L'-l'$, $L''-l''$, etc. by their values from (2) and set

$$\left.\begin{aligned} ae + a'e' + a''e'' + \text{etc.} &= \mathcal{A}\\ be + b'e' + b''e'' + \text{etc.} &= \mathcal{B}\\ ce + c'e' + c''e'' + \text{etc.} &= \mathcal{C} \end{aligned}\right\} \quad (4)$$

etc., we get

$$K = k + \mathcal{A}x^0 + \mathcal{B}y^0 + \mathcal{C}z^0 + \text{etc.} \qquad (5)$$

Since the errors $e$, $e'$, $e''$, etc. are unknown, the quantities $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, etc. cannot be calculated from (4). However, it is clear that they are just the values of $X$, $Y$, $Z$, etc. at the observed values of $v$, $v'$, $v''$, etc. Thus the systems of equations (1), (3), and (5) give a complete solution to our problem, since what we said at the end of Art. 2 about computing $l$, $l'$, $l''$, etc. from the observed values of $v$, $v'$, $v''$, etc. applies equally to the computation of $a$, $a'$, $a''$, etc. $b$, $b'$, $b''$, etc. etc.

## 7.

There are several formulas other than (3) that give the weight of the most reliable estimate, and it will be worthwhile to derive them.

First note that if the equations in (2) are multiplied by $\frac{a}{p}$, $\frac{a'}{p'}$, $\frac{a''}{p''}$, etc. and added we get

$$[aa]x^0 + [ab]y^0 + [ac]z^0 + \text{etc.} = \tfrac{aL}{p} + \tfrac{a'L'}{p'} + \tfrac{a''L''}{p''} + \text{etc.}$$

The left-hand side of this equation is zero. If by analogy we denote the right-hand side by $[aL]$, we have

$$[aL] = 0,$$

et prorsus simili modo

$$[bL] = 0, \quad [cL] = 0, \quad \text{etc.}$$

Multiplicando porro aequatones (2) deinceps per $\frac{L}{p}$, $\frac{L'}{p'}$, $\frac{L''}{p''}$, etc., et addendo, invenimus

$$\frac{lL}{p} + \frac{l'L'}{p'} + \frac{l''L''}{p''} + \text{etc.} = \frac{LL}{p} + \frac{L'L'}{p'} + \frac{L''L''}{p''} + \text{etc.}$$

unde obtinemus expressionem *secundam* pro pondere,

$$P = \frac{1}{\frac{lL}{p} + \frac{l'L'}{p'} + \frac{l''L''}{p''} + \text{etc.}}$$

Denique multiplicando aequationes (2) per $\frac{l}{p}$, $\frac{l'}{p'}$, $\frac{l''}{p''}$, etc., et addendo, pervenimus ad expressionem *tertiam* ponderis

$$P = \frac{1}{[al]x^0 + [bl]y^0 + [cl]z^0 + \text{etc.} + [ll]}$$

si ad instar reliquarum denotationum statuimus

$$\frac{ll}{p} + \frac{l'l'}{p'} + \frac{l''l''}{p''} + \text{etc.} = [ll]$$

Hinc adiumento aequationum (1) facile fit transitus ad *expressionem quartam,* quam exhibemus:

$$\begin{aligned} \frac{1}{P} = [ll] - \ & [aa]x^0x^0 - \ [bb]y^0y^0 - \ [cc]z^0z^0 - \text{etc.} \\ & - 2[ab]x^0y^0 - 2[ac]x^0z^0 - 2[bc]y^0z^0 - \text{etc.} \end{aligned}$$

8.

Solutio generalis, quam hactenus explicavimus, ei potissimum casui adaptata est, ubi *una* incognita a quantitatibus observatis pendens determinanda est. Quoties vero plures incognitae ab iisdem observationibus pendentes valores maxime plausibles exspectant, vel quoties adhuc incertum est, quasnam potissimum incognitas ex observationibus derivare oporteat, has alia ratione praeparare convenient, cuius evolutionem iam aggredimur.

Considerabimus quantitates $x$, $y$, $z$, etc. tamquam indeterminatas, statuemus

$$\left.\begin{aligned} [aa]x + [ab]y + [ac]z + \text{etc.} &= \xi \\ [ab]x + [bb]y + [bc]z + \text{etc.} &= \eta \\ [ac]x + [bc]y + [cc]z + \text{etc.} &= \zeta \end{aligned}\right\} (6)$$

and likewise

$$[bL] = 0, \quad [cL] = 0, \quad \text{etc.}$$

If we now multiply the equations (2) in turn by $\frac{L}{p}$, $\frac{L'}{p'}$, $\frac{L''}{p''}$, etc., we find that

$$\frac{lL}{p} + \frac{l'L'}{p'} + \frac{l''L''}{p''} + \text{etc.} = \frac{LL}{p} + \frac{L'L'}{p'} + \frac{L''L''}{p''} + \text{etc.},$$

from which we obtain the *second* expression for the weight:

$$P = \frac{1}{\frac{lL}{p} + \frac{l'L'}{p'} + \frac{l''L''}{p''} + \text{etc.}}.$$

Next, multiplying the equations (2) in turn by $\frac{l}{p}$, $\frac{l'}{p'}$, $\frac{l''}{p''}$ etc. and summing, we obtain the *third* expression for the weight:

$$P = \frac{1}{[al]x^0 + [bl]y^0 + [cl]z^0 + \text{etc.} + [ll]},$$

where in conformity with the previous notation we have set

$$\frac{ll}{p} + \frac{l'l'}{p'} + \frac{l''l''}{p''} + \text{etc.} = [ll].$$

From this we easily pass via (1) to the *fourth expression*

$$\begin{aligned} \frac{1}{P} = [ll] - \ & [aa]x^0x^0 - \ [bb]y^0y^0 - \ [cc]z^0z^0 - \text{etc.} \\ & - 2[ab]x^0y^0 - 2[ac]x^0z^0 - 2[bc]y^0z^0 - \text{etc.} \end{aligned}$$

8.

The general solution we have just described is best in cases where *one* unknown depending on the observed quantities is to be determined. When the most reliable values of several unknowns depending on the same observations are required or when we do not know at the outset which unknowns we shall have to obtain from the observations, they should be computed in a different way, which we will now derive.

We will take $x$, $y$, $z$, etc. to be variables and set

$$\left.\begin{aligned} [aa]x + [ab]y + [ac]z + \text{etc.} &= \xi \\ [ab]x + [bb]y + [bc]z + \text{etc.} &= \eta \\ [ac]x + [bc]y + [cc]z + \text{etc.} &= \zeta \end{aligned}\right\} (6)$$

etc., supponemusque, per eliminationem hinc sequi

$$\left.\begin{aligned}
[\alpha\alpha]\xi + [\alpha\beta]\eta + [\alpha\gamma]\zeta + \text{etc.} &= x\\
[\beta\alpha]\xi + [\beta\beta]\eta + [\beta\gamma]\zeta + \text{etc.} &= y\\
[\gamma\alpha]\xi + [\gamma\beta]\eta + [\gamma\gamma]\zeta + \text{etc.} &= z
\end{aligned}\right\}\quad(7)$$

etc.

Ante omnia hic observare oportet, coëfficientes symmetrice positos necessario aequales fieri, puta

$$\begin{aligned}
[\beta\alpha] &= [\alpha\beta]\\
[\gamma\alpha] &= [\alpha\gamma]\\
[\gamma\beta] &= [\beta\gamma]\\
&\text{etc.}
\end{aligned}$$

quod quidem e theoria generali eliminationis in aequationibus linearibus sponte sequitur, sed etiam infra, absque illa directe demonstrabitur.

Habebimus itaque

$$\left.\begin{aligned}
&x^0 = -[\alpha\alpha].[al] - [\alpha\beta].[bl] - [\alpha\gamma].[cl] - \text{etc.}\\
&y^0 = -[\beta\alpha].[al] - [\beta\beta].[bl] - [\beta\gamma].[cl] - \text{etc.}\\
&z^0 = -[\gamma\alpha].[al] - [\gamma\beta].[bl] - [\gamma\gamma].[cl] - \text{etc.}\\
&\text{etc.}
\end{aligned}\right\}\quad(8)$$

unde, si statuimus

$$\left.\begin{aligned}
[\alpha\alpha]\mathcal{A} + [\alpha\beta]\mathcal{B} + [\alpha\gamma]\mathcal{C} + \text{etc.} &= A\\
[\alpha\beta]\mathcal{A} + [\beta\beta]\mathcal{B} + [\beta\gamma]\mathcal{C} + \text{etc.} &= B\\
[\alpha\gamma]\mathcal{A} + [\beta\gamma]\mathcal{B} + [\gamma\gamma]\mathcal{C} + \text{etc.} &= C
\end{aligned}\right\}\quad(9)$$

etc., obtinemus

$$L = k - A[al] - B[bl] - C[cl] - \text{etc.}$$

vel si insuper statuimus

$$\left.\begin{aligned}
aA + bB + cC + \text{etc.} &= p\epsilon\\
a'A + b'B + c'C + \text{etc.} &= p'\epsilon'\\
a''A + b''B + c''C + \text{etc.} &= p''\epsilon''
\end{aligned}\right\}\quad(10)$$

etc., erit

$$K = k - l\epsilon - l'\epsilon' - l''\epsilon'' - \text{etc.}\qquad(11)$$

etc. We will also assume that by elimination we have

$$\left.\begin{aligned}
[\alpha\alpha]\xi + [\alpha\beta]\eta + [\alpha\gamma]\zeta + \text{etc.} &= x \\
[\beta\alpha]\xi + [\beta\beta]\eta + [\beta\gamma]\zeta + \text{etc.} &= y \\
[\gamma\alpha]\xi + [\gamma\beta]\eta + [\gamma\gamma]\zeta + \text{etc.} &= z
\end{aligned}\right\} (7)$$

etc.

First of all we note that symmetrically located coefficients must be equal; i.e.,

$$\begin{aligned}
&[\beta\alpha] = [\alpha\beta], \\
&[\gamma\alpha] = [\alpha\gamma], \\
&[\gamma\beta] = [\beta\gamma], \\
&\text{etc.}
\end{aligned}$$

This fact follows directly from the general theory of elimination in linear equations; however, we will give an independent proof below.

We now have

$$\left.\begin{aligned}
&x^0 = -[\alpha\alpha].[al] - [\alpha\beta].[bl] - [\alpha\gamma].[cl] - \text{etc.} \\
&y^0 = -[\beta\alpha].[al] - [\beta\beta].[bl] - [\beta\gamma].[cl] - \text{etc.} \\
&z^0 = -[\gamma\alpha].[al] - [\gamma\beta].[bl] - [\gamma\gamma].[cl] - \text{etc.} \\
&\text{etc.}
\end{aligned}\right\} (8)$$

If we set

$$\left.\begin{aligned}
[\alpha\alpha]\mathcal{A} + [\alpha\beta]\mathcal{B} + [\alpha\gamma]\mathcal{C} + \text{etc.} &= A \\
[\alpha\beta]\mathcal{A} + [\beta\beta]\mathcal{B} + [\beta\gamma]\mathcal{C} + \text{etc.} &= B \\
[\alpha\gamma]\mathcal{A} + [\beta\gamma]\mathcal{B} + [\gamma\gamma]\mathcal{C} + \text{etc.} &= C
\end{aligned}\right\} (9)$$

etc., then we get

$$L = k - \mathcal{A}[al] - \mathcal{B}[bl] - \mathcal{C}[cl] - \text{etc.}$$

If in addition we set

$$\left.\begin{aligned}
aA + bB + cC + \text{etc.} &= p\epsilon \\
a'A + b'B + c'C + \text{etc.} &= p'\epsilon' \\
a''A + b''B + c''C + \text{etc.} &= p''\epsilon''
\end{aligned}\right\} (10)$$

etc., then

$$K = k - l\epsilon - l'\epsilon' - l''\epsilon'' - \text{etc.} \quad (11)$$

9.

Comparatio aequationum (7), (9) docet, quantitates auxiliares $A$, $B$, $C$, etc. esse valores indeterminatarum $x$, $y$, $z$, etc. respondentes valoribus indeterminatarum $\xi$, $\eta$, $\zeta$, etc. his $\xi = \mathcal{A}$, $\eta = \mathcal{B}$, $\zeta = \mathcal{C}$, etc., unde patet haberi

$$\left.\begin{aligned} [aa]A + [ab]B + [ac]C + \text{etc.} &= \mathcal{A} \\ [ab]A + [bb]B + [bc]C + \text{etc.} &= \mathcal{B} \\ [ac]A + [bc]B + [cc]C + \text{etc.} &= \mathcal{C} \end{aligned}\right\} \quad (12)$$

etc. Multiplicando itaque aequationes (10) resp. per $\frac{a}{p}$, $\frac{a'}{p'}$, $\frac{a''}{p''}$, etc. et addendo, obtinemus

$$\left.\begin{aligned} \mathcal{A} &= a\epsilon + a'\epsilon' + a''\epsilon'' + \text{etc.} \\ \text{et prorsus simili modo} \qquad & \\ \mathcal{B} &= b\epsilon + b'\epsilon' + b''\epsilon'' + \text{etc.} \\ \mathcal{C} &= c\epsilon + c'\epsilon' + c''\epsilon'' + \text{etc.} \end{aligned}\right\} \quad (13)$$

etc. Iam quum $\mathcal{A}$ sit valor functionis $X$, si pro $v$, $v'$, $v''$, etc. valores observati substituuntur, facile perspicietur, si his applicentur correctiones $-\epsilon$, $-\epsilon'$, $-\epsilon''$, etc. resp., functionem $X$ hinc adepturam esse valorem 0, et perinde functiones $Y$, $Z$, etc. hinc ad valorem evanescentem reductum iri. Simili ratione ex aequatione (11) colligitur, $K$ esse valorem functionis $u$ ex eadem substitutione emergentem.

Applicationem correctionem $-\epsilon$, $-\epsilon'$, $-\epsilon''$, etc. ad observationes, vocabimus *observationum compensationem,* manifestoque deducti sumus ad conclusionem gravissimam, puta, observationes eo quem docuimus modo compensatas omnibus aequationibus conditionalibus exacte satisfacere, atque cuilibit quantitati ab observationibus quomodocunque pendenti eum ipsum valorem conciliare, qui ex observationum non mutatarum combinatione maxime idonea emergeret. Quum itaque impossible sit, errors ipsos $e$, $e'$, $e''$, etc. ex equationibus conditionalibus eruere, quippe quarum multitudo haud sufficit, saltem *errores maxime plausibiles* nacti sumus qua denominatione quantitiates $\epsilon$, $\epsilon'$, $\epsilon''$, etc. designare licebit.

10.

Quum multitudo observationum maior esse supponatur multitudine aequationum conditionalium, praeter systema correctionum maxime plausibilium $-\epsilon$, $-\epsilon'$, $-\epsilon''$, etc. infinite multa alia inveniri possunt, quae aequationibus conditionalibus satisfaciant, operaeque pretium est indagare, quomodo haec ad illud se habeant.

9.

On comparing equations (7) and (9) we find that the auxiliary quantities $A$, $B$, $C$, etc. are the values of the variables $x$, $y$, $z$, etc. corresponding to values $\xi = \mathcal{A}$, $\eta = \mathcal{B}$, $\zeta = \mathcal{C}$, etc. of the variables $\xi$, $\eta$, $\zeta$, etc. Hence

$$\left.\begin{aligned} [aa]A + [ab]B + [ac]C + \text{etc.} &= \mathcal{A} \\ [ab]A + [bb]B + [bc]C + \text{etc.} &= \mathcal{B} \\ [ac]A + [bc]B + [cc]C + \text{etc.} &= \mathcal{C} \end{aligned}\right\} \quad (12)$$

etc. If we multiply the equations (10) in turn by $\frac{a}{p}$, $\frac{a'}{p'}$, $\frac{a''}{p''}$, etc. and add, we get

$$\left.\begin{aligned} \mathcal{A} &= a\epsilon + a'\epsilon' + a''\epsilon'' + \text{etc.} \\ \text{and likewise} \qquad & \\ \mathcal{B} &= b\epsilon + b'\epsilon' + b''\epsilon'' + \text{etc.} \\ \mathcal{C} &= c\epsilon + c'\epsilon' + c''\epsilon'' + \text{etc.} \end{aligned}\right\} \quad (13)$$

etc. Now $\mathcal{A}$ is the value of the function $X$ when $v$, $v'$, $v''$, etc. are replaced by their observed values. Hence if we make corrections of $-\epsilon$, $-\epsilon'$, $-\epsilon''$, etc. to $v$, $v'$, $v''$, etc., the function $X$ will become zero. Similarly $Y$, $Z$, etc. will vanish. In the same way we find from (11) that with these corrections the value of $u$ is $K$.

We will call the application of the corrections $-\epsilon$, $-\epsilon'$, $-\epsilon''$, etc. to the observations *the adjustment of the observations.* Clearly we have arrived at an extremely important result: observations adjusted as above satisfy the auxiliary equations exactly and give any quantity depending on the observations the value that comes from the best combination of the unadjusted observations. Since the number of conditional equations is not sufficient to allow us to extract the errors $e$, $e'$, $e''$, etc. we may at least obtain the values $\epsilon$, $\epsilon'$, $\epsilon''$, etc., which may be called the *most plausible errors.*

10.

Along with the system $-\epsilon$, $-\epsilon'$, $-\epsilon''$, etc. of most reliable corrections, there are infinitely many others that satisfy the conditional equations, since the number of observations is assumed greater than the number of conditional equations. It will be worthwhile to investigate how these systems relate to the most reliable system.

Constituant itaque $-E$, $-E'$, $-E''$, etc. tale systema a maxime plausibli diversum, habebimusque

$$\begin{aligned} aE + a'E' + a''E'' + \text{etc.} &= \mathcal{A} \\ bE + b'E' + b''E'' + \text{etc.} &= \mathcal{B} \\ cE + c'E' + c''E'' + \text{etc.} &= \mathcal{C} \end{aligned}$$

etc. Multiplicando has aequationes resp. per $A$, $B$, $C$, etc. et addendo, obtinemus adiumento aequationum (10)

$$p\epsilon E + p'\epsilon' E' + p''\epsilon'' E'' + \text{etc.} = A\mathcal{A} + B\mathcal{B} + C\mathcal{C} + \text{etc.}$$

Prorsus vero simili modo aequationes (13) suppeditant

$$p\epsilon\epsilon + p'\epsilon'\epsilon' + p''\epsilon''\epsilon'' + \text{etc.} = A\mathcal{A} + B\mathcal{B} + C\mathcal{C} + \text{etc.} \tag{14}$$

E combinatione harum duarum aequationum facile deducitur

$$\begin{aligned} &pEE + p'E'E' + p''E''E'' + \text{etc.} \\ &= p\epsilon\epsilon + p'\epsilon'\epsilon' + p''\epsilon''\epsilon'' + \text{etc.} + p(E-\epsilon)^2 + p'(E'-\epsilon')^2 + p'(E'-\epsilon')^2 + \text{etc.} \end{aligned}$$

Aggregatum $pEE + p'E'E' + p''E''E'' +$ etc. itaque necessario *maius* erit aggragato $p\epsilon\epsilon + p'\epsilon'\epsilon' + p''\epsilon''\epsilon'' +$ etc., quod enuntiari potest tamquam.

THEOREMA. *Aggregatum quadratorum correctionum, per quas observationes cum aequationibus conditionalibus conciliare licet, per pondera observationum resp. multiplicatorum, fit minimum, si correctiones maxime plausibles adoptantur.*

Hoc est ipsum principium quadratorum minimorum, ex quo etiam aequationes (12), (10) facile immediate derivari possunt. Ceterum pro hoc aggregato minimo, quod in sequentibus per $S$ denotabimus, aequatio (14) nobis suppeditat expressionem $A\mathcal{A} + B\mathcal{B} + C\mathcal{C} +$ etc.

## 11.

Determinatio errorum maxime plausibilium, quum a coëfficientibus $l$, $l'$, $l''$, etc. independens sit, manifesto praeparationem commodissimam sistit, ad quemvis usum, in quem observationes vertere placuerit. Praeterea perspicuum est, ad illud negotium haud opus esse eliminatione *indefinita* seu cognitione coëfficientium $[\alpha\alpha]$, $[\alpha\beta]$, etc., nihilque aliud requiri, nisi ut quantitates auxiliares $A$, $B$, $C$, etc., quas in sequentibus *correlata* aequationum conditionalium $X = 0$, $Y = 0$, $Z = 0$, etc. vocabimus, ex aequationibus (12) per eliminationem definitam eliciantur atque in formulis (10) substituantur.

Let $-E$, $-E'$, $-E''$, etc. form a system of corrections that is different from the most reliable system. We then have

$$\begin{aligned} aE + a'E' + a''E'' + \text{etc.} &= \mathcal{A} \\ bE + b'E' + b''E'' + \text{etc.} &= \mathcal{B} \\ cE + c'E' + c''E'' + \text{etc.} &= \mathcal{C} \end{aligned}$$

etc. If we multiply these equations in turn by $A$, $B$, $C$, etc. and add, we obtain from (10)

$$p\epsilon E + p'\epsilon' E' + p''\epsilon'' E'' + \text{etc.} = A\mathcal{A} + B\mathcal{B} + C\mathcal{C} + \text{etc.}$$

Similarly, (13) gives

$$p\epsilon\epsilon + p'\epsilon'\epsilon' + p''\epsilon''\epsilon'' + \text{etc.} = A\mathcal{A} + B\mathcal{B} + C\mathcal{C} + \text{etc.} \tag{14}$$

On combining these two equations we easily find that

$$\begin{aligned} &pEE + p'E'E' + p''E''E'' + \text{etc.} \\ &= p\epsilon\epsilon + p'\epsilon'\epsilon' + p''\epsilon''\epsilon'' + \text{etc.} + p(E-\epsilon)^2 + p'(E'-\epsilon')^2 + p'(E'-\epsilon')^2 + \text{etc.} \end{aligned}$$

Thus the sum $pEE + p'E'E' + p''E''E'' + \text{etc.}$ is necessarily *greater* than the sum $p\epsilon\epsilon + p'\epsilon'\epsilon' + p''\epsilon''\epsilon'' + \text{etc.}$, a fact which may be stated as follows.

THEOREM. *The sum of the products of the weights of the observations with the squares of any corrections that cause the conditional equations to be satisfied is minimized by the most plausible corrections.*

This is the principle of least squares, from which we can easily derive (12) and (10) directly. Moreover, equation (14) gives us the expression $A\mathcal{A}+B\mathcal{B}+C\mathcal{C}+\text{etc.}$ for the minimum sum, which we will denote by $S$ in the sequel.

## 11.

In applications involving the observations themselves, the numbers to compute are the most reliable errors, since they do not depend on the coefficients $l$, $l'$, $l''$, etc. Moreover, there is no need here for a *general* elimination or a knowledge of the coefficients $[\alpha\alpha]$, $[\alpha\beta]$, etc. All that is required is to calculate the auxiliary quantities $A$, $B$, $C$, etc. from the equations (12) by a definite elimination and to substitute them into the formulas (10). In the sequel we will call the quantities $A$, $B$, $C$, etc. *the correlates* of the conditional equations $X = 0$, $Y = 0$, $Z = 0$, etc.

Quamquam vero haec methodus nihil desiderandum linquat, quoties quantitatum ab observationibus pendentium valores maxime plausibiles tantummodo requiruntur, tamen res secus se habere videtur, quoties insuper pondus alicuius determinationis in votis est, quum ad hunc finem, prout hac vel illa quatuor expressionum supra traditarum uti placuerit, cognitio quantitatum $L$, $L'$, $L''$, etc., vel saltem cognitio harum $x^0$, $y^0$, $z^0$, etc. necessaria videatur. Hac ratione utile erit, negotium eliminationis accuratius perscrutari, unde via facilior ad pondera quoque invenienda se nobis aperiet.

12.

Nexus quantitatum in hac disquisitione occurrentium haud parum illustratur per introductionem functionis indefinitae secundi ordinis

$$[aa]xx + 2[ab]xy + 2[ac]xz + \text{etc.} + [bb]yy + 2[bc]yz + \text{etc.} + [cc]zz + \text{etc.}$$

quam per $T$ denotabimus. Primo statim obvium est, hanc functionem fieri

$$\frac{(ax+by+cz+\text{etc.})^2}{p} + \frac{(a'x+b'y+c'z+\text{etc.})^2}{p'} + \frac{(a''x+b''y+c''z''+\text{etc.})^2}{p''} + \text{etc.} \qquad (15)$$

Porro patet esse

$$T = x\xi + y\eta + z\zeta + \text{etc.} \qquad (16)$$

et si hic denuo $x$, $y$, $z$, etc. adiumento aequationum (7) per $\xi$, $\eta$, $\zeta$, etc. exprimuntur,

$$T = [\alpha\alpha]\xi\xi + 2[\alpha\beta]\xi\eta + 2[\alpha\gamma]\xi\eta + \text{etc.} + [\beta\beta]\eta\eta + 2[\beta\gamma]\eta\zeta + \text{etc.} + [\gamma\gamma]\zeta\zeta + \text{etc.}$$

Theoria supra evoluta bina systemata valorum determinatorum quantitatum $x$, $y$, $z$, etc. atque $\xi$, $\eta$, $\zeta$, etc. continet; priori, in quo $x = x^0$, $y = y^0$, $z = z^0$, etc. $\xi = -[al]$, $\eta = -[bl]$, $\zeta = -[cl]$, etc., respondebit valor ipsius $T$ hic

$$T = [ll] - \tfrac{1}{P}$$

quod vel per expressionem tertiam ponderis $P$ cum aequatione (16) comparatam, vel per quartam sponte elucet; posteriori, in quo $x = A$, $y = B$, $z = C$, etc. atque $\xi = \mathcal{A}$, $\eta = \mathcal{B}$, $\zeta = \mathcal{C}$, etc., respondet valor $T = S$, uti vel e formulis (10) et (15), vel ex his (14) et (16) manifestum est.

This procedure leaves nothing to be desired as long as only the most plausible values of quantities depending on the observations are required. But things appear less satisfactory when it is a matter of estimating weights, since to use any of the four expressions given above for the weight it would seem that we have to know the quantities $L$, $L'$, $L''$, etc., or at least $x^0$, $y^0$, $z^0$, etc. For this reason it will be useful to examine the elimination process more thoroughly, an examination which will show us a better way of computing weights.

12.

The quadratic function

$$[aa]xx + 2[ab]xy + 2[ac]xz + \text{etc.} + [bb]yy + 2[bc]yz + \text{etc.} + [cc]zz + \text{etc.},$$

which we will call $T$, throws considerable light on the relation between the quantities occurring in this inquiry. In the first place, this function is obviously

$$\frac{(ax+by+cz+\text{etc.})^2}{p} + \frac{(a'x+b'y+c'z+\text{etc.})^2}{p'} + \frac{(a''x+b''y+c''z''+\text{etc.})^2}{p''} + \text{etc.} \qquad (15)$$

Moreover,

$$T = x\xi + y\eta + z\zeta + \text{etc.}, \qquad (16)$$

and if we use (7) to represent $x$, $y$, $z$, etc. in terms of $\xi$, $\eta$, $\zeta$, etc., we get

$$T = [\alpha\alpha]\xi\xi + 2[\alpha\beta]\xi\eta + 2[\alpha\gamma]\xi\eta + \text{etc.} + [\beta\beta]\eta\eta + 2[\beta\gamma]\eta\zeta + \text{etc.} + [\gamma\gamma]\zeta\zeta + \text{etc.}$$

The theory developed above contains two systems of constants for the quantities $x$, $y$, $z$, etc. and $\xi$, $\eta$, $\zeta$, etc. The first is $x = x^0$, $y = y^0$, $z = z^0$, etc. and $\xi = -[al]$, $\eta = -[bl]$, $\zeta = -[cl]$, etc., for which the corresponding value of $T$ is

$$T = [ll] - \tfrac{1}{P}.$$

This equation follows either by comparing the third expression for the weight $P$ with (16) or directly from the fourth expression. For the second system, in which $x = A$, $y = B$, $z = C$, etc. and $\xi = \mathcal{A}$, $\eta = \mathcal{B}$, $\zeta = \mathcal{C}$, etc., we have $T = S$, as is obvious from the formulas (10) and (15) or (14) and (16).

13.

Iam negotium principale consistit in transformatione functionis $T$ ei simile, quam in Theoria Motus Corporum Coelestium art. 182 atque fusius in Disquisitione de elementis ellipticis Palladis exposuimus. Scilicet statuemus (17)

$$\begin{aligned}
[bb,1] &= [bb] - \tfrac{[ab]^2}{[aa]} \\
[bc,1] &= [bc] - \tfrac{[ab][ac]}{[aa]} \\
[bd,1] &= [bd] - \tfrac{[ab][ad]}{[aa]} \\
&\text{etc.} \\
[cc,2] &= [cc] - \tfrac{[ac]^2}{[aa]} - \tfrac{[bc,1]^2}{[bb,1]} \\
[cd,2] &= [cd] - \tfrac{[ac][ad]}{[aa]} - \tfrac{[bc,1][bd,1]}{[bb,1]} \\
&\text{etc.} \\
[dd,3] &= [dd] - \tfrac{[ad]^2}{[aa]} - \tfrac{[bd,1]^2}{[bb,1]} - \tfrac{[cd,2]^2}{[cc,2]}
\end{aligned}$$

etc. etc. Dein statuendo*)

$$\begin{aligned}
[bb,1]y + [bc,1]z + [bd,1]w + \text{etc.} &= \eta' \\
[cc,2]z + [cd,2]w + \text{etc.} &= \zeta'' \\
[dd,3]w + \text{etc.} &= \varphi''' \\
\text{etc.}
\end{aligned}$$

erit

$$T = \tfrac{\xi\xi}{[aa]} + \tfrac{\eta'\eta'}{[bb,1]} + \tfrac{\zeta''\zeta''}{[cc,2]} + \tfrac{\varphi'''\varphi'''}{[dd,3]} + \text{etc.}$$

quantitatesque $\eta'$, $\zeta''$, $\varphi'''$, etc. a $\eta$, $\zeta$, $\varphi$, etc. pendebunt per aequationes sequentes

$$\begin{aligned}
\eta' &= \eta - \tfrac{[ab]}{[aa]}\xi \\
\zeta'' &= \zeta - \tfrac{[ac]}{[aa]}\xi - \tfrac{[bc,1]}{[bb,1]}\eta' \\
\varphi''' &= \varphi - \tfrac{[ad]}{[aa]}\xi - \tfrac{[bd,1]}{[bb,1]}\eta' - \tfrac{[cd,2]}{[cc,2]}\zeta'' \\
&\text{etc.}
\end{aligned}$$

*)In pracedentibus sufficere poterant ternae literae pro variis systematibus quantitatum ad tres primas aequationes conditionales referendae: hoc vero loco, ut algorithmi lex clarius eluceat, quartum adiungere visum est; et quum in serie naturali literas $a$, $b$, $c$; $A$, $B$, $C$; $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ sponte sequantur $d$, $D$, and $\mathcal{D}$ in serie $x$, $y$, $z$, deficiente alphabeto, apposuimus $w$, nec non in hac $\xi$, $\eta$, $\zeta$ hanc $\varphi$.

13.

Our chief problem now becomes one of transforming the function $T$ in the manner I described in Art. 182 of *Theoria Motus Corporum Coelestium* and more fully in *Disquisitione de elementis ellipticis Palladis*. Namely, set (17)

$$\begin{aligned}
[bb,1] &= [bb] - \tfrac{[ab]^2}{[aa]}\\
[bc,1] &= [bc] - \tfrac{[ab][ac]}{[aa]}\\
[bd,1] &= [bd] - \tfrac{[ab][ad]}{[aa]}\\
&\text{etc.}\\
[cc,2] &= [cc] - \tfrac{[ac]^2}{[aa]} - \tfrac{[bc,1]^2}{[bb,1]}\\
[cd,2] &= [cd] - \tfrac{[ac][ad]}{[aa]} - \tfrac{[bc,1][bd,1]}{[bb,1]}\\
&\text{etc.}\\
[dd,3] &= [dd] - \tfrac{[ad]^2}{[aa]} - \tfrac{[bd,1]^2}{[bb,1]} - \tfrac{[cd,2]^2}{[cc,2]}
\end{aligned}$$

etc. etc. If we then set*)

$$\begin{aligned}
[bb,1]y + [bc,1]z + [bd,1]w + \text{etc.} &= \eta'\\
[cc,2]z + [cd,2]w + \text{etc.} &= \zeta''\\
[dd,3]w + \text{etc.} &= \varphi'''\\
\text{etc.}&
\end{aligned}$$

we get

$$T = \tfrac{\xi\xi}{[aa]} + \tfrac{\eta'\eta'}{[bb,1]} + \tfrac{\zeta''\zeta''}{[cc,2]} + \tfrac{\varphi'''\varphi'''}{[dd,3]} + \text{etc.},$$

where the quantities $\eta'$, $\zeta''$, $\varphi'''$, etc. are related to $\eta$, $\zeta$, $\varphi$, etc. by the equations

$$\begin{aligned}
\eta' &= \eta - \tfrac{[ab]}{[aa]}\xi\\
\zeta'' &= \zeta - \tfrac{[ac]}{[aa]}\xi - \tfrac{[bc,1]}{[bb,1]}\eta'\\
\varphi''' &= \varphi - \tfrac{[ad]}{[aa]}\xi - \tfrac{[bd,1]}{[bb,1]}\eta' - \tfrac{[cd,2]}{[cc,2]}\zeta''\\
&\text{etc.}
\end{aligned}$$

*)In what went before we could make do with three letters for the various systems of quantities corresponding to the first three conditional equations. But here it is advisable to add a fourth to make the steps of the algorithm clear. Although $d$, $D$, and $\mathcal{D}$ naturally follow in the series $a$, $b$, $c$; $A$, $B$, $C$; and $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, there is no letter after $x$, $y$, $z$, and we have adjoined $w$. We also adjoin $\varphi$ to $\xi$, $\eta$, $\zeta$.

Facile iam omnes formulae ad propositum nostrum necessariae hinc desumuntur. Scilicet ad determinationem correlatorum $A$, $B$, $C$, etc, statuemus (18)

$$\begin{aligned}
\mathcal{B}' &= \mathcal{B} - \tfrac{[ab]}{[aa]}\mathcal{A} \\
\mathcal{C}'' &= \mathcal{C} - \tfrac{[ac]}{[aa]}\mathcal{A} - \tfrac{[bc,1]}{[bb,1]}\,\mathcal{B}' \\
\mathcal{D}''' &= \mathcal{D} - \tfrac{[ad]}{[aa]}\mathcal{A} - \tfrac{[bd,1]}{[bb,1]}\mathcal{B}' - \tfrac{[cd,2]}{[cc,2]}\mathcal{C}''
\end{aligned}$$

etc., ac dien $A$, $B$, $C$, $D$, etc. eruentur per formulas sequentes, et quidem ordine inverso, incipiendo ab ultima,

$$\left.\begin{aligned}
[aa]A + [ab]B \quad + [ac]C \quad + [ad]C \quad + \text{etc.} &= \mathcal{A} \\
[bb,1]B + [bc,1]C + [bd,1]D + \text{etc.} &= \mathcal{B}' \\
[cc,2]C + [cd,2]C + \text{etc.} &= \mathcal{C}'' \\
[dd,3]D + \text{etc.} &= \mathcal{D}''' \\
\text{etc.} \qquad\qquad &
\end{aligned}\right\} \quad (19)$$

Pro aggregato $S$ autem habemus formulam novam (20)

$$S = \tfrac{\mathcal{A}\mathcal{A}}{[aa]} + \tfrac{\mathcal{B}'\mathcal{B}'}{[bb,1]} + \tfrac{\mathcal{C}''\mathcal{C}''}{[cc,2]} + \tfrac{\mathcal{D}'''\mathcal{D}'''}{[dd,3]} + \text{etc.}$$

Denique si pondus $P$, quod determinationi maxime plausibili quantitatis per functionem $u$ expressae tribuendum est, desideratur, faciemus (21)

$$\begin{aligned}
[bl,1] &= [bl] - \tfrac{[ab][al]}{[aa]} \\
[cl,2] &= [cl] - \tfrac{[ac][al]}{[aa]} - \tfrac{[bc,1][bl,1]}{[bb,1]} \\
[dl,3] &= [dl] - \tfrac{[ad][al]}{[aa]} - \tfrac{[bd,1][bl,1]}{[bb,1]} - \tfrac{[cd,2][cl,2]}{[cc,2]}
\end{aligned}$$

etc., quo facto erit (22)

$$\tfrac{1}{P} = [ll] - \tfrac{[al]^2}{[aa]} - \tfrac{[bl,1]^2}{[bb,1]} - \tfrac{[cl,2]^2}{[cc,2]} - \tfrac{[dl,3]^2}{[dd,3]} - \text{etc.}$$

Formulae (17) ... (22), quarum simplicitas nihil desiderandum relinquere videtur, solutionem problematis nostri ab omni parte completam exhibent.

## 14.

Postquam problemata primaria absolvimus, adhuc quasdam quaestiones secundarias attingemus, quae huic argumento maiorem lucem affundent.

All the formulas we need may now be easily extracted from the above. Namely, to determine the correlates $A$, $B$, $C$, etc., we set (18)

$$\begin{aligned}
\mathcal{B}' &= \mathcal{B} - \tfrac{[ab]}{[aa]}\mathcal{A} \\
\mathcal{C}'' &= \mathcal{C} - \tfrac{[ac]}{[aa]}\mathcal{A} - \tfrac{[bc,1]}{[bb,1]}\mathcal{B}' \\
\mathcal{D}''' &= \mathcal{D} - \tfrac{[ad]}{[aa]}\mathcal{A} - \tfrac{[bd,1]}{[bb,1]}\mathcal{B}' - \tfrac{[cd,2]}{[cc,2]}\mathcal{C}''
\end{aligned}$$

etc. Then $A$, $B$, $C$, $D$, etc. are found from the formulas

$$\left.\begin{array}{r}
[aa]A + [ab]B + [ac]C + [ad]C + \text{etc.} = \mathcal{A} \\
[bb,1]B + [bc,1]C + [bd,1]D + \text{etc.} = \mathcal{B}' \\
[cc,2]C + [cd,2]C + \text{etc.} = \mathcal{C}'' \\
[dd,3]D + \text{etc.} = \mathcal{D}''' \\
\text{etc.}
\end{array}\right\} (19)$$

in reverse order beginning with the last. On the other hand, we have a new formula (20)

$$S = \tfrac{\mathcal{A}\mathcal{A}}{[aa]} + \tfrac{\mathcal{B}'\mathcal{B}'}{[bb,1]} + \tfrac{\mathcal{C}''\mathcal{C}''}{[cc,2]} + \tfrac{\mathcal{D}'''\mathcal{D}'''}{[dd,3]} + \text{etc.}$$

for the sum $S$. Finally if we want the weight $P$ of the most plausible estimate of the quantity represented by the function $u$, we compute (21)

$$\begin{aligned}
[bl,1] &= [bl] - \tfrac{[ab][al]}{[aa]} \\
[cl,2] &= [cl] - \tfrac{[ac][al]}{[aa]} - \tfrac{[bc,1][bl,1]}{[bb,1]} \\
[dl,3] &= [dl] - \tfrac{[ad][al]}{[aa]} - \tfrac{[bd,1][bl,1]}{[bb,1]} - \tfrac{[cd,2][cl,2]}{[cc,2]}
\end{aligned}$$

etc., from which we have (22)

$$\tfrac{1}{P} = [ll] - \tfrac{[al]^2}{[aa]} - \tfrac{[bl,1]^2}{[bb,1]} - \tfrac{[cl,2]^2}{[cc,2]} - \tfrac{[dl,3]^2}{[dd,3]} - \text{etc.}$$

The formulas (17) ... (22), whose simplicity leaves nothing to be desired, represent a solution to our problem that is complete in every way.

14.

Having resolved the primary problems, we will now touch upon certain secondary questions that will throw more light on the subject.

Primo inquirendum est, num eliminatio, per quam $x$, $y$, $z$, etc. ex $\xi$, $\eta$, $\zeta$, etc. derivare oportet, umquam impossibilis fieri possit. Manifesto hoc eveniret, si functiones $\xi$, $\eta$, $\zeta$, etc. inter se haud independentes essent. Supponamus itaque aliquantisper, unam earum per reliquas iam determinari, ita ut habeatur aequatio identica

$$\alpha\xi + \beta\eta + \gamma\zeta + \text{etc.} = 0$$

denotantibus $\alpha$, $\beta$, $\gamma$, etc. numeros determinatos. Erit itaque

$$\begin{aligned}
\alpha[aa] + \beta[ab] + \gamma[ac] + \text{etc.} &= 0\\
\alpha[ab] + \beta[bb] + \gamma[bc] + \text{etc.} &= 0\\
\alpha[ac] + \beta[bc] + \gamma[cc] + \text{etc.} &= 0
\end{aligned}$$

etc., unde, si statuimus

$$\begin{aligned}
\alpha a + \beta b + \gamma c + \text{etc.} &= p\theta\\
\alpha a' + \beta b' + \gamma c' + \text{etc.} &= p'\theta'\\
\alpha a'' + \beta b'' + \gamma c'' + \text{etc.} &= p''\theta''
\end{aligned}$$

etc., sponte sequitur

$$\begin{aligned}
a\theta + a'\theta' + a''\theta'' + \text{etc.} &= 0\\
b\theta + b'\theta' + b''\theta'' + \text{etc.} &= 0\\
c\theta + c'\theta' + c''\theta'' + \text{etc.} &= 0
\end{aligned}$$

etc., nec non

$$p\theta\theta + p'\theta'\theta' + p''\theta''\theta'' + \text{etc.} = 0$$

quae aequatio, quum omnes $p$, $p'$, $p''$, etc. natura sua sint quantitates positivae, manifesto consistere nequit, nisi fuerit $\theta = 0$, $\theta' = 0$, $\theta'' = 0$, etc.

Iam consideremus valores differentialium completorum $\mathrm{d}X$, $\mathrm{d}Y$, $\mathrm{d}Z$, etc., respondentes valoribus iis quantitatum $v$, $v'$, $v''$, etc., ad quos referuntur observationes. Haec differentialia, puta

$$\begin{aligned}
&a\mathrm{d}v + a'\mathrm{d}v' + a''\mathrm{d}v'' + \text{etc.}\\
&b\mathrm{d}v + b'\mathrm{d}v' + b''\mathrm{d}v'' + \text{etc.}\\
&c\mathrm{d}v + c'\mathrm{d}v' + c''\mathrm{d}v'' + \text{etc.}
\end{aligned}$$

etc., per conclusionem, ad quam modo delati sumus, inter se ita dependentia erunt, ut per $\alpha$, $\beta$, $\gamma$, etc. resp. multiplicata aggregatum identice evanescens producant, sive quod idem est, quodvis ex ipsis (cui quidem respondet multiplicator $\alpha$, $\beta$, $\gamma$, First we must ask if it is ever impossible to perform the elimination by which $x$, $y$, $z$, etc. is derived from $\xi$, $\eta$, $\zeta$, etc. Obviously this happens when the functions $\xi$, $\eta$, $\zeta$, etc. fail to be mutually independent. Let us suppose for the moment that one of the functions is determined by the others, so that we have the identity

$$\alpha\xi + \beta\eta + \gamma\zeta + \text{etc.} = 0,$$

where $\alpha$, $\beta$, $\gamma$, etc. are constants. Then

$$\begin{aligned}
\alpha[aa] + \beta[ab] + \gamma[ac] + \text{etc.} &= 0\\
\alpha[ab] + \beta[bb] + \gamma[bc] + \text{etc.} &= 0\\
\alpha[ac] + \beta[bc] + \gamma[cc] + \text{etc.} &= 0
\end{aligned}$$

etc. If we set

$$\begin{aligned}
\alpha a + \beta b + \gamma c + \text{etc.} &= p\theta\\
\alpha a' + \beta b' + \gamma c' + \text{etc.} &= p'\theta'\\
\alpha a'' + \beta b'' + \gamma c'' + \text{etc.} &= p''\theta''
\end{aligned}$$

etc., it follows that

$$\begin{aligned}
a\theta + a'\theta' + a''\theta'' + \text{etc.} &= 0\\
b\theta + b'\theta' + b''\theta'' + \text{etc.} &= 0\\
c\theta + c'\theta' + c''\theta'' + \text{etc.} &= 0
\end{aligned}$$

etc., and

$$p\theta\theta + p'\theta'\theta' + p''\theta''\theta'' + \text{etc.} = 0.$$

Since $p$, $p'$, $p''$, etc. are positive, this last equation can only hold if $\theta = 0$, $\theta' = 0$, $\theta'' = 0$, etc.

Now consider the values of the total differentials $\mathrm{d}X$, $\mathrm{d}Y$, $\mathrm{d}Z$, etc. at quantities $v$, $v'$, $v''$, etc. corresponding to the observations. By the above results, these differentials, namely,

$$\begin{aligned}
&a\mathrm{d}v + a'\mathrm{d}v' + a''\mathrm{d}v'' + \text{etc.}\\
&b\mathrm{d}v + b'\mathrm{d}v' + b''\mathrm{d}v'' + \text{etc.}\\
&c\mathrm{d}v + c'\mathrm{d}v' + c''\mathrm{d}v'' + \text{etc.}
\end{aligned}$$

etc., are mutually dependent in that the sum of their products with $\alpha$, $\beta$, $\gamma$, etc. vanishes identically. Equivalently, one of them (corresponding to a nonzero value of $\alpha$, $\beta$, $\gamma$, etc.) automatically vanishes whenever all the others vanish. Hence at etc. non evanescens) sponte evanescet, simulac omnia reliqua evanescere supponuntur. Quamobrem ex aequationibus conditionalibus $X = 0$, $Y = 0$, $Z = 0$, etc., una (ad minimum) pro *superflua* habenda est, quippe cui sponte satisfit, simulac reliquis satisfactum est.

Ceterum, si res profundius inspicitur, apparet, hanc conclusionem per se tantum pro ambitu infinite parvo variabilitatis indeterminatarum valere. Scilicet proprie duo casus distinguendi erunt, alter, ubi una aequationum conditionalium $X = 0$, $Y = 0$, $Z = 0$, etc. absolute et generaliter iamiam in reliquis contenta est, quod facile in quovis casu averti poterit; alter, ubi, quasi fortuito, pro iis valoribus concretis quantitatum $v$, $v'$, $v''$, etc. ad quos observationes referuntur, una functionum $X$, $Y$, $Z$, etc., e.g. prima $X$, valorem maximum vel minimum (vel generalius, stationarium) nanciscitur respectu mutationum omnium, quas quantitatibus $v$, $v'$, $v''$, etc., salvis aequationibus $Y = 0$, $Z = 0$, etc., applicare possemus. Attamen quum in disquisitione nostra variabilitas quantitatum tantummodo intra limites tam arctos consideretur, ut ad instar infinite parvae tractari possit, hic casus secundus (qui in praxi vix umquam occurret) eundem effectum habebit, quem primus, puta una aequationum conditionalium tamquam superflua reiicienda erit, certique esse possumus, si omnes aequationes conditionales retentae eo sensu, quem hic intelligimus, ab invicem independentes sint, eliminationem necessario fore possibilem. Ceterum disquisitionem uberiorem, qua hoc argumentum, propter theoreticam subtilitatem potius qum practicam utilitatem haud indignum est, ad aliam occasionem nobis reservare debemus.

## 15.

In commentatione priore art. 37 sqq. methodum docuimus, observationum praecisionem a posteriori quam proxime eruendi. Scilicet si valores approximati $\pi$ quantitatum per observationes aequali praecisione gaudentes innotuerunt, et cum valoribus iis comparantur, qui e valoribus maxime plausibilibus $\rho$ elementorum, a quibus illae pendent, per calculum prodeunt: differentiarum quadrata addere, aggregatumque per $\pi - \rho$ dividere oportet, quo facto quotiens considerari poterit tamquam valor approximatus quadrati erroris medii tali observationum generi inhaerentis. Quoties observationes inaequali praecisione gaudent, haec praecepta eatenus tantum mutanda sunt, ut quadrata ante additionem per observationum pondera multiplicari debeant, errorque medius hoc modo prodiens ad observationes referatur, quarum pondus pro unitate acceptum est.

Iam in tractatione praesenti illud aggregatum manifesto quadrat cum aggreleast one of the conditional equations $X = 0$, $Y = 0$, $Z = 0$, etc.—the one which is satisfied whenever the others are—must be considered redundant.

Now when we consider the matter more deeply, it is obvious that this conclusion holds only for infinitely small perturbations of the variables. Two cases must be properly distinguished. The first is where one of the conditional equations $X = 0$, $Y = 0$, $Z = 0$, etc. is a fully general combination of the others, something which can easily be avoided in any particular problem. The second case is where by some chance one of the functions $X$, $Y$, $Z$, etc., say $X$, at the actual values of the observed quantities $v$, $v'$, $v''$, etc. is maximal or minimal (or more generally stationary) with respect to all changes of the quantities $v$, $v'$, $v''$, etc. that preserve the equations $Y = 0$, $Z = 0$, etc. Now in our investigation the variation of the quantities is so tightly constrained that it can be treated as if it were infinitely small. Hence the second case (which seldom occurs in practice) has the same consequence as the first, namely, that one of the conditional equations must be rejected as redundant. If all the remaining equations are mutually independent in the sense understood here, we can be sure the elimination will be possible. This topic is worth a more extended treatment, more for its theoretical subtlety than its practical utility; but I will save it for another occasion.

15.

In the preceding memoir (Art. 37 and the following), I described an a posteriori method for determining the precision of the observations as exactly as possible. Specifically, if we observe $\pi$ values with equal precision and compare them with the values obtained from the most reliable estimates of the $\rho$ unknowns on which they depend, then the sum of squares of the differences divided by $\pi - \rho$ approximates the mean square error inherent in these observations. When the observations have unequal precisions, the procedure is modified by multiplying the squared differences by the weights of the observations before summing. In this case, the resulting mean error refers to the observations whose weights have been taken to be one.

In the present investigation, the sum of squares obviously corresponds to $S$

gato $S$, differentiaque $\pi - \rho$ cum multitudine aequationum conditionalium $\sigma$, quamobrem pro errore medio observationum, quarum pondus $= 1$, habebimus expressionem $\sqrt{\frac{S}{\sigma}}$, quae determinatio eo maiore fide digna erit, quo maior fuerit numerus $\sigma$.

Sed operae pretium erit, hoc etiam independenter a disquisitione priore stabilire. Ad hunc finem quasdam novas denotationes introducere conveniet. Scilicet respondeant valoribus indeterminatarum $\xi$, $\eta$, $\zeta$, etc. his

$$\xi = a, \quad \eta = b, \quad \zeta = c, \quad \text{etc.}$$

valores ipsarum $x$, $y$, $z$, etc. hi

$$x = \alpha, \quad y = \beta, \quad z = \gamma, \quad \text{etc.}$$

ita ut habeatur

$$\begin{aligned} \alpha &= a[\alpha\alpha] + b[\alpha\beta] + c[\alpha\gamma] + \text{etc.} \\ \beta &= a[\alpha\beta] + b[\beta\beta] + c[\beta\gamma] + \text{etc.} \\ \gamma &= a[\alpha\gamma] + b[\beta\gamma] + c[\gamma\gamma] + \text{etc.} \end{aligned}$$

etc. Perinde valoribus

$$\xi = a', \quad \eta = b', \quad \zeta = c', \quad \text{etc.}$$

respondere supponemus hos

$$x = \alpha', \quad y = \beta', \quad z = \gamma', \quad \text{etc.}$$

nec non his

$$\xi = a'', \quad \eta = b'', \quad \zeta = c'', \quad \text{etc.}$$

sequentes

$$x = \alpha'', \quad y = \beta'', \quad z = \gamma'', \quad \text{etc.}$$

et sic porro.

His positis combinatio aequationum (4), (9) suppeditat

$$\begin{aligned} A &= \alpha e + \alpha' e' + \alpha'' e'' + \text{etc.} \\ B &= \beta e + \beta' e' + \beta'' e'' + \text{etc.} \\ C &= \gamma e + \gamma' e' + \gamma'' e'' + \text{etc.} \end{aligned}$$

etc. Quare quum habeatur $S = \mathcal{A}A + \mathcal{B}B + \mathcal{C}C + \text{etc.}$, patet fieri

$$\begin{aligned} S = \ & (ae + a'e' + a''e'' + \text{etc.})(\alpha e + \alpha' e' + \alpha'' e'' + \text{etc.}) \\ & + (be + b'e' + b''e'' + \text{etc.})(\beta e + \beta' e' + \beta'' e'' + \text{etc.}) \\ & + (ce + c'e' + c''e'' + \text{etc.})(\gamma e + \gamma' e' + \gamma'' e'' + \text{etc.}) + \text{etc.} \end{aligned}$$

and the difference $\pi - \rho$ corresponds to the number of conditional equations $\sigma$. Hence we have the expression $\sqrt{\frac{S}{\sigma}}$ for the mean error in the observations whose weights are one. The larger the number $\sigma$ the more trustworthy the estimate.

But it is worthwhile to establish this independently of the previous investigation. For this purpose, it will be convenient to introduce some new notation. Specifically, let the values

$$\xi = a, \quad \eta = b, \quad \zeta = c, \quad \text{etc.}$$

of $\xi$, $\eta$, $\zeta$, etc. be associated with the values

$$x = \alpha, \quad y = \beta, \quad z = \gamma, \quad \text{etc.}$$

of $x$, $y$, $z$, etc., so that

$$\begin{aligned}\alpha &= a[\alpha\alpha] + b[\alpha\beta] + c[\alpha\gamma] + \text{etc.}\\ \beta &= a[\alpha\beta] + b[\beta\beta] + c[\beta\gamma] + \text{etc.}\\ \gamma &= a[\alpha\gamma] + b[\beta\gamma] + c[\gamma\gamma] + \text{etc.}\end{aligned}$$

etc. Likewise let

$$\xi = a', \quad \eta = b', \quad \zeta = c', \quad \text{etc.}$$

be associated with

$$x = \alpha', \quad y = \beta', \quad z = \gamma', \quad \text{etc.}$$

and

$$\xi = a'', \quad \eta = b'', \quad \zeta = c'', \quad \text{etc.}$$

with

$$x = \alpha'', \quad y = \beta'', \quad z = \gamma'', \quad \text{etc.}$$

and so on.

With this notation, a combination of (4) and (9) gives

$$\begin{aligned}A &= \alpha e + \alpha' e' + \alpha'' e'' + \text{etc.}\\ B &= \beta e + \beta' e' + \beta'' e'' + \text{etc.}\\ C &= \gamma e + \gamma' e' + \gamma'' e'' + \text{etc.}\end{aligned}$$

etc. From this it is obvious that

$$\begin{aligned}S = {} & (ae + a'e' + a''e'' + \text{etc.})(\alpha e + \alpha' e' + \alpha'' e'' + \text{etc.})\\ & + (be + b'e' + b''e'' + \text{etc.})(\beta e + \beta' e' + \beta'' e'' + \text{etc.})\\ & + (ce + c'e' + c''e'' + \text{etc.})(\gamma e + \gamma' e' + \gamma'' e'' + \text{etc.}) + \text{etc.}\end{aligned}$$

since $S = \mathcal{A}A + \mathcal{B}B + \mathcal{C}C + \text{etc.}$

16.

Institutionem observationum, per quas valores quantitatum $v$, $v'$, $v''$, etc. erroribus fortuitis $e$, $e'$, $e''$, etc. affectos obtinemus, considerare possumus tamquam experimentum, quod quidem singulorum errorem commissorum magnitudinem docere non valet, attamen, praeceptis quae supra explicavimus adhibitis, valorem quantitatis $S$ subministrat, qui per formulam modo inventam est functio data illorum errorum. In tali experimento errores fortuiti utique alii maiores alii minores prodire possunt; sed quo plures errores concurrunt, eo maior spes aderit, valorem quantitatis $S$ in experimento singulari a valore suo medio parum deviaturum esse. Rei cardo itaque in eo veritur, ut valorem medium quantitatis $S$ stabiliamus. Per principia in commentatione priore exposita, quae hic repetere superfluum esset, invenimus hunc valorem medium

$$\begin{gathered}(a\alpha + b\beta + c\gamma + \text{etc.})mm + (a'\alpha' + b'\beta' + c'\gamma' + \text{etc.})m'm' \\ +(a''\alpha'' + b''\beta'' + c''\gamma'' + \text{etc.})m''m'' + \text{etc.}\end{gathered}$$

Denotando errorem medium observationum talium, quarum pondus $= 1$, per $\mu$, ita ut sit $\mu\mu = pmm = p'm'm' = p''m''m''$, etc., expressio modo inventa ita eshiberi potest:

$$\begin{gathered}(\tfrac{a\alpha}{p} + \tfrac{a'\alpha'}{p'} + \tfrac{a''\alpha''}{p''} + \text{etc.})\mu\mu + (\tfrac{b\beta}{p} + \tfrac{b'\beta'}{p'} + \tfrac{b''\beta''}{p''} + \text{etc.})\mu\mu \\ +(\tfrac{c\gamma}{p} + \tfrac{c'\gamma'}{p'} + \tfrac{c''\gamma''}{p''} + \text{etc.})\mu\mu + \text{etc.}\end{gathered}$$

Sed aggregatum $\frac{a\alpha}{p} + \frac{a'\alpha'}{p'} + \frac{a''\alpha''}{p''} + \text{etc.}$ invenitur

$$= [aa]\cdot[\alpha\alpha] + [ab]\cdot[\alpha\beta] + [ac]\cdot[\alpha\gamma] + \text{etc.}$$

adeoque $= 1$, uti e nexu aequationum (6), (7) facile intelligitur. Perinde fit

$$\begin{gathered}\tfrac{b\beta}{p} + \tfrac{b'\beta'}{p'} + \tfrac{b''\beta''}{p''} + \text{etc.} = 1 \\ \tfrac{c\gamma}{p} + \tfrac{c'\gamma'}{p'} + \tfrac{c''\gamma''}{p''} + \text{etc.} = 1\end{gathered}$$

et sic porro.

Hinc tandem valor medius ipsius $S$ fit $\sigma\mu\mu$, quatenusque igitur valorem fortuitum ipsius $S$ pro medio adoptare licet, erit $\mu = \sqrt{\frac{S}{\sigma}}$.

16.

The act of observing the quantities $v$, $v'$, $v''$, etc., which contaminates them with the random errors $e$, $e'$, $e''$, etc., can be regarded as an experiment. This experiment does not suffice to tell us the sizes of the individual errors. Nonetheless, it does furnish the quantity $S$ as described above, and according to the formula just derived $S$ is a function of the errors. In such an experiment, some of the random errors will be larger and others smaller; but the greater the number of errors, the greater the hope that in any particular experiment the quantity $S$ will differ very little from its mean value. Thus the crux of the matter is to determine the mean value of $S$. By the principles of the preceding memoir — it would be superfluous to repeat them here — we find that the mean value is

$$(a\alpha + b\beta + c\gamma + \text{etc.})mm + (a'\alpha' + b'\beta' + c'\gamma' + \text{etc.})m'm' \\ +(a''\alpha'' + b''\beta'' + c''\gamma'' + \text{etc.})m''m'' + \text{etc.}$$

If we denote the mean error of the observations whose weight is one by $\mu$, so that $\mu\mu = pmm = p'm'm' = p''m''m''$, etc., then the above expression becomes

$$(\tfrac{a\alpha}{p} + \tfrac{a'\alpha'}{p'} + \tfrac{a''\alpha''}{p''} + \text{etc.})\mu\mu + (\tfrac{b\beta}{p} + \tfrac{b'\beta'}{p'} + \tfrac{b''\beta''}{p''} + \text{etc.})\mu\mu \\ +(\tfrac{c\gamma}{p} + \tfrac{c'\gamma'}{p'} + \tfrac{c''\gamma''}{p''} + \text{etc.})\mu\mu + \text{etc.}$$

But the sum $\frac{a\alpha}{p} + \frac{a'\alpha'}{p'} + \frac{a''\alpha''}{p''} + \text{etc.}$ is

$$[aa]\cdot[\alpha\alpha] + [ab]\cdot[\alpha\beta] + [ac]\cdot[\alpha\gamma] + \text{etc.}$$

which from (6) and (7) is easily seen to be one. Moreover

$$\tfrac{b\beta}{p} + \tfrac{b'\beta'}{p'} + \tfrac{b''\beta''}{p''} + \text{etc.} = 1$$
$$\tfrac{c\gamma}{p} + \tfrac{c'\gamma'}{p'} + \tfrac{c''\gamma''}{p''} + \text{etc.} = 1$$

and so on.

The mean value of $S$ is therefore $\sigma\mu\mu$; and to the extent that the random value of $S$ may be taken for its mean, we have $\mu = \sqrt{\frac{S}{\sigma}}$.

## 17.

Quanta fides huic determinationi habenda sit, diiudicare oportet per errorem medium vel in ipsa vel in ipsius quadrato metuendum: posterior erit radix quadrata valoris medii expressionis

$$(\tfrac{S}{\sigma} - \mu\mu)^2$$

cuius evolutio absolvetur per ratiocinia similia iis, quae in commentatione priore artt. 39 sqq. exposita sunt. Quibus brevitatis caussa hic suppressis, formulam ipsam tantum hic apponimus. Scilicet error medius in determinatione quadrati $\mu\mu$ metuendus exprimitur per

$$\sqrt{\tfrac{2\mu^4}{\sigma} + \tfrac{\nu^4 - 3\mu^4}{\sigma\sigma} \cdot N}$$

denotante $\nu^4$ valorem medium biquadratorum errorum, quorum pondus $= 1$, atque $N$ aggregatum

$$(a\alpha + b\beta + c\gamma + \text{etc.})^2 + (a'\alpha' + b'\beta' + c'\gamma' + \text{etc.})^2 + (a''\alpha'' + b''\beta'' + c''\gamma'' + \text{etc.})^2 + \text{etc.}$$

Hoc aggregatum in genere ad formam simpliciorem reduci nequit, sed simili modo ut in art. 40 prioris commentationis ostendi potest, euis valorem semper contineri intra limites $\pi$ et $\frac{\sigma\sigma}{\pi}$. In hypothesi ea, cui theoria quadratorum minimorum ab initio superstructa erat, terminus hoc aggregatum continens, propter $\nu^4 = 3\mu^4$, omnino excidit, praecisioque, quae errori medio, per formulam $\sqrt{\frac{S}{\sigma}}$ determinato, tribuenda est, eadem erit, ac si ex $\sigma$ erroribus exacte cognitis secundum artt. 15, 16 prioris commentationis erutus fuisset.

## 18.

Ad compensationem observationum duo, ut supra vidimus, requiruntur: primum, ut aequationum conditionalium correlata, i.e. numeri $A$, $B$, $C$, etc. aequationibus (12) satisfacientes eruantur, secundum, ut hi numeri in aequationibus (10) substitiuantur. Compensatio hoc modo prodiens dici poterit *perfecta* seu *completa,* ut distinguatur a compensatione *imperfecta* seu *manca*: hac scilicet denominatione designabimus, quae resultant ex iisdem quidem aequationibus (10), sed substratis valoribus quantitatum $A$, $B$, $C$, etc., qui non satisfaciunt aequationibus (12), i.e. qui vel parti tantum satisfaciunt vel nullis. Quod vero attinet ad tales observationum mutationes, quae sub formulis (10) comprehendi nequeunt, a disquisitione praesenti, nec non a denominatione compensationum exclusae sunto.

17.

How far we can trust this last estimate may be judged from the mean error to be feared in the estimate itself or in its square. The latter is the square root of the mean value of the expression

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

which can be derived as in Art. 39 and the following. For brevity we will omit the derivation and only write down the formula. Specifically, the error to be feared in the estimate of the square $\mu\mu$ is

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

where $\nu^4$ is the mean of the fourth power of the errors whose weights are one and $N$ is the sum

$$(a\alpha + b\beta + c\gamma + \text{etc.})^2 + (a'\alpha' + b'\beta' + c'\gamma' + \text{etc.})^2 + (a''\alpha'' + b''\beta'' + c''\gamma'' + \text{etc.})^2 + \text{etc.}$$

In general this sum cannot be simplified. But as in Art. 40 of the earlier memoir, we can show that its value always lies between $\pi$ and $\frac{\sigma\sigma}{\pi}$. Under the hypothesis on which the theory of least squares was originally based,[12] $\nu^4 = 3\mu^4$ and the term containing $N$ drops out. In this case, the precision of the mean error as estimated by $\sqrt{\frac{S}{\sigma}}$ will be the same as if it had been calculated from $\sigma$ actual errors according to Arts. 15–16 of the earlier memoir.

18.

As we have seen above, two things are required to adjust observations. First, we must calculate the correlates of the conditional equations; i.e., numbers $A$, $B$, $C$, etc. satisfying the equations (12). Second we must substitute these numbers in (10). Such an adjustment may be called *perfect* or *complete* as opposed to an *imperfect* or *defective* adjustment. By the latter we mean adjustments that are of the form (10) but are based on values of $A$, $B$, $C$, etc. that fail to satisfy some or all of the equations (12).

In our present investigation we will exclude changes in the observations that are not of the form (10); in fact, we will not even call them adjustments. Since the

[12] I.e., that the distribution of the errors is normal.

Quum, quatenus aequationes (10) locum habent, aequationes (13) ipsis (12) omnino sint aequivalentes, illud discrimen ita quoque enunciari potest: Observationes complete compensatae omnibus aequationibus conditionalibus $X = 0$, $Y = 0$, $Z = 0$, etc. satisfaciunt, incomplete compensatae vero vel nullis vel saltem non omnibus; compensatio itaque, per quam omnibus aequationibus conditionalibus satisfit, necessario est ipsa completa.

19.

Iam quum ex ipsa notione compensationis sponte sequatur, aggregata duarum compensationum iterum constituere compensationem, facile perspicitur, nihil interesse, utrum praecepta, per quae compensatio perfecta eruenda est, immediate ad observationes primitivas applicentur, an ad observationes incomplete iam compensatas.

Revera constituant $-\theta$, $-\theta'$, $-\theta''$, etc. systema compensationis incompletae, quod prodierit e formulis (I)

$$\begin{aligned}
\theta p &= A^0 a + B^0 b + C^0 c + \text{etc.}\\
\theta' p' &= A^0 a' + B^0 b' + C^0 c' + \text{etc.}\\
\theta'' p'' &= A^0 a'' + B^0 b'' + C^0 c'' + \text{etc.}\\
&\text{etc.}
\end{aligned}$$

Quum observationes his compensationibus mutatae omnibus aequationibus conditionalibus non satisfacere supponantur, sint $\mathcal{A}^*$, $\mathcal{B}^*$, $\mathcal{C}^*$, etc. valores, quos $X$, $Y$, $Z$, etc. ex illarum substitutione nanciscuntur. Quarendi sunt numeri $A^*$, $B^*$, $C^*$, etc. aequationibus (II) satisfacientes

$$\begin{aligned}
\mathcal{A}^* &= A^*[aa] + B^*[ab] + C^*[ac] + \text{etc.}\\
\mathcal{B}^* &= A^*[ab] + B^*[bb] + C^*[bc] + \text{etc.}\\
\mathcal{C}^* &= A^*[ac] + B^*[bc] + C^*[cc] + \text{etc.}
\end{aligned}$$

etc., quo facto compensatio completa observationum isto modo mutatarum efficitur per mutationes novas $-\kappa$, $-\kappa'$, $-\kappa''$, etc., ubi $\kappa$, $\kappa'$, $\kappa''$, etc. computandae sunt per formulas (III)

$$\begin{aligned}
\kappa p &= A^* a + B^* b + C^* c + \text{etc.}\\
\kappa' p' &= A^* a' + B^* b' + C^* c' + \text{etc.}\\
\kappa'' p'' &= A^* a'' + B^* b'' + C^* c'' + \text{etc.}
\end{aligned}$$

equations (12) and (13) are equivalent when (10) holds, we may make the following distinction: completely adjusted observations satisfy the all conditional equations $X = 0$, $Y = 0$, $Z = 0$, etc., incompletely adjusted observations satisfy none or at least not all. Thus an adjustment that causes all the conditional equations to be satisfied is necessarily complete.

19.

It immediately follows from the notion of adjustment that the sum of two adjustments is itself an adjustment. It is then easy to see that it makes no difference whether the technique of complete adjustment is applied to the original observations or to observations that have already been incompletely adjusted.

In fact, let $-\theta$, $-\theta'$, $-\theta''$, etc. form a system of adjustments that comes from the formula (I)

$$\begin{aligned}
\theta p &= A^0 a + B^0 b + C^0 c + \text{etc.}\\
\theta' p' &= A^0 a' + B^0 b' + C^0 c' + \text{etc.}\\
\theta'' p'' &= A^0 a'' + B^0 b'' + C^0 c'' + \text{etc.}\\
&\text{etc.}
\end{aligned}$$

Since the observations changed by these adjustments do not satisfy all the conditional equations, let $\mathcal{A}^*$, $\mathcal{B}^*$, $\mathcal{C}^*$, etc. be the values of the functions $X$, $Y$, $Z$, etc. at the adjusted observations. We must find numbers $A^*$, $B^*$, $C^*$, etc. that satisfy the equations (II)

$$\begin{aligned}
\mathcal{A}^* &= A^*[aa] + B^*[ab] + C^*[ac] + \text{etc.}\\
\mathcal{B}^* &= A^*[ab] + B^*[bb] + C^*[bc] + \text{etc.}\\
\mathcal{C}^* &= A^*[ac] + B^*[bc] + C^*[cc] + \text{etc.}
\end{aligned}$$

etc. Given these quantities, a complete adjustment of the altered observations may be effected by means of new alterations $-\kappa$, $-\kappa'$, $-\kappa''$, etc., where $\kappa$, $\kappa'$, $\kappa''$, etc. are computed from (III)

$$\begin{aligned}
\kappa p &= A^* a + B^* b + C^* c + \text{etc.}\\
\kappa' p' &= A^* a' + B^* b' + C^* c' + \text{etc.}\\
\kappa'' p'' &= A^* a'' + B^* b'' + C^* c'' + \text{etc.}
\end{aligned}$$

etc.

etc. Iam inquiramus, quomodo hae correctiones cum compensatione completa observationum primitivarum cohaereant. Primo manifestum est, haberi

$$\begin{aligned}\mathcal{A}^* &= \mathcal{A} - a\theta - a'\theta' - a''\theta'' - \text{etc.}\\ \mathcal{B}^* &= \mathcal{B} - b\theta - b'\theta' - b''\theta'' - \text{etc.}\\ \mathcal{C}^* &= \mathcal{C} - c\theta - c'\theta' - c''\theta'' - \text{etc.}\end{aligned}$$

etc. Substituendo in his aequationibus pro $\theta$, $\theta'$, $\theta''$, etc. valores ex I, nec non pro $\mathcal{A}^*$, $\mathcal{B}^*$, $\mathcal{C}^*$, etc. valores ex II, invenimus

$$\begin{aligned}\mathcal{A} &= (A^0 + A^*)[aa] + (B^0 + B^*)[ab] + (C^0 + C^*)[ac] + \text{etc.}\\ \mathcal{B} &= (A^0 + A^*)[ab] + (B^0 + B^*)[bb] + (C^0 + C^*)[bc] + \text{etc.}\\ \mathcal{C} &= (A^0 + A^*)[ac] + (B^0 + B^*)[bc] + (C^0 + C^*)[cc] + \text{etc.}\end{aligned}$$

etc., unde patet, correlata aequationum conditionalium aequationibus (12) satisfacientia esse

$$A = A^0 + A^*, \quad B = B^0 + B^*, \quad C = C^0 + C^*, \quad \text{etc.}$$

Hinc vero aequationes (10), I et III docent, esse

$$\epsilon = \theta + \kappa, \quad \epsilon' = \theta' + \kappa', \quad \epsilon'' = \theta'' + \kappa'', \quad \text{etc.}$$

i.e. compensatio observationum perfecta eadem prodit, sive immediate computetur, sive mediate profiscendo a compensatione manca.

## 20.

Quoties multitudo aequationum conditionalium permagna est, determinatio correlatorum $A$, $B$, $C$, etc. per eliminationem directam tam prolixa evadere potest, ut calculatoris patientia ei impar sit: tunc saepenumero commodum esse poterit, compensationem completam per approximationes successivas adiumento theorematis art. prac. eruere. Distribuantur aequationes conditionales in duas pluresve classes, investigeturque primo compensatio, per quam aequationibus primae classis satisfit, neglectis reliquis. Dein tractentur observationes per hanc compensationem mutatae ita, ut solarum aequationum secundae classis ratio habeatur. Generaliter loquendo applicatio secundi compensationum systematis consensum cum aequationibus primae classis turbabit; quare, si duae tantummodo classes factae sunt, ad aequationes primae classis revertemur, tertiumque systema, quod

Let us now ask what the relation is between these corrections and the complete adjustment of the original observations. First, it is obvious that

$$\begin{aligned}
\mathcal{A}^* &= \mathcal{A} - a\theta - a'\theta' - a''\theta'' - \text{etc.}\\
\mathcal{B}^* &= \mathcal{B} - b\theta - b'\theta' - b''\theta'' - \text{etc.}\\
\mathcal{C}^* &= \mathcal{C} - c\theta - c'\theta' - c''\theta'' - \text{etc.}
\end{aligned}$$

etc. If we replace $\theta$, $\theta'$, $\theta''$, etc. in these equations by the values in I and replace $\mathcal{A}^*$, $\mathcal{B}^*$, $\mathcal{C}^*$, etc. by their values in II, we find that

$$\begin{aligned}
\mathcal{A} &= (A^0 + A^*)[aa] + (B^0 + B^*)[ab] + (C^0 + C^*)[ac] + \text{etc.}\\
\mathcal{B} &= (A^0 + A^*)[ab] + (B^0 + B^*)[bb] + (C^0 + C^*)[bc] + \text{etc.}\\
\mathcal{C} &= (A^0 + A^*)[ac] + (B^0 + B^*)[bc] + (C^0 + C^*)[cc] + \text{etc.}
\end{aligned}$$

etc. From this it is clear that the correlates of the conditional equations that satisfy the equations (12) are

$$A = A^0 + A^*, \quad B = B^0 + B^*, \quad C = C^0 + C^*, \quad \text{etc.}$$

Hence the equations (10), I, and III give

$$\epsilon = \theta + \kappa, \quad \epsilon' = \theta' + \kappa', \quad \epsilon'' = \theta'' + \kappa'', \quad \text{etc.};$$

i.e., a perfect adjustment of the observations gives the same results whether it is computed directly or computed indirectly starting from a deficient adjustment.

20.

When the number of conditional equations is very large, the calculation of the correlates $A$, $B$, $C$, etc. by direct elimination becomes so laborious that the endurance of the calculator is not equal to the task. In such cases, it is often better to use the theory of the preceding section to compute the complete adjustment by successive approximations.

The conditional equations are distributed into two or more classes and an adjustment that causes the first class to be satisfied is determined, the other classes being ignored. Then the observations altered by this adjustment are treated as if only the second class counted. Generally speaking, the application of this second adjustment will disturb the agreement with the first system. Hence, if there are only two classes, we return to the first class and calculate a third system huic satisfaciat, eruemus; dein observationes ter correctas compensationi quartae subiiciemus, ubi solae aequationes secundae classis respiciuntur. Ita alternis vicibus, modo priorem classem modo posteriorem respicientes, compensationes continuo decrescentes obtinebimus, et si distributio scite adornata fuerat, post paucas iterationes ad numeros stabiles perveniemus. Si plures quam duae classes factae sund, res simile modo se habebit: classes singulae deinceps in computum venient, post ultimam iterum prima et sic porro. Sed sufficiat hoc loco, hunc modum addigitavisse, cuius efficacia multum utique a scita applicatione pendebit.

21.

Restat, ut suppleamus demonstrationem lemmatis in art. 8 suppositi, ubi tamen perspicuitatis caussa alias denotationes huic negotio magis adaptatas adhibebimus.

Sint itaque $x^0$, $x'$, $x''$, $x'''$, etc. indeterminatae, supponamusque, ex aequationibus

$$\begin{aligned}
&n^{00}x^0 + n^{01}x' + n^{02}x'' + n^{03}x''' + \text{etc.} = X^0\\
&n^{10}x^0 + n^{11}x' + n^{12}x'' + n^{13}x''' + \text{etc.} = X'\\
&n^{20}x^0 + n^{21}x' + n^{22}x'' + n^{23}x''' + \text{etc.} = X''\\
&n^{30}x^0 + n^{31}x' + n^{32}x'' + n^{33}x''' + \text{etc.} = X'''\\
&\text{etc.}
\end{aligned}$$

sequi per eliminationem has

$$\begin{aligned}
&N^{00}X^0 + N^{01}X' + N^{02}X'' + N^{03}X''' + \text{etc.} = x^0\\
&N^{10}X^0 + N^{11}X' + N^{12}X'' + N^{13}X''' + \text{etc.} = x'\\
&N^{20}X^0 + N^{21}X' + N^{22}X'' + N^{23}X''' + \text{etc.} = x''\\
&N^{30}X^0 + N^{31}X' + N^{32}X'' + N^{33}X''' + \text{etc.} = x'''\\
&\text{etc.}
\end{aligned}$$

Substitutis itaque in aequatione prima et secunda secundi systematis valoribus quantitatum $X$, $X'$, $X''$, $X'''$, etc. e primo systemate, obtinemus

$$\begin{aligned}
x^0 = \quad & N^{00}(n^{00}x^0 + n^{01}x' + n^{02}x'' + n^{03}x''' + \text{etc.})\\
&+ N^{01}(n^{10}x^0 + n^{11}x' + n^{12}x'' + n^{13}x''' + \text{etc.})\\
&+ N^{02}(n^{20}x^0 + n^{21}x' + n^{22}x'' + n^{23}x''' + \text{etc.})\\
&+ N^{03}(n^{30}x^0 + n^{31}x' + n^{32}x'' + n^{33}x''' + \text{etc.})\ \text{etc.}
\end{aligned}$$

nec non

which satisfies it. Then we subject the thrice-adjusted observations to a fourth adjustment with respect to the second class. Alternating in this way between the first and second classes, we obtain continually decreasing adjustments. If the initial distribution into classes is done skillfully, the numbers will settle down after a few iterations.

The process is similar if there are more than two classes: the classes are treated one at a time in order, the first following the last, and so on. But the effectiveness of this method depends on its skillful implementation, and it suffices for now simply to mention it.

## 21.

It remains to prove the lemma assumed in Art. 8. For clarity we will use a different, more suitable, notation.

Let $x^0$, $x'$, $x''$, $x'''$, etc. be variables and assume that from the equations

$$\begin{aligned}
&n^{00}x^0 + n^{01}x' + n^{02}x'' + n^{03}x''' + \text{etc.} = X^0\\
&n^{10}x^0 + n^{11}x' + n^{12}x'' + n^{13}x''' + \text{etc.} = X'\\
&n^{20}x^0 + n^{21}x' + n^{22}x'' + n^{23}x''' + \text{etc.} = X''\\
&n^{30}x^0 + n^{31}x' + n^{32}x'' + n^{33}x''' + \text{etc.} = X'''\\
&\text{etc.}
\end{aligned}$$

we obtain by elimination the equations

$$\begin{aligned}
&N^{00}X^0 + N^{01}X' + N^{02}X'' + N^{03}X''' + \text{etc.} = x^0\\
&N^{10}X^0 + N^{11}X' + N^{12}X'' + N^{13}X''' + \text{etc.} = x'\\
&N^{20}X^0 + N^{21}X' + N^{22}X'' + N^{23}X''' + \text{etc.} = x''\\
&N^{30}X^0 + N^{31}X' + N^{32}X'' + N^{33}X''' + \text{etc.} = x'''\\
&\text{etc.}
\end{aligned}$$

If we substitute $X$, $X'$, $X''$, $X'''$, etc. from the first system of equation into the first and second equations of the second system, we get

$$\begin{aligned}
x^0 = \quad & N^{00}(n^{00}x^0 + n^{01}x' + n^{02}x'' + n^{03}x''' + \text{etc.})\\
&+ N^{01}(n^{10}x^0 + n^{11}x' + n^{12}x'' + n^{13}x''' + \text{etc.})\\
&+ N^{02}(n^{20}x^0 + n^{21}x' + n^{22}x'' + n^{23}x''' + \text{etc.})\\
&+ N^{03}(n^{30}x^0 + n^{31}x' + n^{32}x'' + n^{33}x''' + \text{etc.})\ \text{etc.}
\end{aligned}$$

$$\begin{aligned}x' = \quad & N^{10}(n^{00}x^0 + n^{01}x' + n^{02}x'' + n^{03}x''' + \text{etc.})\\ & + N^{11}(n^{10}x^0 + n^{11}x' + n^{12}x'' + n^{13}x''' + \text{etc.})\\ & + N^{12}(n^{20}x^0 + n^{21}x' + n^{22}x'' + n^{23}x''' + \text{etc.})\\ & + N^{13}(n^{30}x^0 + n^{31}x' + n^{32}x'' + n^{33}x''' + \text{etc.}) \text{ etc.}\end{aligned}$$

Quum utraque aequatio manifesto esse debeat aequatio identica, tum in priore tum in posteriore pro $x^0$, $x'$, $x''$, $x'''$, etc. valores quoslibet determinatos substituere licebit. Substituamus in priore

$$x^0 = N^{10}, \quad x' = N^{11}, \quad x'' = N^{12}, \quad x''' = N^{13}, \quad \text{etc.},$$

in posteriore vero

$$x^0 = N^{00}, \quad x' = N^{01}, \quad x'' = N^{02}, \quad x''' = N^{03}, \quad \text{etc.}$$

His ita factis subtractio producit

$$\begin{aligned}N^{10} - N^{01} = \quad & (N^{00}N^{11} - N^{10}N^{01})(n^{01} - n^{10})\\ & + (N^{00}N^{12} - N^{10}N^{02})(n^{02} - n^{20})\\ & + (N^{00}N^{13} - N^{10}N^{03})(n^{03} - n^{30})\\ & + \text{etc.}\\ & + (N^{01}N^{12} - N^{11}N^{02})(n^{12} - n^{21})\\ & + (N^{01}N^{13} - N^{11}N^{03})(n^{13} - n^{31})\\ & + \text{etc.}\\ & + (N^{02}N^{13} - N^{12}N^{03})(n^{23} - n^{32})\\ & + \text{etc. etc.}\end{aligned}$$

quae aequatio ita quoque exhiberi potest

$$N^{10} - N^{01} = \Sigma(N^{0\alpha}N^{1\beta} - N^{1\alpha}N^{0\beta})(n^{\alpha\beta} - n^{\beta\alpha})$$

denotantibus $\alpha\beta$ omnes combinationes indicum inaequalium.

Hinc colligitur, si fuerit

$$n^{01} = n^{10}, \quad n^{02} = n^{20}, \quad n^{03} = n^{30}, \quad n^{12} = n^{21}, \quad n^{13} = n^{31}, \quad n^{23} = n^{32}, \quad \text{etc.}$$

sivi generaliter $n^{\alpha\beta} = n^{\beta\alpha}$, fore etiam

$$N^{10} = N^{01}$$

Et quum ordo indeterminatarum in aequationibus propositis sit arbitrarius, manifest in illa suppositione erit generaliter

$$N^{\alpha\beta} = N^{\beta\alpha}$$

and

$$
\begin{aligned}
x' = \quad & N^{10}(n^{00}x^0 + n^{01}x' + n^{02}x'' + n^{03}x''' + \text{etc.}) \\
& + N^{11}(n^{10}x^0 + n^{11}x' + n^{12}x'' + n^{13}x''' + \text{etc.}) \\
& + N^{12}(n^{20}x^0 + n^{21}x' + n^{22}x'' + n^{23}x''' + \text{etc.}) \\
& + N^{13}(n^{30}x^0 + n^{31}x' + n^{32}x'' + n^{33}x''' + \text{etc.}) \text{ etc.}
\end{aligned}
$$

Both equations are clearly identities. Hence we may replace $x^0$, $x'$, $x''$, $x'''$, etc. by any values in the first equation, and by different values in the second. In the first equation let

$$x^0 = N^{10}, \quad x' = N^{11}, \quad x'' = N^{12}, \quad x''' = N^{13}, \quad \text{etc.,}$$

and in the second equation let

$$x^0 = N^{00}, \quad x' = N^{01}, \quad x'' = N^{02}, \quad x''' = N^{03}, \ \text{etc.}$$

If we subtract these two equations we get

$$
\begin{aligned}
N^{10} - N^{01} = \quad & (N^{00}N^{11} - N^{10}N^{01})(n^{01} - n^{10}) \\
& + (N^{00}N^{12} - N^{10}N^{02})(n^{02} - n^{20}) \\
& + (N^{00}N^{13} - N^{10}N^{03})(n^{03} - n^{30}) \\
& + \text{etc.} \\
& + (N^{01}N^{12} - N^{11}N^{02})(n^{12} - n^{21}) \\
& + (N^{01}N^{13} - N^{11}N^{03})(n^{13} - n^{31}) \\
& + \text{etc.} \\
& + (N^{02}N^{13} - N^{12}N^{03})(n^{23} - n^{32}) \\
& + \text{etc. etc.}
\end{aligned}
$$

This last equation can be written

$$N^{10} - N^{01} = \Sigma(N^{0\alpha}N^{1\beta} - N^{1\alpha}N^{0\beta})(n^{\alpha\beta} - n^{\beta\alpha}),$$

where $\alpha\beta$ denotes all combinations of unequal indices.

From this we see that if

$$n^{01} = n^{10}, \quad n^{02} = n^{20}, \quad n^{03} = n^{30}, \quad n^{12} = n^{21}, \quad n^{13} = n^{31}, \quad n^{23} = n^{32}, \quad \text{etc.}$$

or more generally $n^{\alpha\beta} = n^{\beta\alpha}$, then

$$N^{10} = N^{01}.$$

Since the order of the unknowns in the above equations is arbitrary, we have in general

$$N^{\alpha\beta} = N^{\beta\alpha}.$$

## 22.

Quum methodus in hac commentatione exposita applicationem imprimis frequentem et commodam inveniat in calculis ad geodaesiam sublimiorem pertinentibus, lectoribus gratam fore speramus illustrationem praeceptorum per nonnulla exempla hinc desumta.

Aequationes conditionales inter angulos systematis triangulorum e triplici potissimum fonte sunt petendae.

I. Aggregatum angulorum horizontalium, qui circa eundem verticem gyrum integrum horizontis complent, aequare debet quatuor rectos.

II. Summa trium angulorum in quovis triangulo quantitati datae aequalis est, quum, quoties trangulum est in superficie curva, excessum illius summae supra duos rectos tam accurate computare liceat, ut pro ablosute exacto haberi possit.

III. Fons tertius est ratio laterum in triangulis catenam clausam formantibus. Scilicet si series triangulorum ita nexa est, ut secundum triangulum habeat latus unum $a$ commune cum triangulo primo, aliud $b$ cum tertio; perinde quartum triangulum cum tertio habeat latus commune $c$, cum quinto latus commune $d$, et sic porro usque ad ultimum triangulum, cui cum praecedenti latus commune sit $k$, et cum triangulo primo rursus latus $l$, valores quotientium $\frac{a}{l}$, $\frac{b}{a}$, $\frac{c}{b}$, $\frac{d}{c}$, $\ldots$, $\frac{l}{k}$, innotescent resp. e binis angulis triangulorum successivorum, lateribus communibus oppositis, per methodos notas, unde quum productum illarum fractionum fieri debeat $= 1$, prodibit aequatio conditionalis inter sinus illorum angulorum (parte tertia excessus sphaerici vel sphaeroidici, si triangula sunt in superficie curva, resp. diminutorum).

Ceterum in systematibus triangulorum complicatioribus saepissime accidit, ut aequationes conditionales tum secundi tum tertii generis plures se offerant, quam retinere fas est, quoniam pars earum in reliquis iam contenta est. Contra rarior erit casus, ubi aequationibus conditionalibus secundi generis adiungere oportet aequationes similes ad figuras plurium laterum spectantes, puta tunc tantum, ubi polygona formantur, in triangula per mensurationes non divisa. Sed de his rebus ab instituto praesenti nimis alienis, alia occasione fusius agemus. Silentio tamen praeterire non possumus monitum, quod theoria nostra, si applicatio pura atque rigorosa in votis est, supponit, quantitates per $v$, $v'$, $v''$, etc. designatas revera vel immediate observatas esse, vel ex observationibus ita derivatas, ut inter se inde-

22.

Since the method described in this memoir finds its most frequent and natural application in calculations arising from the higher geodesy, I hope my readers will be grateful if I illustrate the principles with some examples drawn from this subject.

There are three principle sources of conditional equations among the angles of systems of triangles.

I. The sum of horizontal angles that make a complete circuit about the same vertex must be equal to four right angles.

II. Even when a triangle lies in a curved surface, the excess of the sum of its angles over two right angles can be so accurately computed that it is effectively exact. Hence the sum of the three angles in any given triangle is equal to a known quantity.

III. The third source is from the ratio of sides of triangles that form a closed chain. Specifically, let a series of triangles be connected so that the second triangle has the side $a$ in common with the first and side $b$ in common with the third. Let the fourth triangle have the side $c$ in common with the third and $d$ with the fifth, and so on. Finally, let the last triangle have side $k$ in common with its predecessor and side $l$ in common with the first. By a well-known formula, the ratios $\frac{a}{l}$, $\frac{b}{a}$, $\frac{c}{b}$, $\frac{d}{c}, \ldots, \frac{l}{k}$ are determined respectively by the two angles lying opposite the common sides of each triangle. The fact that the product of the quotients is one gives a conditional equation among the sines of the angles (each of which must be reduced by one-third the excess due to sphericity or spheroidicity if they lie in a curved surface).

Of course, in very complicated systems of triangles there are more conditional equations of the second and third kinds than can properly be used, since some of them depend on the others. On the other hand, there will be infrequent cases where one must augment the conditional equations of the second kind with analogous equations for figures with many sides. This occurs whenever there are polygons that are not divided into triangles by the observations. But these matters are rather far removed from the present work, and I will treat them in more detail at another time.

However, there is one problem that we cannot pass over in silence. If we wish to apply our theory in its pure and rigorous form, we must assume that the quantities $v$, $v'$, $v''$, etc. are either observed directly or have been derived from pendentes maneant, vel saltem tales censeri possint. In praxi vulgari observantur anguli triangulorum ipse, qui proin pro $v$, $v'$, $v''$, etc. accipi possunt; sed memores esse debemus, si forte systema insuper contineat triangula talia, quorum anguli non sint immediate observati, sed prodeant tamquam summae vel differentiae angulorum revera observatorum, illos non inter observatorum numerum referendos, sed in forma compositionis suae in calculis retinendos esse. Aliter vero res se habebit in modo observandi ei simili, quem sequutus est clar. STRUVE (Astronomische Nachrichten II, p. 431), ubi directiones singulorum laterum ab eodem vertice proficiscentium obtinentur per comparationem cum una eademque directione arbitraria. Tunc scilicet hi ipsi anguli pro $v$, $v'$, $v''$, etc. accipiendi sunt, quo pacto omnes anguli triangulorum in forma differentiarum se offerent, aequationesque conditionales primi generis, quibus per rei naturam sponte satisfit, tamquam superfluae cessabunt. Modus observationis, quem ipse sequutus sum in dimensione triangulorum annis praecedentibus perfecta, differt quidem tum a priore tum a posteriore mode, attamen respectu effectus posteriori aequiparari potest, ita ut in singulis stationibus directiones laterum inde profiscentium ab inition quasi arbitrario numeratas pro quantitatibus $v$, $v'$, $v''$, etc. accipere oporteat. Duo iam exempla elaborabimus, alterum ad modum priorem, alterum ad posteriorem pertinens.

## 23.

Exemplum primum nobis suppeditabit opus clar. DE KRAYENHOF, *Précis historique des opérations trigonométriques faites en Hollande*, et quidem compensationi subiiciemus partem eam systematis triangulorum, quae inter novem puncta Harlingen, Sneek, Oldeholtpade, Ballum, Leeuwarden, Dockum, Drachten, Oosterwolde, Gröningen coninentur. Formantur inter haec puncta novem triangula in opere illo per numeros 121, 122, 123, 124, 125, 127, 128, 131, 132 denotata, quorum anguli (a nobis indicibus praescriptis distincti) secundam tabulam p. 77–81 ita sunt observati:

Triangulum 121.

| | | |
|---|---|---|
| 0. | Harlingen | $50^\circ\ 58'\ 15.238''$ |
| 1. | Leeuwarden | $82^\circ\ 47'\ 15.351''$ |
| 2. | Ballum | $46^\circ\ 14'\ 27.202''$ |

observations in such a way that they remain independent—or at least can be considered so. Usually one observes the angles of the triangles themselves, which can then be taken for $v$, $v'$, $v''$, etc. But whenever the system contains triangles whose angles have not been observed directly, but are sums or differences of angles that have been observed, these angles should not be counted among the observed angles. Instead they must be kept in their composite forms in the calculation.

Things are different, however, for a way observing similar to one followed by STRUVE (*Astronomische Nachtrichten* II, p. 431), in which the directions of individual sides proceeding from the same vertex are obtained relative to one of the directions, arbitrarily chosen, and these angles are then taken as the quantities $v$, $v'$, $v''$, etc. In this case the angles are differences, and the conditional equations of the first kind are superfluous, since they are automatically satisfied.

The method of observing triangles I have followed over the past years differs from both the first and second methods. Yet it is effectively like the second, in that the directions of sides proceeding from an individual station, reckoned from an arbitrary origin, are taken as the quantities $v$, $v'$, $v''$, etc. We will now work out two examples, one of which pertains to the first method and the other to the second.

## 23.

Our first example is from the work of DE KRAYENHOF, *Précis historique des opérations trigonométriques faites en Hollande.* We will adjust the part of the system of triangles that are contained between the nine points Harlingen, Sneek, Oldeholtpade, Ballum, Leeuwarden, Dockum, Drachten, Oosterwolde, Gröningen. These points form nine triangles which are denoted in DE KRAYENHOF's work by the numbers 121, 122, 123, 124, 125, 127, 128, 131, 132. According to the table on pp. 77–81, the angles (here prefixed by distinguishing numbers and presented in degrees–minutes–seconds) were measured as follows.

Triangle 121.

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

Triangulum 122.

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

Triangulum 123.

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

Triangulum 124.

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

Triangulum 125.

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

Triangulum 127.

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

Triangulum 128.

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

Triangulum 131.

| | | |
|---|---|---|
| 21. | Dockum | 57° 1′ 55.292″ |
| 22. | Drachten | 83° 33′ 14.515″ |
| 23. | Gröingen | 39° 24′ 52.397″ |

Triangulum 132.

| | | |
|---|---|---|
| 24. | Oosterwolde | 81° 54′ 17.447″ |
| 25. | Gröningen | 31° 52′ 46.094″ |
| 26. | Drachten | 66° 12′ 57.246″ |

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. | Gröingen | 39° 24′ 52.397″ |

Triangle 132.

| | | |
|---|---|---|
| 24. | Oosterwolde | 81° 54′ 17.447″ |
| 25. | Gröningen | 31° 52′ 46.094″ |
| 26. | Drachten | 66° 12′ 57.246″ |

Consideratio nexus inter haec triangula monstrat, inter 27 angulos, quorum valores approximati per observationem innotuerunt, 13 aequationes conditionales haberi, puta duas primi generis, novem secundi, duas tertii. Sed haud opus erit, has aequationes omnes in forma sua finita hic adscribere, quum ad calculos tantummodo requirantur quantitates in theoria generali per $\mathcal{A}$, $a$, $a'$, $a''$, etc., $\mathcal{B}$, $b$, $b'$, $b''$, etc. etc. denotatae: quare illarum loco statim adscribimus aequationes supra per (13) denotatas, quae illas quantitates ob oculos ponunt: loco signorum $\epsilon$, $\epsilon'$, $\epsilon''$, etc. simpliciter hic scribemus (0), (1), (2), etc.

Hoc modo duabus aequationibus conditionalibus primi generis respondent sequentes:

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

Excessus sphaeroidicos novem triangulorum invenimus deinceps: 1.749″,1.147″, 1.243″, 1.698″, 0.873″, 1.167″, 1.104″, 2.161″, 1.403″. Oritur itaque aequatio conditionalis secundi generis prima haec:*) $v^{(0)} + v^{(1)} + v^{(2)} - 180^\circ\, 0'\, 1.749'' = 0$, et perinde reliquae: hinc habemus novem aequationes sequentes:

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

Aequationes conditionales tertii generis commodius in forma logarithmica exhibentur: ita prior est

$$\begin{aligned}
&\log\sin(v^{(0)}-0.583'') - \log\sin(v^{(2)}-0.583'') - \log\sin(v^{(3)}-0.382'')\\
&+\log\sin(v^{(4)}-0.382'') - \log\sin(v^{(6)}-0.414'') + \log\sin(v^{(7)}-0\ .414'')\\
&-\log\sin(v^{(16)}-0.389'') + \log\sin(v^{(17)}-0.389'') - \log\sin(v^{(19)}-0.368'')\\
&+\log\sin(v^{(20)}-0.368'') = 0.
\end{aligned}$$

*)Indices in hoc exemplo per figuras arabicas exprimere praeferimus.

A consideration of the relation between these triangles shows that among the 27 angles whose approximate values have been observed there are 13 conditional equations: two of the first kind, nine of the second, and two of the third. However, there is no need to write down all these equations in closed form, since for calculation we only need the quantities denoted in our general theory by $\mathcal{A}$, $a$, $a'$, $a''$, etc., $\mathcal{B}$, $b$, $b'$, $b''$, etc. etc. For this reason we will write down instead the equations numbered (13) above, which exhibit these quantities explicitly. Instead of the variables $\epsilon$, $\epsilon'$, $\epsilon''$, etc. we will simply write (0), (1), (2), etc.

The two conditional equations of the first kind correspond to the following equations:

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

The spheroidal excesses of the nine triangles are respectively 1.749″, 1.147″, 1.243″, 1.698″, 0.873″, 1.167″, 1.104″, 2.161″, 1.403″. From the first we get a conditional equation of the second kind,*) $v^{(0)} + v^{(1)} + v^{(2)} - 180^\circ\, 0'\, 1.749'' = 0$, and likewise for the rest. Thus we have the following nine equations.

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

The conditional equations of the third kind are best exhibited in logarithmic form. The first is

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

*) In this example we prefer to represent indices by Arabic numerals.

Superfluum videtur, alteram in forma finita adscribere. His duabus aequationibus respondent sequentes, ubi singuli coëfficientes referuntur ad figuram septimam logarithorum briggicorum:

$$\begin{aligned}
&17.068(0) - 20.174(2) + 16.993(3) + 7.328(4) - 17.976(6) + 22.672(7)\\
&\quad - 5.028(16) + 21.780(17) - 19.710(19) = 11.671(20) = -371,\\
&17.976(6) - 0.880(8) - 20.617(9) + 8.564(10) - 19.082(13) + 4.375(14)\\
&\quad + 6.798(18) - 11.671(20) + 13.657(21) - 25.620(23) - 2.995(24)\\
&\quad + 33.854(25) = +370.
\end{aligned}$$

Quum nulla ratio indicata sit, cur observationibus pondera inaequalia tribuamus, statuemus $p^{(0)} = p^{(1)} = p^{(2)}$ etc. $= 1$. Denotatis itaque correlatis aequationum conditionalium eo ordine, quo aequationes ipsis respondentes exhibuimus, per $A$, $B$, $C$, $D$, $E$, $F$, $G$, $H$, $I$, $K$, $L$, $M$, $N$, prodeunt ad illorum determinationem aequationes sequentes:

$$\begin{aligned}
-2.197 &= 5A + C + D + E + h + I + 5.917N\\
-0.436 &= 6B + E + F + G + I + K + L + 2.962M\\
-3.958 &= A + 3C - 3.106M\\
+0.722 &= A + 3C - 9.665M\\
-0.753 &= A + B + 3E + 4.696M + 17.096N\\
+2.355 &= B + 3F + 12.053N\\
-1.201 &= B + 3G - 14.707N\\
-0.461 &= A + 3H + 16.752N\\
+2.596 &= A + B + 3I - 8.039M - 4.874N\\
+0.043 &= B + 3K - 11.963N\\
-0.616 &= B + 3L - 30.859N\\
-371 &= +2.962B - 3.106C - 9.665D + 4.696E + 16.752H - 8.039I\\
&\quad +2902.27M - 459.33N\\
+370 &= +5.917A + 17.096E - 12.053F - 14.707G - 4.874I\\
&\quad -11.963K + 30.859L - 459.33M + 3385.96N
\end{aligned}$$

It is superfluous to write down the closed form of the other. To these two equations there correspond the following equations, in which the individual coefficients are from seven place common logarithms.

$$\begin{aligned}17.068(0) &- 20.174(2) + 16.993(3) + 7.328(4) - 17.976(6) + 22.672(7)\\ &- 5.028(16) + 21.780(17) - 19.710(19) = 11.671(20) = -371,\\ 17.976(6) &- 0.880(8) - 20.617(9) + 8.564(10) - 19.082(13) + 4.375(14)\\ &+ 6.798(18) - 11.671(20) + 13.657(21) - 25.620(23) - 2.995(24)\\ &+ 33.854(25) = +370.\end{aligned}$$

Since there is no reason to think that we should assign unequal weights to the observations, we will set $p^{(0)} = p^{(1)} = p^{(2)}$ etc. $= 1$. Thus if we denote the correlates of the conditional equations (in the same order as the equations themselves were written down) by $A$, $B$, $C$, $D$, $E$, $F$, $G$, $H$, $I$, $K$, $L$, $M$, $N$, they may be calculated from the following equations.

$$\begin{aligned}
-2.197 &= 5A + C + D + E + h + I + 5.917N\\
-0.436 &= 6B + E + F + G + I + K + L + 2.962M\\
-3.958 &= A + 3C - 3.106M\\
+0.722 &= A + 3C - 9.665M\\
-0.753 &= A + B + 3E + 4.696M + 17.096N\\
+2.355 &= B + 3F + 12.053N\\
-1.201 &= B + 3G - 14.707N\\
-0.461 &= A + 3H + 16.752N\\
+2.596 &= A + B + 3I - 8.039M - 4.874N\\
+0.043 &= B + 3K - 11.963N\\
-0.616 &= B + 3L - 30.859N\\
-371 &= +2.962B - 3.106C - 9.665D + 4.696E + 16.752H - 8.039I\\
&\quad +2902.27M - 459.33N\\
+370 &= +5.917A + 17.096E - 12.053F - 14.707G - 4.874I\\
&\quad -11.963K + 30.859L - 459.33M + 3385.96N
\end{aligned}$$

Hinc eruimus per eliminationem:

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

Denique errores maxime plausible prodeunt per formulas

$$
\begin{aligned}
(0) &= C + 17.068M \\
(1) &= A + C \\
(2) &= C - 20.174M \\
(3) &= D - 16.993M
\end{aligned}
$$

etc., unde obtinemus valores numericos sequentes; in gratiam comparationis apponimus (mutatis signis) correctiones a clar. DE KRAYENHOF observationibus applicatas:

| | DE KR. | | DE KR. |
|---|---|---|---|
| $(0) = -3.108$ | $-2.090$ | $(14) = +0.795$ | $+2.400$ |
| $(1) = -1.832$ | $+0.116$ | $(15) = +0.061$ | $+1.273$ |
| $(2) = +0.981$ | $-1.982$ | $(16) = +1.211$ | $+5.945$ |
| $(3) = +1.952$ | $+1.722$ | $(17) = -1.732$ | $-7.674$ |
| $(4) = -0.719$ | $+2.848$ | $(18) = +1.265$ | $+1.876$ |
| $(5) = -0.512$ | $-3.848$ | $(19) = +2.959$ | $+6.251$ |
| $(6) = +3.648$ | $-0.137$ | $(20) = -1.628$ | $-5.530$ |
| $(7) = -3.221$ | $+1.000$ | $(21) = +2.211$ | $-3.486$ |
| $(8) = -1.180$ | $-1.614$ | $(22) = +0.322$ | $-3.454$ |
| $(9) = -1.116$ | $0.000$ | $(23) = -2.489$ | $0.000$ |
| $(10) = +2.376$ | $+5.928$ | $(24) = -1.709$ | $+0.400$ |
| $(11) = +1.096$ | $-3.570$ | $(25) = +2.701$ | $+2.054$ |
| $(12) = +0.016$ | $+2.414$ | $(26) = -1.606$ | $-3.077$ |
| $(13) = -2.013$ | $-6.014$ | | |

Aggregatum quadratorum nostrarum compensationum invenitur = 97.8845. Hinc error medius, quatenus ex 27 angulis observatis colligi potest,

$$\sqrt{\frac{97.8845}{13}} = 2.7440''$$

By elimination we have

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

The most reliable errors are then given by the formulas

$$\begin{aligned}(0) &= C + 17.068M\\ (1) &= A + C\\ (2) &= C - 20.174M\\ (3) &= D - 16.993M\end{aligned}$$

etc. From these we obtain the following numerical values. For the sake of comparison we append the corrections (with the signs changed) applied by the DE KRAYENHOF to the observations.

| | DE KR. | | DE KR. |
|---|---|---|---|
| (0) = −3.108 | −2.090 | (14) = +0.795 | +2.400 |
| (1) = −1.832 | +0.116 | (15) = +0.061 | +1.273 |
| (2) = +0.981 | −1.982 | (16) = +1.211 | +5.945 |
| (3) = +1.952 | +1.722 | (17) = −1.732 | −7.674 |
| (4) = −0.719 | +2.848 | (18) = +1.265 | +1.876 |
| (5) = −0.512 | −3.848 | (19) = +2.959 | +6.251 |
| (6) = +3.648 | −0.137 | (20) = −1.628 | −5.530 |
| (7) = −3.221 | +1.000 | (21) = +2.211 | −3.486 |
| (8) = −1.180 | −1.614 | (22) = +0.322 | −3.454 |
| (9) = −1.116 | 0.000 | (23) = −2.489 | 0.000 |
| (10) = +2.376 | +5.928 | (24) = −1.709 | +0.400 |
| (11) = +1.096 | −3.570 | (25) = +2.701 | +2.054 |
| (12) = +0.016 | +2.414 | (26) = −1.606 | −3.077 |
| (13) = −2.013 | −6.014 | | |

The sum of squares of our adjustments is 97.8845. Hence the mean error, as far as it can be extracted from the 27 observed angles, is

$$\sqrt{\tfrac{97.8845}{13}} = 2.7440''.$$

Aggregatum quadratorum mutationum, quas clar. DE KRAYENHOF ipse angulis obervatis applicavit, invenitur = 341.4201.

24.

Exemplum alterum suppeditabunt triangula inter quinque puncta triangulationis Hannoveranae, Falkenberg, Breithorn, Hauselberg, Wulfsode, Wilsede. Observatae sunt directiones:*)

In statione *Falkenberg*

| | | |
|---|---|---|
| 0. | Wilsede ......... | 187° 47′ 30.311″ |
| 1. | Wulfsode ........ | 225° 9′ 39.676″ |
| 2. | Hauselberg ...... | 266° 13′ 56.239″ |
| 3. | Breithorn ....... | 274° 14′ 43.634″ |

In statione *Breithorn*

| | | |
|---|---|---|
| 4. | Falkenberg ....... | 94° 33′ 40.755″ |
| 5. | Hauselberg ...... | 122° 51′ 23.054″ |
| 6. | Wilsede ......... | 150° 18′ 35.100″ |

In statione *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″ |

In statione *Wulfsode*

| | | |
|---|---|---|
| 11. | Hauselberg ........ | 9° 5′ 36.593″ |
| 12. | Falkenberg ....... | 45° 27′ 33.556″ |
| 13. | Wilsede ......... | 118° 44′ 13.159″ |

In statione *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″ |

*)Initia, ad quae singulae directiones referuntur, hic tamquam arbitraria considerantur, quamquam revera cum lineis meridianis stationum coincidunt. Observationes in posterum complete publici iuris fient; interim figura invenitur in *Astronomische Nachtrichten* Vol. I. p. 441.

The sum of squares of the changes which DE KRAYENHOF himself applied to the observed angles is 341.4201.

24.

Our second example is supplied by triangles between five points from the triangulation of Hannover: Falkenberg, Breithorn, Hauselberg, Wulfsode, Wilsede. The following directions were observed:*)

From station *Falkenberg*

| | | |
|---|---|---|
| 0. | Wilsede | 187° 47′ 30.311″ |
| 1. | Wulfsode | 225° 9′ 39.676″ |
| 2. | Hauselberg | 266° 13′ 56.239″ |
| 3. | Breithorn | 274° 14′ 43.634″ |

From station *Breithorn*

| | | |
|---|---|---|
| 4. | Falkenberg | 94° 33′ 40.755″ |
| 5. | Hauselberg | 122° 51′ 23.054″ |
| 6. | Wilsede | 150° 18′ 35.100″ |

From station *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″ |

From station *Wulfsode*

| | | |
|---|---|---|
| 11. | Hauselberg | 9° 5′ 36.593″ |
| 12. | Falkenberg | 45° 27′ 33.556″ |
| 13. | Wilsede | 118° 44′ 13.159″ |

From station *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″ |

---

*) The base lines from which individual directions were observed can here be taken as arbitrary, although they actually coincide with the meridian lines of the stations. At a later time the observations will be made public. For the time being, a figure can be found in *Astronomische Nachtrichten*, Vol. I., p. 441.

Ex his observationibus septem triangula formare licet.

Triangulum I.

| | |
|---|---|
| Falkenberg ........ | 8° 0′ 47.395″ |
| Breithorn ........ | 28° 17′ 42.299″ |
| Hauselberg ...... | 143° 41′ 29.140″ |

Triangulum II.

| | |
|---|---|
| Falkenberg ....... | 86° 27′ 13.323″ |
| Breithorn ........ | 55° 44′ 54.345″ |
| Wilsede .......... | 37° 47′ 53.635″ |

Triangulum III.

| | |
|---|---|
| Falkenberg ....... | 41° 4′ 16.563″ |
| Hauselberg ...... | 102° 33′ 49.504″ |
| Wulfsode ......... | 36° 21′ 56.963″ |

Triangulum IV.

| | |
|---|---|
| Falkenberg ....... | 78° 26′ 25.928″ |
| Hauselberg ....... | 68° 8′ 2.752″ |
| Wilsede .......... | 35° 25′ 34.281″ |

Triangulum V.

| | |
|---|---|
| Falkenberg ....... | 37° 22′ 9.365″ |
| Wulfsode ......... | 73° 16′ 39.603″ |
| Wilsede .......... | 69° 21′ 11.508″ |

Triangulum VI.

| | |
|---|---|
| Breithorn ........ | 27° 27′ 12.046″ |
| Hauselberg ...... | 148° 10′ 28.108″ |
| Wilsede ........... | 4° 22′ 19.354″ |

Triangulum VII.

| | |
|---|---|
| Hauselberg ....... | 34° 25′ 46.752″ |
| Wulfsode ........ | 109° 38′ 36.566″ |
| Wilsede .......... | 35° 55′ 37.227″ |

From these observations seven triangles may be formed.

Triangle I.

| | |
|---|---|
| Falkenberg | 8° 0′ 47.395″ |
| Breithorn | 28° 17′ 42.299″ |
| Hauselberg | 143° 41′ 29.140″ |

Triangle II.

| | |
|---|---|
| Falkenberg | 86° 27′ 13.323″ |
| Breithorn | 55° 44′ 54.345″ |
| Wilsede | 37° 47′ 53.635″ |

Triangle III.

| | |
|---|---|
| Falkenberg | 41° 4′ 16.563″ |
| Hauselberg | 102° 33′ 49.504″ |
| Wulfsode | 36° 21′ 56.963″ |

Triangle IV.

| | |
|---|---|
| Falkenberg | 78° 26′ 25.928″ |
| Hauselberg | 68° 8′ 2.752″ |
| Wilsede | 35° 25′ 34.281″ |

Triangle V.

| | |
|---|---|
| Falkenberg | 37° 22′ 9.365″ |
| Wulfsode | 73° 16′ 39.603″ |
| Wilsede | 69° 21′ 11.508″ |

Triangle VI.

| | |
|---|---|
| Breithorn | 27° 27′ 12.046″ |
| Hauselberg | 148° 10′ 28.108″ |
| Wilsede | 4° 22′ 19.354″ |

Triangle VII.

| | |
|---|---|
| Hauselberg | 34° 25′ 46.752″ |
| Wulfsode | 109° 38′ 36.566″ |
| Wilsede | 35° 55′ 37.227″ |

Aderunt itaque septem aequationes conditionales secundi generis (aequationes primi generis manifest cessant), quas ut eruamus, computandi sunt ante omnia excessus sphaeroidici septem triangulorum. Ah hunc finem requiritur cognitio magnitudinis absolutae saltem unius lateris: latus inter puncta Wilsede and Wulfsode est 22, 877.94 metrorum. Hinc prodeunt excessus spaeroidici triangulorum I ... 0.202″; II ... 2.442″; III ... 1.257″; IV ... 1.919″; V ... 1.957″; VI ... 0.321″; VII ... 1.295″.

Iam si directiones eo ordine, quo supra allatae indicibusque distinctae sunt, per $v^{(0)}$, $v^{(1)}$, $v^{(2)}$, $v^{(3)}$, etc. designantur, trianguli I anguli fiunt

$$v^{(3)} - v^{(2)}, \quad v^{(5)} - v^{(4)}, \quad 360^\circ + v^{(7)} - v^{(10)}$$

adeoque aequatio conditionalis prima

$$-v^{(2)} + v^{(3)} - v^{(4)} + v^{(5)} + v^{(7)} - v^{(10)} + 179^\circ\, 59'\, 59.798'' = 0.$$

Perinde triangula reliqua sex alias suppeditant; sed levis attentio docebit, has septem aequationes non esse independentes, sed secundam identicam cum summa primae, quartae et sextae; nec non summam tertiae et quintae identicam cum summa quartae et septimae: quapropter secundam et quintam negligemus. Loco remanentium aequationum conditionalium in forma finita, adscribimus aequationes correspondentes e complexu (13), dum pro characteribus $\epsilon$, $\epsilon'$, etc. his (0), (1), (2) etc., utimur:

$$\begin{aligned}
-1.368 &= -(2) + (3) - (4) \;\; + (5) \;\; + (7) \;\; - (10)\\
+1.773 &= -(1) + (2) - (7) \;\; + (9) \;\; - (11) + (12)\\
+1.042 &= -(0) + (2) - (7) \;\; + (8) \;\; + (14) - (17)\\
-0.813 &= -(5) + (6) - (8) \;\; + (10) - (16) + (17)\\
-0.750 &= -(8) + (9) - (11) + (13) - (15) + (17)
\end{aligned}$$

Aequationes conditionales tertii generis *octo* e trangulorum systemate peti possent, quum tum terna quatuor triangulorum I, II, IV, VI, tum terna ex his III, IV, V, VII ad hunc finem combinare liceat; attamen levis attentio docet, *duas* sufficere, alteram ex illis, alteram ex his, quum reliquae in his atque prioribus aequationibus conditionalibus iam contentae esse debeant. Aequatio itaque conditinalis sexta nobis erit

$$\begin{aligned}
&\log\sin(v^{(3)} - v^{(2)} - 0.067'') \quad - \log\sin(v^{(5)} - v^{(4)} - 0.067'')\\
&+ \log\sin(v^{(14)} - v^{(17)} - 0.640'') - \log\sin(v^{(2)} - v^{(0)} - 0.640'')\\
&+ \log\sin(v^{(6)} - v^{(5)} - 0.107'') \quad - \log\sin(v^{(17)} - v^{(16)} - 0.107'') = 0
\end{aligned}$$

There are seven conditional equations of the second kind (equations of the first kind obviously do not apply here). To derive them, we must first of all compute the spheroidal excesses of the seven triangles. For this we need to know the length of at least one side. The side between the points Wilsede and Wulfsode is 22,877.94 meters. This gives the following spheroidal excesses: I ... 0.202″; II ... 2.442″; III ... 1.257″; IV ... 1.919″; V ... 1.957″; VI ... 0.321″; VII ... 1.295″.

If we now denote the directions, in the same order that they were introduced and numbered above, by $v^{(0)}$, $v^{(1)}$, $v^{(2)}$, $v^{(3)}$, etc., then the angles of triangle I become

$$v^{(3)} - v^{(2)}, \quad v^{(5)} - v^{(4)}, \quad 360^\circ + v^{(7)} - v^{(10)}.$$

Thus the first conditional equation is

$$-v^{(2)} + v^{(3)} - v^{(4)} + v^{(5)} + v^{(7)} - v^{(10)} + 179^\circ\, 59'\, 59.798'' = 0.$$

Similarly the remaining triangles yield six other conditions. But a little reflection shows that these seven equations are not independent: the second is identical to the sum of the second, fourth, and sixth, and the sum of the of the third and the fifth is identical to the sum of the fourth and the seventh. Consequently we will ignore the second and the fifth equations. Instead of closed forms for the remaining conditional equations, we will write down the equations corresponding to the system (13), using (0), (1), (2), etc. for the variables $\epsilon$, $\epsilon'$, etc. These equations are

$$\begin{aligned}
-1.368 &= -(2) + (3) - (4) + (5) + (7) - (10)\\
+1.773 &= -(1) + (2) - (7) + (9) - (11) + (12)\\
+1.042 &= -(0) + (2) - (7) + (8) + (14) - (17)\\
-0.813 &= -(5) + (6) - (8) + (10) - (16) + (17)\\
-0.750 &= -(8) + (9) - (11) + (13) - (15) + (17)
\end{aligned}$$

There are *eight* conditional equations of the third kind which can be obtained from the system of triangles, since any three of the four triangles I, II, IV, VI can be combined for this purpose, as well as any three of the triangles III, IV, V, VII. However, a little reflection shows that only *two* suffice, one from the first group and another from the second, since these subsume the others in their respective groups. Thus our sixth conditional equation will be

$$\begin{aligned}
&\log\sin(v^{(3)} - v^{(2)} - 0.067'') - \log\sin(v^{(5)} - v^{(4)} - 0.067'')\\
&+ \log\sin(v^{(14)} - v^{(17)} - 0.640'') - \log\sin(v^{(2)} - v^{(0)} - 0.640'')\\
&+ \log\sin(v^{(6)} - v^{(5)} - 0.107'') - \log\sin(v^{(17)} - v^{(16)} - 0.107'') = 0
\end{aligned}$$

atque septima

$$\begin{aligned}
&\log\sin(v^{(2)}-v^{(1)}-0.419'') \quad -\log\sin(v^{(12)}-v^{(11)}-0.419'')\\
&+\log\sin(v^{(14)}-v^{(17)}-0.640'')-\log\sin(v^{(2)}-v^{(0)}-0.640'')\\
&+\log\sin(v^{(13)}-v^{(11)}-0.432'')-\log\sin(v^{(17)}-v^{(15)}-0.432'')=0
\end{aligned}$$

quibus respondent aequationes complexus (13)

$$\begin{aligned}
+25=&+\ \ 4.31\,(0)\ -153.88\,(2)\ +149.57\,(3)\ +\ \ 39.11\,(4)\ \ -79.64\,(5)\\
&+40.53\,(6)\ +\ \ 31.90\,(14)+275.39\,(16)-307.29\,(17),\\
-3=&+\ \ 4.31\,(0)\ -\ \ 24.16\,(1)\ +\ \ 19.85\,(2)\ +\ \ 36.11\,(11)\ -28.59\,(12)\\
&-\ \ 7.52\,(13)+\ \ 31.90\,(14)+\ \ 29.06\,(15)-\ \ 60.96\,(17)
\end{aligned}$$

Quodsi iam singulis directionibus eandem certitudinem tribuimus, statuendo $p^{(0)}=p^{(1)}=p^{(2)}$ etc. $=1$, correlaqtaque septem aequationum conditionalium, eo ordine, quem hic sequuti sumus, $A$, $B$, $C$, $D$, $E$, $F$, $G$ denotamus, horum determinatio petenda erit ex aequationibus sequentibus:

$$\begin{aligned}
-1.368&=+6A-2B-2C-2D+184.72F-19.85G\\
+1.773&=-2A+6B+2C+2E-153.88F-20.69G\\
+1.042&=-2A+2B+6C-2D-2E+181.00F+108.40G\\
-0.813&=-2A-2C+6D+2E-462.51F-60.96G\\
-0.750&=+2B-2C+2D+6E-307.29F-133.65G\\
+25\ &=+184.72A-153.88B+181.00C-462.51D-307.29E\\
&\quad+224868F+16694.1G\\
-3\ &=-19.85A-20.69B+108.40C-60.96D-133.65E\\
&\quad+16694.1F+8752.39G
\end{aligned}$$

Hinc deducimus per eliminationem

$$\begin{aligned}
A&=-0.225\\
B&=+0.344\\
C&=-0.088\\
D&=-0.171\\
E&=-0.323\\
F&=+0.000215915\\
G&=-0.00547462
\end{aligned}$$

and our seventh

$$\begin{aligned}&\log\sin(v^{(2)}-v^{(1)}-0.419'')\quad-\log\sin(v^{(12)}-v^{(11)}-0.419'')\\+&\log\sin(v^{(14)}-v^{(17)}-0.640'')-\log\sin(v^{(2)}-v^{(0)}-0.640'')\\+&\log\sin(v^{(13)}-v^{(11)}-0.432'')-\log\sin(v^{(17)}-v^{(15)}-0.432'')=0\end{aligned}$$

for which the corresponding equations (13) are

$$\begin{aligned}+25=&+4.31\,(0)-153.88\,(2)+149.57\,(3)+39.11\,(4)-79.64\,(5)\\&+40.53\,(6)+31.90\,(14)+275.39\,(16)-307.29\,(17),\\-3=&+4.31\,(0)-24.16\,(1)+19.85\,(2)+36.11\,(11)-28.59\,(12)\\&-7.52\,(13)+31.90\,(14)+29.06\,(15)-60.96\,(17).\end{aligned}$$

If we assign the same certainty to each direction, setting $p^{(0)}=p^{(1)}=p^{(2)}$ etc. $=1$, and denote the correlates of the conditional equations in order by $A$, $B$, $C$, $D$, $E$, $F$, $G$, then we must calculate them from the following equations:

$$\begin{aligned}-1.368&=+6A-2B-2C-2D+184.72F-19.85G\\+1.773&=-2A+6B+2C+2E-153.88F-20.69G\\+1.042&=-2A+2B+6C-2D-2E+181.00F+108.40G\\-0.813&=-2A-2C+6D+2E-462.51F-60.96G\\-0.750&=+2B-2C+2D+6E-307.29F-133.65G\\+25\ &=+184.72A-153.88B+181.00C-462.51D-307.29E\\&\quad+224868F+16694.1G\\-3\ &=-19.85A-20.69B+108.40C-60.96D-133.65E\\&\quad+16694.1F+8752.39G\end{aligned}$$

By elimination

$$\begin{aligned}A&=-0.225\\B&=+0.344\\C&=-0.088\\D&=-0.171\\E&=-0.323\\F&=+0.000215915\\G&=-0.00547462\end{aligned}$$

Iam errores maxime plausibiles habentur per formulas:

$$\begin{aligned}(0)&=-C+4.31F+4.31G\\(1)&=-B-24.16G\\(2)&=-A+B+C-153.88F+19.85G\end{aligned}$$

etc., unde prodeunt valores numerici

| | |
|---|---|
| $(0)=+0.065''$ | $(9)=+0.021''$ |
| $(1)=-0.212''$ | $(10)=+0.054''$ |
| $(2)=+0.339''$ | $(11)=-0.219''$ |
| $(3)=-0.193''$ | $(12)=+0.501''$ |
| $(4)=+0.233''$ | $(13)=-0.282''$ |
| $(5)=-0.071''$ | $(14)=-0.256''$ |
| $(6)=-0.162''$ | $(15)=+0.164''$ |
| $(7)=-0.481''$ | $(16)=+0.230''$ |
| $(8)=+0.406''$ | $(17)=-0.139''$ |

Summa quadratorum horum errorum invenitur = 1.2288; hinc error medius unius directionis, quatenus e 18 directionibus observatis erui potest,

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

25.

Ut etiam pars altera theoriae nostrae exemplo illustretur, indagamus praecisiones, qua latus Falkenberg-Breithorn e latere Wilsede-Wulfsode adiumento observationum compensatarum determinatur. Functio $u$, per quam illud in hoc casu exprimitur, est

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

Huius valor e valoribus correctis directionum $v^{(0)}$, $v^{(1)}$, etc. invenitur

$$26766.68\,\text{m.}$$

Differentiatio autem illius expressionis suppeditat, si differentialia $dv^{(0)}$, $dv^{(1)}$, etc. minutis secundis expressa concipiuntur,

$$\begin{aligned}du=\;&0.16991\text{m.}(dv^{(0)}-dv^{(1)})+0.08836\text{m.}(dv^{(4)}-dv^{(6)})\\&-0.03899\text{m.}(dv^{(12)}-dv^{(13)})+0.16731\text{m.}(dv^{14}-dv^{(16)})\end{aligned}$$

Now the most reliable errors are obtained from the formulas

$$\begin{aligned}(0) &= -C + 4.31F + 4.31G\\ (1) &= -B - 24.16G\\ (2) &= -A + B + C - 153.88F + 19.85G\end{aligned}$$

etc., which gives the numerical values

| | |
|---|---|
| $(0) = +0.065''$ | $(9) = +0.021''$ |
| $(1) = -0.212''$ | $(10) = +0.054''$ |
| $(2) = +0.339''$ | $(11) = -0.219''$ |
| $(3) = -0.193''$ | $(12) = +0.501''$ |
| $(4) = +0.233''$ | $(13) = -0.282''$ |
| $(5) = -0.071''$ | $(14) = -0.256''$ |
| $(6) = -0.162''$ | $(15) = +0.164''$ |
| $(7) = -0.481''$ | $(16) = +0.230''$ |
| $(8) = +0.406''$ | $(17) = -0.139''$ |

The sum of the squares of these errors is 1.2288. Hence, as far as it can be derived from 18 observed directions, the mean error in any one direction is

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

25.

In order to illustrate the other part of our theory, let us investigate the precision with which the side Falkenberg–Breithorn is determined from the side Wilsede–Wulfsode by means of the adjusted observations. For this case the function $u$ is

$$u = 22877.94\,\mathrm{m.} \times \frac{\sin(v^{(18)} - v^{(12)} - 0.652).\sin(v^{(14)} - v^{(16)} - 0.814)}{\sin(v^{(4)} - v^{(0)} - 0.652).\sin(v^{(6)} - v^{(4)} - 0.814)}.$$

From the corrected values of the directions $v^{(0)}$, $v^{(1)}$, etc. its value is found to be

$$26766.68\,\mathrm{m.}$$

By differentiating the above expression we have

$$\begin{aligned}du = \;& 0.16991\mathrm{m.}(dv^{(0)} - dv^{(1)}) + 0.08836\mathrm{m.}(dv^{(4)} - dv^{(6)})\\ & -0.03899\mathrm{m.}(dv^{(12)} - dv^{(13)}) + 0.16731\mathrm{m.}(dv^{14} - dv^{(16)}),\end{aligned}$$

Hinc porro invenitur

$$
\begin{aligned}
[al] &= -\ \ 0.08836\\
[bl] &= +\ \ 0.13092\\
[cl] &= -\ \ 0.00260\\
[dl] &= +\ \ 0.07895\\
[el] &= +\ \ 0.03899\\
[fl] &= -40.13150\\
[gl] &= +10.99570\\
[ll] &= +\ \ 0.13238
\end{aligned}
$$

Hinc denique per methodos supra traditas invenitur, quatenus metrum pro unitate dimensionum linearium accipimus

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

unde error medius in valore lateis Falkenberg–Breithorn metuendus $= 0.2886m$ metris, (ubi $m$ error medius in directionibus observatis metuendus, et quidem in minutis secundis expressus), adeoque, si valorem ipsius $m$ supra erutum adoptamus,

$$0.1209\,\mathrm{m}$$

Ceterum inspectio systematis triangulorum sponte docet, punctum Hauselberg omnino ex illo elidi potuisse, incolumi manente nexu inter latera Wilsede–Wulfsode atque Falkenberg–Breithorn. Sed a bona methodo abhorret, *supprimere* idcirco observationes, quae ad punctum Hauselberg referuntur,*) quum certe ad praecisionem augendam conferre valeant. Ut clarius appreret, quantum praecisionis augmentum inde redundent, calculum denuo fecimus excludendo omnia, quae ad punctum Hauselberg referuntur, quo pacto e 18 directionibus supra traditis octo excidunt, atque reliquarum errores maxime plausibles ita inveniuntur:

| | |
|---|---|
| (0) = +0.327 | (12) = +0.206 |
| (1) = −0.206 | (13) = −0.206 |
| (3) = −0.121 | (14) = +0.327 |
| (4) = +0.121 | (15) = +0.206 |
| (6) = −0.121 | (16) = +0.121 |

*) Maior pars harum observationum iam facta erat, antequam punctum Breithorn repertum, atque in systema receptum esset.

where the differentials $dv^{(0)}$, $dv^{(1)}$, etc. are in seconds.

Hence we easily find that

$$
\begin{aligned}
[al] &= -\ \ 0.08836\\
[bl] &= +\ \ 0.13092\\
[cl] &= -\ \ 0.00260\\
[dl] &= +\ \ 0.07895\\
[el] &= +\ \ 0.03899\\
[fl] &= -40.13150\\
[gl] &= +10.99570\\
[ll] &= +\ \ 0.13238
\end{aligned}
$$

By the method given above

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

provided we accept the meter as the unit of linear dimension. Hence the mean error to be feared in the value of the side Falkenberg–Breithorn is $0.2886m$ meters (where $m$ is the mean error to be anticipated in the observations of the directions, expressed in seconds). Thus if we accept the value of $m$ derived above, we have a mean error of

$$0.1209\,\text{meters}.$$

Finally, an inspection of the system of triangles shows that the point Hauselberg can be deleted from the others, without destroying the connection between the sides Wilsede–Wulfsode and Falkenberg–Breithorn. But it would be a departure from good practice to *suppress* the observations relating to Hauselberg*) on that account, since they certainly increase the precision. To highlight how much increase in precision derives from this source, we have redone the calculation with all points relating to Hauselberg excluded. When this is done, eight of the eighteen directions fall out, and the most reliable errors of the remaining directions are

| | |
|---|---|
| $(0) = +0.327$ | $(12) = +0.206$ |
| $(1) = -0.206$ | $(13) = -0.206$ |
| $(3) = -0.121$ | $(14) = +0.327$ |
| $(4) = +0.121$ | $(15) = +0.206$ |
| $(6) = -0.121$ | $(16) = +0.121$ |

*)The greater part of the observations had already been made before the point Breithorn was discovered and brought into the system.

Valor lateris Falkenberg–Breithor tunc prodit = 26766.63 m, parum quidem a valore supra eruto discrepans, sed calculus ponders producit

$$\tfrac{1}{P} = 0.13082 \quad \text{sive} \quad P = 7.644$$

adeoque error medius metuendus = $0.36169m$ metris = 0.1515 m. Patet itaque, per accessionem observationum, quae ad punctum Hauselberg referuntur, pondus determinationis lateris Falkenberg–Breithorn auctum esse in ratione numeri 7.644 ad 12.006, sivi unitatis ad 1.571.

The value of the side Falkenberg–Breithorn is now 26766.63 meters, which differs little from the value derived above. However, a calculation of the weights gives

$$\frac{1}{P} = 0.13082 \quad \text{or} \quad P = 7.644,$$

and hence the mean error is $0.36169m$ meters or 0.1515 m. It is therefore clear that the addition of the observations relating to Hauselberg has increased the weight of the estimate of the side Falkenberg–Breithorn by the ratio of 12.006 to 7.644, or 1.571.

*This page intentionally left blank*

# Theoria
# Combinationis Observationum
# Erroribus Minimis Obnoxiae

*Anzeigen*

# Theory of the
# Combination of Observations
# Least Subject to Errors

*Notices*

# Theoria Combinationis Observationum Erroribus Minimis Obnoxiae

## Anzeigen

Göttingische gelehrte Anzeigen. Stück 33. Seite 321 bis 327. 1821. Februar 26.

Am 15. Februar wurde der Königl. Societät von Hrn. Hofr. GAUSS eine Vorlesung übergeben, überschrieben

*Theoria combinationis observationum erroribus minimis obnoxiae, pars prior,*

die eine der wichtigsten Anwendungen der Wahrscheinlichkeitsrechnung zum Gegenstande hat. Alle Beobachtungen, die sich auf Grössenbestimmungen aus der Sinnenwelt beziehen, können, mit welcher Genauigkeit und mit wie vortrefflichen Werkzeugen sie auch angestellt werden, nie *absolute* Genauigkeit haben; sie bleiben immer nur Näherungen, grössern oder kleinern Fehlern ausgesetzt. Nicht von solchen Fehlern ist hier die Rede, deren Quellen genau bekannnt sind, und deren Grösse bei bestimmten Beobachtungen jedesmal berechnet wedern kann; denn da dergleichen Fehler bei den beobachteten Grössen in Abzug gebracht werden können und sollen, so ist es dasselbe, als ob sie gar nicht da wären. Ganz anders verhält es sich dagegen mit den als zufällig zu betrachtenden Fehlern, die aus der beschränkten Schärfe der Sinne, aus mancherlei unvermeidlichen und keiner Regel folgenden Unvollkommenheiten der Instrumente, und aus mancherlei regelloss (wenigstens für uns) wirkenden Störungen durch äussere Umstände (z.B. das Wallen der Atmosphäre beim Sehen, Mangel absoluter Festigkeit beim Aufstellen der Instrumente) herrühren. Diese zufälligen Fehler, die dem Calcül nicht unterworfen werden können, lassen sich nicht *wegschaffen,* und der Beobachter kann sie durch sorgfältige Aufmerksamkeit und durch Vervielfältigung der Beobachtungen nur *vermindern*: allein nachdem der Beobachter das seinige gethan hat, ist es an dem Geometer, die Unsicherheit der Beobachtungen und der durch

# Theory of the Combination of Observations Least Subject to Errors

## Notices

Göttingische gelehrte Anzeigen. Volume 33. Pages 321 to 327. February 26, 1821.

On February 15 Herr Hofr. GAUSS presented a lecture to the Royal Society entitled

*Theoria combinationis observationum erroribus minimis obnoxiae, pars prior.*

The lecture concerns one of the most important applications of probability calculus.

Observations of quantities in the real world are never *absolutely* precise. However carefully they are measured and however fine the instruments, they remain approximations that are contaminated by errors of various sizes. We are not concerned here with errors whose sources are known exactly and whose magnitudes can be calculated for any given observation. Such errors can and should be removed from the measured quantities, so that it is as if the errors never were.

The case is entirely different for errors that can be thought of as random—errors arising from the finite precision of the senses, from the many unavoidable and irregular imperfections of instruments, and from the many unpredictable (at least by us) disturbances caused by external conditions (e.g., the shimmering of the atmosphere in vision, the lack of absolute stability in the support of the instrument). These random errors are not subject to calculation and cannot be *eliminated*. The observer can only *reduce* them by attentive care and by replicating his observations. After the observer has done his part, it is then up to the mathematician to assess the uncertainty in the observations—and in any quanti-

Rechnung daraus abgeleiteten Grössen nach streng mathematischen Principien zu würdigen, und was das wichtigste ist, da, wo die mit den Beobachtungen zusammenhängenden Grössen aus denselben durch verschiedene Combinationen abgeleitet werden können, diejenige Art vorzuschreiben, wobei so wenig Unsicherheit als möglich zu befürchten bleibt.

Obgleich die zufälligen Fehler als solche keinem Gesetze folgen, sondern ohne Ordnung in einer Beobachtung grösser, in einer andern kleiner ausfallen, so ist doch gewiss, dass bei einer bestimmten Beobachtungsart, auch die Individualität des Beobachters und seiner Werkzeuge als bestimmt betrachtet, die aus jeder einfachen Fehlerquelle fliessenden Fehler nicht bloss in gewissen Grenzen eingeschlossen sind, sondern dass auch alle möglichen Fehler zwischen diesen Grenzen ihre bestimmte relative Wahrscheinlichkeit haben, der zu Folge sie nach Maassgabe ihrer Grösse häufiger oder seltener zu erwarten sind, und derjenige, der eine genaue und vollständige Einsicht in die Beschaffenheit einer solchen Fehlerquelle hätte, würde diese Grenzen und den Zusammenhang zwischen der Wahrscheinlichkeit der einzelnen Fehler und ihrer Grösse zu bestimmen im Stande sein, auf eine ähnliche Weise, wie sich bei Glücksspielen, so bald man ihre Regeln kennt, die Grenzen der möglichen Gewinne und Verluste, und deren relative Wahrscheinlichkeiten berechnen lassen. Dasselbe gilt auch von dem aus dem Zusammenwirken der einfachen Fehlerquellen entspringenden Totalfehler. Auch sind diese Begriffe nicht auf unmittelbare Beobachtungen beschränkt, sondern auch auf mittelbare aus Beobachtungen abgeleitete Grössenbestimmungen anwendbar. In der Wirklichkeit werden uns freilich fast allemal die Mittel fehlen, das Gesetz der Wahrscheinlichkeiten der Fehler *a priori* anzugeben.

Wie wir die Unzulässigkeit einer bestimmten Art von Beobachtungen im Allgemeinen abschätzen wollen, hängt zum Theil von unserer Willkür ab. Man kann dabei entweder bloss die Grösse der äussersten möglichen Fehler zum Maassstabe wählen, oder zugleich auf die grössere oder geringere Wahrscheinlichkeit der einzelnen möglichen Fehler mit Rücksicht nehmen. Das letztere scheint angemessener zu sein. Allein diese Berücksichtigung kann auf vielfache Weise geschehen. Man kann, wie es die Berechner bisher gemacht haben, den sogenannten wahrscheinlichen (nicht wahrschein*lichsten*) Fehler zum Maassstabe wählen, welches derjenige ist, über welchen hinaus alle möglichen Fehler zusammen noch eben so viele Wahrscheinlichkeit haben, wie alle diesseits liegenden zusammen; allein es wird *weit vortheilhafter* sein, zu diesem Zweck statt des wahrscheinlichen Fehlers ties calculated from them — according to rigorous mathematical principles. Most important, when the quantities that depend on the observations can be derived from them by means of different combinations, the mathematician should specify the combination for which the uncertainty is as small as possible.

Random errors, as such, follow no law but are, in no particular order, larger in one observation and smaller in another. Nonetheless, in a given class of observation, where the individual observer and his instruments are taken as fixed, the error from any simple source of error lies within certain limits. Moreover, all possible errors between these limits have fixed relative probabilities and consequently can be expected to occur more or less frequently depending on their sizes. Someone with a complete and precise understanding of the workings of such an error would be in a position to determine its limits and the relation between the magnitude of an individual error and its probability — just as the limits of the possible gains and losses in a game of chance along with their relative probabilities can be calculated as soon as the rules of the game are known. The same is true of the total error generated by the combined effects of simple sources of error. Moreover, these ideas are not confined to observations alone but apply to quantities derived from the observations. In practice, of course, we will almost never have the means to specify the probability law of the errors a priori.

How to make a general assessment of the admissibility of a particular class of observations is partly a matter of individual judgment. We could adopt the size of the largest possible error as our measure; alternatively, we could take into account the various probabilities of the individual errors. The latter seems more appropriate; but it can be done in many ways. We can take the probable (not the *most* probable) error for a measure, as calculators have done previously. The probable error is the one for which all the errors lying beyond it have the same collective probability as all the errors lying within it.

Instead of the probable error, however, it will be *far more profitable* for our purposes to use the *mean error*, provided we come to a proper understanding den *mittlern* zu gebrauchen, vorausgesetzt, dass man diesen an sich noch schwankenden Begriff auf the rechte Art bestimmt. Man lege jedem Fehler ein von seiner Grösse abhängendes Moment bei, multiplicire das Moment jedes möglichen Fehlers in dessen Wahrscheinlichkeit und addire die Producte: der Fehler, dessen Moment diesem Aggregat gleich ist, wird als mittlerer betrachtet werden müssen. Allein welche Function der Grösse des Fehlers wir für dessen Moment wählen wollen, bleibt wieder unsrer *Wilkür* überlassen, wenn nur der Werth derselben immer positiv ist, und für grössere Fehler grösser als für kleinere. Der Verf. hat die einfachste Function dieser Art gewählt, nemlich das Quadrat; diese Wahl ist aber noch mit manchen andern höchst wesentlichen Vortheilen verknüpft, die bei keiner andern statt finden. Denn sonst könnte auch jede andere Potentz mit geraden Exponenten gebraucht werden, und je grösser dieser Exponent gewählt würde, desto näher würde man dem Princip kommen, wo bloss die äussersten Fehler zum Maasstabe der Genauigkeit dienen. Gegen die Art, wie ein grosser Geometer den Begriff des mittlern Fehlers genommen hat, indem er die Momente der Fehler diesen gleich setzt, wenn sie positiv sind, und die ihnen entgegengesetzten Grössen dafür gebraucht, wenn sie negativ sind, lässt sich bemerken, dass dabei gegen die mathematische Continuität angestossen wird, dass sie so gut wie jede andere auch willkürlich gewählt ist, dass die Resultate viel weniger einfach und genugthuend ausfallen, und dass es auch an sich schon natürlicher scheint, das Moment der Fehler in einem stärkern Verhältniss, wie diese selbst, wachsen zu lassen, indem man sich gewiss lieber den einfachen Fehler zweimal, als den doppelten einmal gefallen lässt.

Diese Erläuterungen mussten vorangeschickt werden, wenn auch nur etwas von dem Inhalt der Untersuchung hier angeführt werden sollte, wovon die gegenwärtige Abhandlung die erste Abtheilung ausmacht.

Wenn die Grössen, deren Werthe durch Beobachtungen gefunden sind, mit einer gleichen Anzahl unbekannter Grössen auf eine bekannte Art zusammenhangen, so lassen sich, allgemein zu reden, die Werthe der unbekannten Grössen aus den Beobachtungen durch Rechnung ableiten. Freilich werden jene Werthe auch nur näherungsweise richtig sein, in so fern die Beobachtungen es waren: allein die Wahrscheinlichkeitsrechnung hat nichts dabei zu thun, als die Unsicherheit jener Bestimmungen zu würdigen, indem sie die der Beobachtungen voraussetzt. Ist die Anzal der unbekannten Grössen grösser als die der Beobachtungen, so lassen sich jene aus diesen noch gar nicht bestimmen. Allein wenn die Anzal der of this slippery notion. Let us assign each error a moment depending on its size, multiply the moment of each error by its probability, and sum the products. Then the mean error is the error whose moment is equal to this sum. Now which function of the size of the error is to be taken for the moment is once again a matter of *individual judgment,* provided of course that the function be positive and be greater for larger errors than for smaller. The author has chosen the simplest function of this kind, namely, the square. But this choice has many other intrinsic advantages that other choices do not possess. For otherwise we could use any other power to an even exponent, and the larger the exponent, the nearer we would come to using only the extreme error as a measure of precision.

A great mathematician has adopted a concept of mean error in which the moment is set equal to the error itself when the error is positive and to the opposite of the error when the error is negative. Against this kind of mean error it can be said that it violates mathematical continuity, that it is as arbitrary as any other choice, that it gives far less simple and satisfactory results, and finally that it seems more natural to allow the moment of the error to grow in greater proportion to the error than the error itself, since we would rather put up with a simple error twice than a single error twice the size.

These observations are a necessary preface to any discussion of our inquiries, the first part of which are contained in the present work.

When quantities whose values have been determined by observations are connected with an equal number of unknown quantities in a known way, the unknowns, generally speaking, can be calculated from the observations. Insofar as the observations are only approximately correct, these values will, of course, be only approximately correct. In this case the probability calculus has nothing to do except to assess the uncertainty of the calculated values, since the latter is completely determined by the uncertainty of the observations. When the number of the unknown quantities is greater than the number of observations, there is no way to determine any of them. However, when the number of unknown unbekannten Grössen kleiner ist, als die der Beobachtungen, so ist die Aufgabe mehr als bestimmt: es sind dann unendlich viele Combinationen möglich, um aus den Beobachtungen die unbekannten Grössen abzuleiten, die freilich alle zu einerlei Resultaten führen müssten, wenn die Beobachtungen absolute Genauigkeit hätten, aber unter den obwaltenden Umständen mehr oder weniger von einander abweichende Resultate hervorbringen. Aus dieser ins Unendliche gehenden Mannichfaltigkeit von Combinationen die zweckmässigste auszuwählen, d.i. diejenige, wobei die Unsicherheit der Resultate die möglich kleinste wird, ist unstreitig eine der wichtigsten Aufgaben bei der Anwendung der Mathematik auf die Naturwissenschaften.

Der Verfasser gegenwärtiger Abhandlung, welcher im Jahr 1797 diese Aufgabe nach den Grundsätzen der Wahrscheinlichkeitsrechnung zuerst untersuchte, fand bald, dass die Ausmittelung der *wahrscheinlichsten* Werthe der unbekannten Grösse unmöglich sei, wenn nicht die Function, die die Wahrscheinlichkeit der Fehler darstellt, bekannt ist. In so fern sie dies aber nicht ist, bleibt nichts übrig, als hypothetisch eine solche Function anzunehmen. Es schien ihm das natürlichste, zuerst den umgekehrten Weg einzuschlagen und die Function zu suchen, die zum Grunde gelegt werden muss, wenn eine allgemein als gut anerkannte Regel für den einfachsten aller Fälle daraus hervorgehen soll, die nemlich dass das arithmetische Mittle aus mehreren für eine und dieselbe unbekannte Grösse durch Beobachtungen von gleicher Zuverlässigkeit gefundenen Werthen als der wahrscheinlichste betrachtet werden müsse. Es ergab sich daraus, dass die Wahrscheinlichkeit eines Fehlers $x$, einer Exponentialgrösse von der Form $e^{-hhxx}$ proportional angenommen werden müsse, und dass dann gerade diejenige Methode, auf die er schon einige Jahre zuvor durch andere Betrachtungen gekommen war, allgemein nothwendig werde. Diese Methode, welche er nachher besonders seit 1801 bei allerlei astronomischen Rechnungen fast täglich anzuwenden Gelegenheit hatte, und auf welche auch LEGENDRE inswischen gekommen war, ist jetzt unter dem Namen Methode der Kleinsten Quadrate im allgemeinen Gebrauch: und ihre Begründung durch die Wahrscheinlichkeitsrechnung, so wie die Bestimmung der Genauigkeit der Resultate selbst, nebst andern damit zusammenhängenden Untersuchungen sind in der *Theoria Motus Corporum Coelestium* ausfürlich entwickelt.

Der Marquis DELAPLACE, welcher nachher diesen Gegenstand aus einem neuen Gesichtspunkte betrachtete, indem er nicht die wahrscheinlichsten Werthe der unbekannten Grössen suchte, sondern die zweckmässigste Combination der Beobquantities is smaller than the number of observations, the problem is overdetermined, and there are infinitely many possible combinations of the observations from which the unknowns may be derived. Of course, if the observations had absolute precision, these combinations would have to yield a single result; but under the circumstances they must produce results that differ more or less greatly from one another. Undoubtedly one of the most important problems in the application of mathematics to the natural sciences is to choose from this infinite multiplicity of combinations the one that is most suitable; i.e., the one for which the uncertainty in the result is as small as possible.

In 1797 the author of the present work first examined this problem in light of the principles of probability calculus. He soon found that it is impossible to determine the *most probable* values of the unknown quantities if the function representing the probability of the error is not known. In such cases there is nothing to do but adopt a hypothetical function. It seemed to the author that the most natural thing was to proceed conversely and look, in the simplest case, for the function that necessarily makes the most probable value of a quantity measured by several equally reliable observations equal to the arithmetic mean of the observations — a rule that is generally recognized to be a good one. The result was that the probability of an error $x$ had to be taken proportional to an exponential quantity of the form $e^{-hhxx}$ and that a general method, one that the author had arrived at from other considerations a few years before, was a direct and necessary consequence. This method, which the author later had occasion to use almost daily in astronomical calculations, especially since 1801, and which in the meantime LEGENDRE also arrived at, is now in general use under the name of the method of least squares. The justification of this method by probability theory along with the determination of the precision of the results and other related inquiries is fully developed in the *Theoria Motus Corporum Coelestium*.

Later, the Marquis DELAPLACE considered the topic from a different point of view, seeking not the most probable value of the unknown quantities but the most suitable combination of the observations. He discovered the remarkable result that achtungen, fand das merkwürdige Resultat, dass, wenn die Anzahl der Beobachtungen als unendlich gross betrachtet wird, die Methode der kleinsten Quadrate allemal und unabhängig von der Function, die die Wahrscheinlichkeit der Fehler ausdrückt, die zweckmässigste Combination sei.

Man siet hieraus, dass beide Begründungen noch etwas zu wünschen übrig lassen. Die erstere ist ganz von der hypothetischen Form für die Wahrscheinlichkeit der Fehler abhängig, und sobald man diese verwirft, sind wirklich die durch die Methode der kleinsten Quadrate gefundenen Werthe der unbekannten Grössen nicht mehr die wahrscheinlichsten, eben so wenig wie die arithmetischen Mittel in dem vorhin angeführten einfachsten aller Fälle. Die zweite Begründungsart lässt uns ganz im Dunkeln, was bei einer mässigen Anzahl von Beobachtungen zu thun sei. Die Methode der kleinsten Quadrate hat dann nicht mehr den Rang eines von der Wahrscheinlichkeitsrechnung gebotenen Gesetzes, sondern empfiehlt sich nur durch die Einfachheit der damit verknüpften Operationen.

Der Verfasser, welcher in gegewärtiger Abhandlung diese Untersuchung aufs neue vorgenommen hat, indem er von einem ähnlichen Gesichtspunkte ausging, wie Delaplace, aber den Begriff des mittlern zu befürchtenden Fehlers auf eine andere, und wie ihm scheint, schon und für sich natürliche Art, feststellt, hofft, dass die Freunde der Mathematik mit Vergnügen sehen werden, wie die Methode der kleinsten Quadrate in ihrer neuen hier gegebenen Begründung allgemein als die zweckmässigste Combination der Beobachtungen erscheint, nicht näherungsweise, sondern nach mathematischer Schärfe, die Function für die Wahrscheinlichkeit der Fehler sei, welche sie wolle, und die Anzahl der Beobachtungen möge gross oder klein sein.

Mit dem Hauptgegenstande ist eine Menge anderer merkwürdiger Untersuchungen enge verbunden, deren Umfang aber den Verfasser nöthigte, die Entwickelung des grössten Theils derselben einer künftige zweiten Vorlesung vorzubehalten. Von denjenigen, die schon in der gegenwärtigen ersten Abtheilung vorkommen, sei es uns erlaubt, hier nur ein Result anzuführen. Wenn die Function, welche die relative Wahrscheinlichkeit jedes einzelnen Fehlers ausdrückt, unbekannt ist, so bleibt natürlich auch die Bestimmung der Wahrscheinlichkeit, dass der Fehler zwischen gegebene Grenzen falle, unmöglich: dessenungeachtet muss, wenn nur allemal grössere Fehler geringere (wenigstens nicht grössere) Wahrscheinlichkeit haben als kleinere, die Wahrscheinlichkeit, dass der Fehler zwischen die Grenzen $-x$ und $+x$ falle, nothwendig grösser (wenigstens nicht kleiner) sein, als $\frac{x}{m}\sqrt{\frac{1}{3}}$, when the number of observations can be regarded as infinitely large the method of least squares always gives the most suitable combination, independent of the function that expresses the probability of the error.

From the foregoing we see that the two justifications each leave something to be desired. The first depends entirely on the hypothetical form of the probability of the error; as soon as that form is rejected, the values of the unknowns produced by the method of least squares are no more the most probable values than is the arithmetic mean in the simplest case mentioned above. The second justification leaves us entirely in the dark about what to do when the number of observations is not large. In this case the method of least squares no longer has the status of a law ordained by the probability calculus but has only the simplicity of the operations it entails to recommend it.

In the present work the author has taken up the inquiry anew. He starts from a viewpoint similar to DELAPLACE's, but defines the notion of the mean error in another (it seems to him) intrinsically more natural way. He hopes that the friends of mathematics will view with pleasure how in its new justification the method of least squares emerges as the generally most suitable combination of the observations — not approximately, but with mathematical rigor, whatever the probability function of the error and however large or small the number of observations.

The main topic is closely connected with a number of other noteworthy inquiries. However, the range of these inquiries has forced the author to reserve them for a future, second work. Here we can only present a single result from the inquiries of the first part. When the function representing the relative probability of the individual errors is unknown, it is, of course, impossible to determine the probability that the error falls between given limits. Nonetheless, if $m$ denotes the mean error to be feared in the observations and if larger errors always have smaller (or at least not greater) probabilities than smaller errors, then the probability that the error falls between the limits $-x$ and $+x$ is necessarily greater wenn $x$ kleiner ist als $m\sqrt{\frac{3}{4}}$, and nicht kleiner als $1 - \frac{4mm}{9xx}$, wenn $x$ grösser ist als $m\sqrt{\frac{3}{4}}$, wobei $m$ den bei den Beobachtungen zu befürchtenden mittlern Fehler bedeutet. Für $x = m\sqrt{\frac{3}{4}}$ fallen wie man sieht beide Ausdrücke zusammen.

(or at least not less) than $\frac{x}{m}\sqrt{\frac{1}{3}}$, if $x$ is smaller than $m\sqrt{\frac{3}{4}}$, and not smaller than $1 - \frac{4mm}{9xx}$, when $x$ is greater than $m\sqrt{\frac{3}{4}}$. When $x = m\sqrt{\frac{3}{4}}$, the two expressions coincide.[13]

[13]The fractions $\sqrt{\frac{3}{4}}$ should be replaced by $\sqrt{\frac{4}{3}}$. However, even with this correction, the first inequality does not follow from the theorem stated in Part One.

Göttingische gelehrte Anzeigen. Stück 32. Seite 313 bis 318. 1823. February 24.

Eine am 2. Febr. der Königl. Societät von Hrn. Hofr. GAUSS überreichte Vorlesung, überschrieben

*Theoria combinationis observationum erroribus minimis obnoxiae, pars posterior,*

steht im unmittelbaren Zusammenhange mit einer frühern, wovon in diesen Blättern [1821 Februar 26] eine Anzeige gegeben ist. Wir bringen darüber nur kurz in Erinnerung, dass ihr Zweck war, die sogenannte Methode der kleinsten Quadrate auf eine neue Art zu begründen, wobei diese Methode nicht näherungsweise, sondern in mathematischer Schärfe, nicht mit der Beschränkung auf den Fall einer sehr grossen Anzahl von Beobachtungen, und nicht abhäning von einem hypothetischen Gesetze für die Wahrscheinlichkeit der Beobachtungsfehler, sondern in vollkommener Allgemeinheit, als die zweckmässigste Combinationsart der Beobachtungen erscheint. Der gegenwärtige zweite Theil der Untersuchung enthält nun eine weitere Ausführung dieser Lehre in einer Reihe von Lehrsätzen und Problemen, die damit in genauester Verbindung stehen. Es würde der Einrichtung dieser Blätter nicht angemessen sein, diesen Untersuchungen hier Schritt vor Schritt zu folgen, auch unnöthig, da die Abhandlung selbst bereits unter der Presse ist. Wir begnügen uns daher, nur die Gegenstände von einigen dieser Untersuchungen, die sich leichter isolirt herausheben lassen, hier anzuführen.

Die Werthe der unbekannten Grössen, welche der Methode der kleinsten Quadrate gemäss sind, und die man die *sichersten Werthe* nennen kann, werden vermittelst einer bestimmten Elimination gefunden, und die diesen Bestimmungen beizulegenden Gewichte vermittelest einer unbestimmten Elimination, wie dies schon aus der *Theoria Motus Corporum Coelestium* bekannt ist: auf eine neue Art wird hier *a priori* bewiesen, dass unter den obwaltenden Voraussetzungen diese Elimination allemal möglich ist. Zugleich wird eine merkwürdige Symmetrie unter den bei der unbestimmten Elimination hervorgehenden Coöfficienten nachgewiesen.

So leicht und klar sich diese Eliminationsgeschäfte im Allgemeinen übersehen lassen, so ist doch nicht zu läugnen, dass die wirkliche numerische Ausführung, bei einer beträchtlichen Anzahl von unbekannten Grössen, beschwerlich wird. Was die

Göttingische gelehrte Anzeigen. Volume 32. Pages 313 to 318. February 24, 1823.

On February 2 Herr Hofr. GAUSS presented a lecture to the Royal Society entitled

*Theoria combinationis observationum erroribus minimis obnoxiae, pars posterior.*

The lecture is a direct continuation of an earlier lecture, whose notice appeared in these pages [February 26, 1821]. Regarding this earlier lecture, we recall only that its purpose was to give a new justification of the so-called method of least squares, a justification that establishes it rigorously as the most suitable way of combining observations—not approximately but in full generality. There is no requirement that the number of observations be large, and there is no hypothetical law for the probability of the errors in the observations.

The second part of the investigation contains a further development of the subject in a sequence of theorems and problems that are intimately connected with it. Here it would not be suitable to follow these investigations step by step. Moreover, it is unnecessary, since the memoir is already in press. Therefore, we will content ourselves with presenting the substance of a few of the investigations that can be easily isolated from the others.

The values of the unknowns from the method of least squares, which we may call the *most reliable* values, are found by a definite elimination. The corresponding weights are found by a general elimination, a fact which was already established in the *Theoria Motus Corporum Coelestium.* Here we give a new, a priori proof that the elimination is always possible under the underlying assumptions. At the same time, we establish a remarkable symmetry among the coefficients from the general elimination.

Although it is easy to give a clear description of the elimination as a general process, its numerical realization is undeniably burdensome when the number of *bestimmte* Elimination, die zur Ausmittelung der sichersten Werthe für die unbekannten Grössen zureicht, betrifft, so hat der Verfasser ein Verfahren, wodurch die wirkliche Rechnung, so viel es nur die Natur der Sache verträgt, abgekürzt wird, bereits in der *Theoria Motus Corporum Coelestium* angedeutet, und in einer im ersten Bande der *Commentt. Rec. Soc. R. Gott.* befindlichen Abhandlung, *Disquisitio de elementis ellipticis Palladis,* ausführlich entwickelt. Dieses Verfahren gewährt zugleich den Vortheil, dass das Gewicht der Bestimmung der einen unbekannten Grösse, welche man bei dem Geschäft als die letzte betrachtet hat, sich von selbst mit ergibt. Dan nun die Ordnung unter den unbekannten Grösse gänzlich willkürlich ist, und man also welche man will, als die letzte behandeln kann, so ist dies Verfahren in allen Fällen zureichend, wo nur für Eine der unbekannten Grössen das Gewicht mit verlangt wird, und die beschwerliche unbestimmte Elimination wird dann umgangen.

Die seitdem bei den rechnenden Astronomen so allgemein gewordene Gewohnheit, die Methode der kleinsten Quadrate auf schwierige astronomische Rechnungen anzuwenden, wie auf die volständige Bestimmung von Cometenbahnen, wobei die Anzahl der unbekannten Grössen bis auf sechs steigt, hat indess das Bedürfniss, das Gewicht der sichersten Werthe *aller* unbekannten Grössen auf eine bequemere Art als durh die unbestimmte Elimination, zu finden, fühlbar gemacht, und da die bemühungen einiger Geometer*) keinen Erfolg gehabt hatten, so hat man sich nur so geholfen, dass man den oben erwähnten Algorithmus so viele male mit veränderter Ordnung der unbekannten Grössen durchführte, als unbekannte Grössen waren, indem man jeder einmal den letzten Platz anwies. Es scheint uns jedoch, dass durch dieses kunstlose Verfahren in Vergleichung mit der unbestimmten Elimination in Rücksicht auf Kürze der Rechnung nichts gewonnen wird. Der Verfasser hat daher diesen wichtigen Gegenstand einer besondern Untersuchung unterworfen, und einen neuen Algorithmus zur Bestimmung der Gewichte der Werthe *sämmtlicher* unbekannten Grössen mitgetheilt, der alle Geschmeidigkeit und Kürze zu haben scheint, welcher die Sache ihrer Natur nach fähig ist.

Der sicherste Werth einer Grösse, welche eine gegebene Function der unbekannten Grössen der Aufgabe ist, wird gefunden, indem man für letztere ihre durch die Methode der kleinsten Quadrate erhaltenen sichersten Werthe substituirt. Allein eine bisher noch nicht behandelte Aufgabe ist es, wie das jener Bestimmung

*) z.B. PLANA'S. Siehe Zeitschrift für Astronomie und verwandte Wissenschaften Band 6, S. 258.

unknowns is large. For the *definite* elimination, which suffices to find the most reliable values of the unknowns, the author had already sketched a method in the *Theoria Motus Corporum Coelestium* that shortens the actual calculations as far as possible. This method was later described explicitly in a memoir entitled *Disquisitio de elementis ellipticis Palladis*, which appeared in the first volume of the *Commentt. Rec. Soc. R. Gott.* This method also has the advantage that it immediately gives the weight of the single unknown appearing last in the elimination process. Since the order of the unknowns is entirely arbitrary and any unknown can be taken for the last, the method is sufficient in all cases where the weight of only one unknown is sought. In this case, the burdensome general elimination can be avoided.

Since then it has become the custom of computational astronomers to apply the method of least squares to difficult astronomical calculations, such as the complete determination of the orbits of comets, in which the number of unknown quantities can be as large as six. Here the necessity of finding the weights of *all* the unknowns by a more convenient method than general elimination has made itself felt. Since the efforts of certain mathematicians*) have met with no success, people have been reduced to executing the above algorithm as many times as there are unknowns, ordering the unknowns so that each in turn falls in the last place. It seems to me, however, that this artless method is no better than general elimination with respect to economy of calculation. Therefore, the author has undertaken a special investigation of this important problem and has provided a new algorithm for determining the weights of the values of *all* the unknowns, which is as flexible and economical as the nature of the problem permits.

The most reliable value of a function of the unknown quantities under consideration is found by replacing the unknowns with the most reliable values determined by the method of least squares. A previously untreated problem is to find the as-

*) E.g., Plana's. See *Zeitschrift für Astronomie und verwandte Wissenschaften*, Vol. 6, p. 258.

beizulegende Gewicht zu finden sei. Die hier gegebene Auflösung dieser Aufgabe verdient um so mehr von den rechenden Astronomen beherzigt zu werden, da sich findet, dass mehrere derselben dabei früher auf eine nicht richtige Art zu Werke gegangen sind.

Die Summe der Quadrate der Unterschiede zwischen den unmittelbar beobachteten Grössen, und denjenigen Werthen, welche ihre Ausdrücke, als Functionen der unbekannten Grössen, durch Substitution der sichersten Werthe für letztere erhalten (welche Quadrate, im Fall die Beobachtungen ungleiche Zuverläsigkeit haben, vor der Addition erst noch durch die respectiven Gewichte multiplicirt werden müssen) bildet bekanntlich ein absolutes Minimum. Sobald man daher einer der unbekannten Grössen einen Werth beilegt, der von dem sichersten verschieden ist, wird ein ähnliches Aggregat, wie man auch die übrigen unbekannten grössen bestimmen mag, allezeit grösser ausfallen, als das erwänte Minimum. Allein die übrigen unbekannten Grössen werden sich nur auf Eine Art so bestimmen lassen, dass die Vergrösserung des Aggregats so klein wie möglich, oder dass das Aggregat selbst ein relatives Minimum werde. Diese von dem Verfasser hier ausgeführte Untersuchung führt zu einigen interessanten Wahrheiten, die über die ganze Lehre noch ein vielseitigeres Licht verbreiten.

Es fügt sich zuweilen, dass man erst, nachdem man schon eine ausgedehnte Rechnung über eine Reihe von Beobachtungen in allen Theilen durchgeführt hat, Kenntnisse von einer neuen Beobachtung erhält, die man gern noch mit zugezogenen hätte. Es kann in vielen Fällen erwünscht sein, wenn man nicht nöthig hat, deshalb die ganze Eliminationsarbeit von vorne wieder anzufangen, sondern im Stande ist, die durch das Hinzukommen der neuen Beobachtung entstehende Modification in den sichersten Werthen und deren Gewichten zu finden. Der Verfasser hat daher dieser Aufgabe hier besonders abgehandelt, eben so wie die verwandte, wo man einer schon angewandten Beobachtung hintennach ein anderes Gewicht, als ihr beigelegt war, zu ertheilen sich veranlasst sieht, und, ohne die Rechnung von vorne zu wiederhollen, die Veränderungen der Endresultate zu erhalten wünscht.

Wie der *wahrscheinliche* Fehler einer Beobachtungsgattung (als bisher üblicher Massstab ihrer Unsicherheit) aus einer hinlänglichen Anzahl wirklicher Beobachtungsfehler näherungsweise zu finden sei, hatte der Verfasser in einer besondern Abhandlung in der Zeitschrift für Astronomie und verwandte Wissenschaften [1816 März u. April] gezeigt: diese Verfahren, so wie der Gebrauch des wahrsociated weight. The solution to the problem given here should be all the more valued by computational astronomers, since it turns out that some of them have previously gone about it in the wrong way.

It is well known that the sum of squares of the differences between the observed quantities and the values obtained by substituting the most reliable values of the unknowns into their representations as functions of the unknowns (and in the case of observations of unequal accuracy multiplying by their respective weights before summing) is an absolute minimum. Thus, whenever an unknown is assigned a value different from its most reliable value, the corresponding sum of squares must be greater than the minimum, whatever values are given to the other unknowns. However, there is only *one* way to determine the other unknowns so that the sum of squares is as small as possible, i.e., so that the sum attains a relative minimum. Investigations of the author along these lines lead to some interesting results that throw a revealing light on the entire theory.

It occasionally happens that after we have completed all parts of an extended calculation on a sequence of observations, we learn of a new observation that we would like to include. In many cases we will not want to have to redo the entire elimination but instead to find the modifications due to the new observation in the most reliable values of the unknowns and in their weights. The author has given special treatment to this problem. He has also treated the related problem where one finds it necessary to give a different weight to an observation that has already been used and wishes to obtain the changes in the final results without repeating the calculation from the beginning.

In a special article in the *Zeitschrift für Astronomie und verwandte Wissenschaften* [March and April 1816], the author showed how to approximate the scheinlichen Fehlers überhaupt, ist aber von der hypothetischen Form der Grösse der Wahrscheinlichkeit der einzelnen Fehler abhängig, und musste es sein. Im ersten Theile der gegenwärtigen Abhandlung ist nun zwar gezeigt, wie aus denselben Datis der mittlere Fehler der Beobachtungen (als zweckmässiger Maassstab ihrer Ungenauigkeit) näherungweise gefunden wird. Allein immer bleibt hierbei die Bedenklichkeit übrig, dass man nach aller Schärfe selten oder fast nie im Besitz der Kenntniss der wahren Grösse von einer Anzahl wirklicher Beobachtungsfehler sein kann. Bei der Ausübung hat man dafür bisher immer die Unterschiede zwischen dem, was die Beobachtungen ergeben haben, und den Resultaten der Rechnung nach den durch die Methode der kleinsten Quadrate gefundenen sichersten Werthen der unbekannten Grössen, wovon die Beobachtungen abhangen, zum Grunde gelegt. Allein da man nicht berechtigt ist, die sichersten Werthe für die wahren Werthe selbst zu halten, so überzeugt man sich leicht, dass man durch dieses Verfahren allemal den wahrscheinlichen und mittlern Fehler *zu klein* finden muss, und daher den Beobachtungen und den daraus gezogenen Resultaten eine grössere Genauigheit beilegt, als sie wirklic besitzen. Freilich hat in dem Falle, wo die Anzahl der Beobachtungen vielemale grösser ist als die der unbekannten Grössen, diese Unrichtigkeit wenig zu bedeuten; allein theils erfordert die Würde der Wissenschaft, dass man vollständig und bestimmt übersehe, wieviel man hierdurch zu fehlen Gefahr läuft, theils sind auch wirklich öfters nach jenem fehlerhaften Verfahren Rechnungsresultate in wichtigen Fällen aufgestellt, wo jene Voraussetzung nicht Statt fand. Der Verfasser hat daher diesen Gegenstand einer besondern Untersuchung unterworfen, die zu einem sehr merkwürdigen höchst einfachen Resultate geführt hat. Man braucht nemlich den nach dem angezeigten fehlerhaften Verfahren gefundenen mittlern Fehler, um ihn in den richtigen zu verwandeln, nur mit

$$\sqrt{\frac{\pi-\rho}{\pi}}$$

zu multipliciren, wo $\pi$ die Anzahl der Beobachtungen und $\rho$ die Anzahl der unbekannten Grössen bedeutet.

Die letzte Untersuchung betrifft noch die Ausmittelung des Grades von Genauigkeit, welcher dieser Bestimmung des mittlern Fehlers selbst beigelegt werden muss: die Resultate derselben müssen aber in der Abhandlung selbst nachgelesen werden.

*probable* error of a class of observations (up to now the customary measure of its uncertainty) from a sufficient number of actual observation errors. This method—as well as the use of probable errors in general—depends, as it must, on the hypothetical form of the probabilities of individual errors. Now in the first part of the present memoir it is shown how the mean error of the observations (an appropriate measure of the imprecision) may be approximated from the same data. Yet there is always the fact that we seldom if ever have a rigorous knowledge of the true values of a collection of actual observation errors. In practice, therefore, we replace them by the difference between the values given by the observations and the results of calculating the most reliable values of the unknowns on which the observations depend by the method of least squares. But since we are not justified in taking the most reliable values for the true values, we can easily convince ourselves that this method always yields probable and mean errors that are too small, with the consequence that the observations and any results derived from them are assigned a greater precision than they really have. Of course, in the case where the number of observations is many times greater than the number of unknowns this lapse counts for little. But even so, in the name of science, we should determine completely and definitively the amount of error. Moreover, the results of this erroneous method are often reported in important cases where the above stipulation does not hold. Therefore, the author has devoted a special investigation to the subject, which has lead to a remarkable and extremely simple result. Namely, to correct the mean errors found by the above erroneous method, we need only multiply by

$$\sqrt{\frac{\pi-\rho}{\pi}},$$

where $\pi$ denotes the number of observations and $\rho$ the number of unknowns.

The last investigation concerns the determination of the degree of precision that must be assigned to the above estimate of the mean error itself. For the results of this investigation the reader must consult the memoir itself.

Göttingische gelehrte Anzeigen. Stück 153. Seite 1521 bis 1527. 1826. September 25.

Am 16. September überreichte der Herr Hofr. GAUSS der königl. Societät eine Vorlesung:

*Supplementum theoriae combinationis observationum erroribus minimis obnoxiae.*

Bei allen frühern Arbeiten über die Anwendung der Wahrscheinlichkeitsrechnung auf die zweckmässigste Benutzung der Beobachtungen, und namentlich auch in der Behandlung dieses Gegenstandes im fünften Bande der *Commentationes recentiores* liegt in Beziehung auf die Form der Hauptaufgabe eine bestimmte Voraussetzung zum Grunde, die allerdings den meisten in der Ausübung vorkommenden Fällen angemessen ist. Diese Voraussetzung besteht darin, dass die beobachteten Grössen auf eine bekannte Art von gewissen unbekannten Grössen (Elementen) abhängen, d.i. bekannte Functionen dieser Elemente sind. Die Anzahl dieser Elemente muss, damit die Aufgabe überhaupt hierher gehöre, kleiner sein, als die Anzahl der beobachteten Grössen, also diese selbst abhängig von einander.

Inzwischen sind doch auch die Fälle nicht selten, wo die gedachte Voraussetzung nicht unmittelbar Statt finden, d.i. wo die beobachteten Grössen noch nicht in der Form von bekannten Functionen gewisser unbekannter Elemente gegeben sind, und wo man auch nicht sogleich siet, wie jene sich in eine solche Form bringen lassen; wo hingegen zum Ersatz die gegenseitige Abhängigkeit der beobachteten Grössen (die natürlich auf irgend eine Weise gegeben sein muss) durch gewisse Bedingungsgleichungen gegeben ist, welche die wahren Werthe von jenen, der Natur der Sache nach, nothwendig genau Genüge leisten müssen. Zwar sieht man bei näherer Betrachtung bald ein, dass dieser Fall von dem andern nicht wesentlicht, sondern bloss in der Form verschieden ist, und sich wirklich, der Theorie nach leicht, auf denselben zurüchfüren lässt; allein häufig bleibt dies doch ein unnatürlicher Umweg, der in der Anwendung viel beschwerlichere Rechnungen herbeiführt, als eine eigne der ursprünglichen Gestalt der Aufgabe besonders angemessene Auflösung. Diese ist daher der Gegenstand der gegenwärtigen Abhandlung, und die Auflösung der Aufgabe, welche sie als ein selbstständiges von der frühern Abhandlung unabhängiges Ganze gibt, hat ihrerseits eine solche Geschmeidigkeit, dass es sogar in manchen Fällen vortheilhaft sein kann, sie selbst da

Göttingische gelehrte Anzeigen. Volume 153. Pages 1521 to 1527. September 25, 1826.

On September 16 Herr Hofr. GAUSS presented a lecture to the Royal Society entitled

*Supplementum theoriae combinationis observationum erroribus minimis obnoxiae.*

In all previous work on the application of probability calculus to the most suitable use of observations and especially in the treatment of this subject in the fifth volume of the *Commentationes recentiores*, the main problem is assumed to have a certain form, one that is appropriate for most cases arising in practice. Specifically, it is assumed that the observed quantities depend in a known way on certain unknown quantities (elements); i.e., they are known functions of the elements. In the problems considered here, the number of elements must be less than the number of observations, which consequently depend on one another.

However, it is not uncommonly the case that the above assumption is not directly satisfied; i.e., the observed quantities are not given in the form of known functions of certain unknown elements, and it is not immediately obvious how to cast them in such a form. On the contrary, the mutual dependence of the observations (which, of course, must be specified in one way or another) is given by certain conditional equations, which the true values by the nature of the problem have to satisfy exactly. Now a closer examination quickly shows that the difference between the two cases is merely a matter of form, not substance: in theory the latter can easily be reduced to the former. But in many cases this is nothing more than an unnatural dodge that gives rise to more burdensome calculations than would a special solution tailored to the original form of the problem. Such a solution is the object of the present memoir. The solution itself, which is self-contained and independent of the earlier memoir, is in turn so adaptable that in anzuwenden, wo die bei der ältern Methode zum Grunde liegende Voraussetzung schon von selbst erfüllt war.

Die Hauptaufgabe stellt sich hier nun unter folgender Gestalt dar. Wenn von den Grössen $v$, $v'$, $v''$ u.s.w., zwischen welchen ein durch eine oder mehrere Bedingungsgleichungen gegebener Zusammenhang Statt findet, eine andere auf irgend eine Art abhänigig ist, z.B. durch die Function $u$ ausgedrücket werden kann, so wird eben dieselbe auch auf unendlich viele andere Arten aus jener bestimmt, oder durch unendlich viele andere Functionen, statt $u$, ausgedrückt werden können, die aber natürlich alle einerlei Resultate geben, in so fern die Werthe von $v$, $v'$, $v''$ u.s.w., welche allen Bedingungsgleichungen Genüge leisten, substituirt werden. Hat man aber nur genäherte Werthe von $v$, $v'$, $v''$ u.s.w., wie sie Beobachtungen von beschränkter Genauigkeit immer nur liefern können, so können auch die daraus abgeleiteten Grössen auf keine absolute Richtigkeit Anspruch machen: die verschiedenen für $u$ angewandten Functionen werden, algemein zu reden, ungleiche, aber was die Hauptsache ist, ungleich zuverlässige Resultate geben. Die Aufgabe ist nun, aus der unendlichen Mannifaltigkeit von Functionen, durch welche die unbekannte Grösse ausgedrückt werden kann, diejenige auszuwählen, bei deren Resultat die möglich kleinste Unzuverlässigkeit zu befürchten bleibt.

Die Abhandlung gibt eigentlich *zwei* Auflösungen dieser Aufgabe. Die erste Auflösung erreicht das Ziel auf dem kürzesten Wege, wenn wirklich nur Eine unbekannte von den Beobachtungen auf eine vorgeschriebene Art abhängige Grösse abzuleiten ist. Allein die nähere Betrachtung dieser Auflösung führt zugleich auf das merkwürdige Theorem, dass man für die unbekannte Grösse genau denselben Werth, welcher aus der zweckmässigsten Combination der Beobachtungen folgt, erhählt, wenn man an die Beobachtungen gewisse nach bestimmten Regeln berechnete Veränderungen anbringt, und sie dann in irgend einer beleibigen Function, welche die unbekannte Grösse ausdrückt, substituirt. Diese Veränderungen haben neben der Eigenschaft, dass sie allen Bedingungsgleichungen Genüge leisten, noch die, dass unter allen denkbaren Systemen, welche dasselbe thun, die Summe ihrer Quatrate (in so fern die Beobachtungen als gleich zuverlässig vorausgesetzt wurden) die möglich kleinste ist. Man sieht also, dass hierdurch zugleich eine neue Begründung der Methode der kleinsten Quadrate gewonnen wird, und dass diese von der Function $u$ ganz unabhängige *Ausgleichung* der Beobachtungen eine zweite Auflösungsart abgibt, die vor der ersten einen grossen Vorzug hat, wenn man mehr als Eine unbekannte Grösse aus den Beobachtungen auf die many cases it is advantageous to use it even when the basic assumptions of the first method hold.

The main problem may be stated in the following form. Let $v$, $v'$, $v''$, etc. be quantities between which there is a dependence specified by one or more conditional equations. Let another quantity be dependent on the given quantities in some way — say it can be expressed as a function $u$. Then $u$ can be determined in infinitely many different ways from the $v$'s; that is, it can be represented by infinitely many functions other than $u$. Now these functions have to give identical results whenever they are evaluated at the true values of $v$, $v'$, $v''$, etc. satisfying the conditional equations. However, if we have only approximate values of $v$, $v'$, $v''$, etc., such as observations of limited precision have to produce, no quantities derived from them can lay claim to absolute accuracy. Generally speaking, the various functions used in place of $u$ become unequal; however, the crux of the matter is that they give results of unequal reliability. The problem is then to select from this infinite multiplicity of functions that represent the unknown the one that gives a value whose unreliability is least.

The memoir actually gives *two* solutions to this problem. The first solution is the most direct when only *one* unknown depending on the observations in a prescribed manner is to be derived. However, a closer consideration of this solution leads directly to a remarkable theorem. Specifically, the value of the unknown quantity that is obtained from the most suitable combination can also be obtained by altering the observations according to a fixed rule and substituting them into any function that represents the unknown. These alterations have the additional property that they cause the conditional equations to be satisfied and that among all such systems their sum of squares is as small as possible (insofar as the observations are assumed to be equally reliable). We see, therefore, that we have found a new justification for the principle of least squares. Moreover, the *adjustments* of the observations are entirely independent of the function $u$ and give a second solution that is considerably superior to the first when there is more than *one* unknown to be derived from the observations in the most suitable way.

zweckmässigste Art abzuleiten ist: in der That werden die Beobachtungen dadurch zu *jeder* von ihnen zu machenden Anwendung fertig vorbereitet. Nur musste bei dieser zweiten Auflösung noch eine besondere Anleitung hinzukommen, den Grad der Genauigkeit, der bei jeder einzelnen Anwendung erreicht wird, zu bestimmen. Für dies alles enthält die Abhandlung vollständige und nach Möglichkeit einfache Vorschriften, die natürlich hier keines Auszuges fähig sind. Eben so wenig können wir hier in Beziehung auf die, nach der Entwicklung der Hauptaufgaben, noch ausgeführten anderweitigen Untersuchungen, welche mit dem Gegenstande in innigem Zusammenhange stehen, uns in das Einzelne einlassen. Nur das Eine merkwürdige Theorem führen wir hier an, dass die Vorschriften zur vollständigen Ausgleichung der Beobachtungen immer einerlei Resultat geben, sie mögen auf die ursprünglichen Beobachtungen selbst, oder auf die bereits einstweilen *unvollständig* ausgeglichenen Beobachtungen angewandt werden, in so fern dieser Begriff in der Abhandlung näher bestimmten Bedeutung genommen wird, unter welcher, als specieller Fall, derjenige begriffen ist, wo mit den Beobachtungen schon eine zwar vorschriftsmässig ausgeführte, aber nur einen Theil der Bedingungsgleichungen berücksichtigende Ausgleichung vorgenommen war.

Den letzten Theil der Abhandlung machen ein paar mit Sorgfalt ausgearbeitete Beispiele der Anwendung der Methode aus, die theils von den geodätischen Messungen des Generals von Krayenhoff, theils von der vom Verfasser selbst im Königreich Hannover ausgeführten Triangulirung entlehnt sind, und die dazu dienen können, sowohl die Anwendung dieser Theorie mehr zu erläutern, als auch manche, dergleichen Messungen betreffende, Umstände überhaupt in ein helleres Licht zu stellen.

Die trigonometrischen Messungen gehören ganz besonders in das Feld, wo die Wahrscheinlichkeitsrechnung Anwendung findet, und namentlich in derjenigen Form Anwendung findet, die in der gegenwärtigen Abhandlung entwickelt ist. Gerade hier ist es Regel, dass mehr beobachtet wird, als unumgänglich nöthig ist, und dass so die Messungen einander vielfältig controlliren. Nur durch die Benutzung der strengen Grundsätze der Wahrscheinlichkeitsrechnung kann man von diesem Umstande den Vortheil ganz ziehen, der sich davon ziehen lässt, und den Resultaten die grösste Genauigkeit geben, deren sie fähig sind. Ausserdem aber geben jene Grundsätze zugleich das Mittel, die Genauigkeit der Messungen selbst, und die Zulässigkeit der darauf gegründeten Resultate zu bestimmen. Endlich dienen sie dazu, bei der Anordnung des Dreieckssystems, aus mehreren, In fact, the adjustments prepare the observations so that they are ready for *any* application of them. The only other thing this second solution needs is a means of calculating the degree of precision that holds for any particular application. The memoir contains a complete procedure for this that is as simple as possible and which naturally cannot be summarized here.

Nor can we enter into the details of the subsequent investigations that are intimately connected with the subject of this memoir. Here we will only cite a single remarkable theorem. The prescription for complete adjustment of the observations always gives the same results whether it is applied to the original observation or to observations that have already been *partially* adjusted, provided the notion of partial adjustment is understood in the narrow sense it is given in the memoir. A special case is where the observations are adjusted according to the prescription but only with respect to some of the conditional equations.

The last part of the memoir consists of some carefully worked out examples of the method, which are borrowed from the geodetic measurements of General VON KRAYENHOFF and from a triangulation in the Kingdom of Hanover carried out by the author himself. They serve not only to further illustrate the application of the theory, but they also throw a brighter light on the many circumstances that affect these kinds of measurements.

Trigonometric measurements belong *par excellence* to an area in which the probability calculus applies and particularly applies in the form developed in the present memoir. Here one typically observes more than is strictly necessary, so that the measurements constrain one another in many ways. Only by using rigorous principles of probability calculus can we take fullest advantage of this circumstance to obtain results of greatest possible precision. Moreover, these principles also furnish the means to estimate the precision of the measurements themselves and the admissibility of the results based on them. Finally, the principles help us select the most suitable ordering of the system of triangles from among the several unter denen man vielleicht die Wahl hat, das zweckmässigste auszuwählen. Und alles dieses nach festen sichern Regeln, mit Ausschliessung aller Willkürlichkeiten. Allein sowohl die sichere Würdigung, als die vollkommenste Benutzung der Messungen ist nur dan möglich, wenn sie in reiner Autenthicität und Vollständigkeit vorliegen, und es wäre daher sehr zu wünschen, dass alle grösseren auf besondere Genauigkeit Anspruch machenden Messungen dieser Art immer mit aller nöthigen Ausführlichkeit bekannt gemacht werden möchten. Nur zu gewöhnlich is das Gegentheil, wo nur Endresultate für die einzelnen gemessenen Winkel mitgetheilt werden. Wenn solche Endresultate nach richtigen Grundsätzen gebildet werden, indem man durchaus alle einzelnen Beobachtungsreihen, die nicht einen durchaus unstatthaften Fehler gewiss enthalten, dazu concurriren lässt, so ist der Nachtheil freilich lange nicht so gross, als wenn man etwa nur diejenigen Reihen beibehält, die am besten zu den nahe liegenden Prüfungsmitteln passen, welche die Summen der Winkel jedes Dreiecks und die Summen der Horizontalwinkel um jeden Punkt herum darbieten. Wo dies durchaus verwerfliche Verfahren angewandt ist, sei es aus Unbekanntschaft mit den wahren Grundsätzen einer richtigen Theorie, oder aus dem geheimen Wunsche, den Messungen das Ansehen grösserer Genauigkeit zu geben, geht der Maassstab zu einer gerechten Würdigung der Beobachtungen und der aus ihnen abzuleitenden Resultate verloren; die gewöhnliche Prüfung nach den Winkelsummen in den einzelnen Drieecken, und bei den Punkten, wo die gemessenen Winkel den ganzen Horizont umfassen, scheint dann eine Genaugkeit der Messungen zu beweisen, von der sie vielleicht ser weit entfernt sind, und wenn andere Prüfungsmittel, durch die Seitenverhältnisse in geschlossenen Polygonen oder durch Diagonalrichtungen vorhanden sind, werden diese die Gewissheit des Daseins von viel grössern Fehlern verrathen. Umgekehrt aber, wenn die zuletzt erwähnte Voraussetzung Statt findet, und das Ausgleichen der Beobachtungen in Beziehung auf die Prüfungsmittel ohne die sichern Vorschriften der Wahrscheinlichkeitsrechnung versucht ist, wo es immer ein Herumtappen im Dunkel bleiben muss, und grössere, oft viel grössere, Correctionen herbeiführt, als nöthig sind, kann leicht dadurch ein *zu* ungünstiges Urtheil über die Messungen veranlasst werden. Diese Bemerkungen zeigen die Wichtigkeit sowohl einer hinlänglich ausführlichen Bekanntmachung, als einer auf strenge Principien gegründeten mathematischen Combination der geodätischen Messungen: sie gelten aber offenbar that one may be able to choose. And all this according to hard and fast rules in which nothing is arbitrary.

However, the reliable assessment and the best use of the measurements is only possible when authentic and complete measurements are available. It is therefore highly desirable that all the more important measurements of this kind having claim to special accuracy be made known in full detail. Only too often the opposite happens, and only the final results for the individual measured angles are reported. Now the loss is not too great when such final results are produced according to correct principles and all observations not containing completely inadmissible errors are combined. It is worse when, say, someone only keeps the sequence that best satisfies the easy tests of validity furnished by the sum of the angles of each triangle and the sum of the horizontal angles around each point. Wherever this thoroughly objectional method is used — whether because the true principles of a correct theory are unknown or because of a secret wish to give the measurements the appearance of greater accuracy — we loose the yardstick by which the measurements and their derived results can be correctly assessed. In this case, the usual tests of summing the angles of individual triangles and the angles around points where the measured angles span the entire horizon lend an appearance of accuracy to the measurements that can be quite spurious; and when different means of testing are available — by the ratios of the sides of closed polygons or by diagonal lines — they will disclose the certain presence of very large errors.

On the other hand, when the above-mentioned assumption holds and one tries to adjust the observations with respect to the test conditions without the secure prescriptions of probability calculus, which amounts to nothing more than groping around in the dark, one will introduce larger corrections than necessary — often very much larger. In this case the assessment of the errors can easily be *too unfavorable.*

These remarks show the importance both of making geodetic measurements known in adequate detail and of combining them according to rigorous princimehr oder weniger bei Beobachtungen jeder Art, astronomischen, physikalischen u.s.w., die sich auf das Quantitative beziehen, insofern die Mannigfaltigkeit der dabei Statt findenden Umstände zu wechselseitigen Controllen Mittel darbietet.

ples. Obviously the remarks hold for quantitative observations of all kinds—astronomical, physical, etc.—to the extent that the particularities of the problem furnish mutual constraints.

# Afterword

*This page intentionally left blank*

# Afterword

The *Theoria Combinationis* was Gauss's last word on least squares. But it was not his first; nor was Gauss the first to address the problem of combining systems of measurements. The story is a long one, spanning a century or more, with a cast of famous characters — Laplace, Legendre, Gauss — and some less well known players. It is against this setting that Gauss's achievements in the *Theoria Combinationis* must be evaluated. Consequently, this survey divides naturally into two parts. The first is a sketch of the background, with emphasis on the earlier work of Gauss and Laplace. The second is an amplification of the material in the *Theoria Combinationis* itself. Here my purpose has been to strike a balance between Gauss the probabilist and statistician, who has been well served by the secondary literature, and Gauss the linear algebraist and numerical analyst, who has been somewhat slighted.

For the most part this afterword is based on my own reading of the original sources; however, I have been helped by the large body of secondary literature. Special mention should be made of Plackett's balanced treatment of the discovery of least squares [34], which contains translations of most of the correspondence cited below. Stigler's *History of Statistics* [42], though it does not treat the *Theoria Combinationis*, contains indispensable background information, as does a paper on Boskovich by Eisenhart [2]. I have also benefited from two papers on the contributions of Laplace and Gauss by Sheynin [38], [39].

## Gauss's Schooldays

During his years as a student at Göttingen, Gauss experienced a period of intense creativity, comparable to the *anni mirabiles* of Newton's youth. Gauss was accustomed to recording his discoveries in a diary, which, with great good fortune, has come down to us.[14] The entries are brief, often obscure, but they confirm the truth of Gauss's off-hand claims to priority that irritated many of his contemporaries. Two entries also substantiate his more insistent claim to the method of least squares.

*Calculus probabilitatis contra La Place defensus*—probability calculus defended against Laplace. We know from Gauss's correspondence with his friend Olbers

[14]A facsimile of the diary may be found in Gauss's collected works [14, v. 10, pp. 483–572].

[14, v. 8, p. 140] that this entry (dated June 17, 1798) refers to a probabilistic justification of the method of least squares. Gauss was defending against Laplace's method for combining observations (really due to Boscovich — see below), in which the sum of the absolute value of the residuals is to be minimized subject to the condition that the residuals sum to zero. Twenty-three years later in the Notice to the *Pars Prior* [6], he gives us the details.

> In 1797 the author of the present work first examined this problem in light of the principles of probability calculus. He soon found that it is impossible to determine the *most probable* values of the unknown quantities if the function representing the probability of the error is not known. In such cases there is nothing to do but adopt a hypothetical function. It seemed to the author that the most natural thing was to proceed conversely and look, in the simplest case, for the function that necessarily makes the most probable value of a quantity measured by several equally reliable observations equal to the arithmetic mean of the observations — a rule that is generally recognized to be a good one. The result was that the probability of an error $x$ had to be taken proportional to an exponential quantity of the form $e^{-hhxx}$ and that a general method, one that the author had arrived at from other considerations a few years before, was a direct and necessary consequence. This method, which the author later had occasion to use almost daily in astronomical calculations, especially since 1801, and which in the meantime LEGENDRE also arrived at, is now in general use under the name of the method of least squares. The justification of this method by probability theory along with the determination of the precision of the results and other related inquiries is fully developed in the *Theoria Motus Corporum Coelestium*.

In view of the subsequent priority controversy with Legendre, we must ask how reliable this statement is. Most of Gauss's commentators take his version of the discovery at face value, and they have good reasons for doing so. We know from his diary that Gauss was punctilious in his claims. Moreover, his statements on least squares, public and private, are to the point, consistent, and many were made at a time when Gauss could be expected to have a clear memory of what he had done. The statement quoted above, for example, was made by Gauss in his mid forties about work done in his early twenties. And that is by no means the earliest statement. In 1806, in a letter to Olbers [14, v. 8, pp. 138–139], Gauss claimed the principle of least squares, expressing regret that once again his work had become entangled with that Legendre. Finally, Olbers, a person of unquestioned probity [33], testified that Gauss communicated the method to him in 1803, two years before Legendre published it.

But there is more. The next entry in Gauss's diary, dated June 1798, reads, *Problema eliminationis ita solutum ut nihil amplius desiderari possit*—the problem of elimination resolved in such a way that nothing more can be desired. I take this entry to refer to Gaussian elimination—not just the algorithm, but the flexible theoretical tool that Gauss employed throughout his work on least squares. The accompanying comment in Gauss's collected works refers this entry to his dissertation, where, as part of a review of past attempts to prove the fundamental theorem of algebra, he discusses the problem of elimination of coefficients among polynomials and promises a future treatment of elimination in general. Either reading is possible, but I think the former is more convincing. Having developed a probabilistic justification of least squares, Gauss would naturally go on to determine the precision of his estimates, a task which requires a theory of elimination. It is therefore significant that the entry occurs immediately after the one on least squares. It may be only coincidence that the words of the entry are echoed in the *Pars Posterior* (Art. 31), where the context is the calculation of weights by elimination: *Gratum itaque geometris fore speramus, si modum novum pondera determinationum calculandi ... hic exponamus, qui nihil amplius desiderandum relinquere videtur.* [I therefore hope that mathematicians will be grateful if I set forth a new method for calculating the weights of the estimates ... which seems to leave nothing more to be desired.]

Thus, we can say with some confidence that in 1794 or 1795 (both dates appear in Gauss's correspondence) Gauss arrived at the principle of least squares, presumably on grounds of utility. In 1798, responding to Laplace and Boscovich, Gauss formulated his first probabilistic justification of the method of least squares, essentially as it appears in the *Theoria Motus*. If the diary entry regarding elimination is admitted to refer to least squares, Gauss knew how to compute the weights of his least squares estimates. Weights or not, Gauss went on to use the method almost daily in his astronomical work (see the letter to Laplace dated 1812 in [34]).

## Legendre and the Priority Controversy

Filling in the story of Gauss's discovery of least squares would be largely an academic exercise if it were not for the priority dispute with Legendre. In 1805 Legendre [29] published the method of least squares in an appendix to his treatise *Nouvelles méthodes pour la détermination des orbites des comètes.* The appendix, which is reproduced in [42, p. 58], is a model of clear exposition; the only improvement we might make on it today is to use matrix notation. Although Legendre gave no formal justification for his method, other than to say that it prevents extreme errors from prevailing by establishing an equilibrium among all the errors, he wrote down the normal equations, pointed out that the mean was a special case, and noted that observations found to have large errors could be conveniently removed from the equations after the fact.

The priority controversy arose when Gauss published the method in 1809 in his *Theoria Motus* and called it "my principle" (*principium nostrum*). Legendre took exception in a letter to Gauss, in which he asserted his claim to the method by virtue of first publication.[15] Gauss never responded, but continued privately and publicly to claim priority of discovery. The story has been well worked over, and there is no need to go into details here (e.g., see [34], [43]). I would like, however, to make one observation.

It has become a commonplace that Gauss got away with murder on matters of priority simply because he was Gauss. The facts suggest otherwise. We know now that Gauss, had he published, could have been the founder of complex analysis, of the theory of elliptic functions, and of non-Euclidean geometry. Yet though we give a polite nod toward Gauss in these matters, we assign credit to the people who actually published—Cauchy for complex analysis, Abel and Jacobi for elliptic functions, and Lobachevsky and Bolyai for non-Euclidean geometry. Least squares is different. Gauss was openly involved in its development. He published two justifications of the principle and went on to explore its consequences with a thoroughness not matched until the twentieth century. Moreover, he developed powerful algorithms for implementing all aspects of least squares. Thus, Gauss's claim to least squares should not be confounded with his casual remarks on unpublished results.

The question of who should be credited with the discovery of least squares

[15] Plackett [34] gives a translation of the entire letter.

is, therefore, one whose answer depends on how you feel about publication and priority. My own view is that since Gauss was working on his *Theoria Motus* at the time Legendre published and since its appearance was delayed owing to circumstances beyond Gauss's control,[16] Legendre and Gauss should be regarded as independent codiscoverers. Legendre caused the method of least squares to become widely known. Gauss and Laplace went on to give it a sound theoretical basis.

## Beginnings: Mayer, Boscovich, and Laplace

Gauss and Legendre did not work in a vacuum. Astronomers had been combining observations and mathematicians had been grappling with the slippery issues of probability long before Gauss was born. Stigler has told this story in his *History of Statistics* [42, Ch. 1], and here we will just glance briefly at the highlights.

It is convenient to cast the problem of combining observations in terms of matrices. Suppose we are given a set of overdetermined linear equations of the form

$$0 = \mathbf{A}\mathbf{x} + \mathbf{l}, \tag{1}$$

where $\mathbf{A}$ is an $\pi \times \rho$ matrix with $\pi > \rho$, and $\mathbf{l}$ is a vector of observations.[17] The problem is to combine these equations into a square system so that they can be solved for $\mathbf{x}$. If we restrict ourselves to linear combinations of the equations, we can write the general solution in the following form. Choose an $\pi \times \rho$ matrix $\mathbf{N}$ and form the square system

$$0 = \mathbf{N}^{\mathrm{T}}\mathbf{A}\mathbf{x} + \mathbf{N}^{\mathrm{T}}\mathbf{l}.$$

Different choices of $\mathbf{N}$ give different combinations. For example, the choice

$$\mathbf{N} = \begin{pmatrix} \mathbf{I}_\rho \\ 0 \end{pmatrix},$$

where $\mathbf{I}_\rho$ is the identity matrix of order $\rho$, amounts to selecting the first $\rho$ equations and throwing out the others. Least squares results from choosing $\mathbf{N} = \mathbf{A}$.

[16]In a letter to Laplace (1812) quoted by Plackett [34], Gauss gives a brief statement of the genesis of the *Theoria Motus*. Owing to the conquest of the German states by Napoleon, Gauss's publisher required him to translate his manuscript, which was started in 1805 and finished in 1806, from the original German into Latin.

[17]The notation here, and later, echoes that of the *Theoria Combinationis*.

It should be stressed that this matrix representation is thoroughly anachronistic — and not just because it uses matrices. Even written out in scalar form, the idea of a general combination of equations (as represented by the matrix $\mathbf{N}$) did not appear until 1811, when Laplace [25] gave his justification of the principle of least squares. Be that as it may, matrix notation can help extract essential ideas from a welter of scalar formulas, and we will use it throughout this survey. But the reader is warned to consult the original works before making assertions based on the matrix equations found here.

The difficulty with the above approach is that in the absence of a rigorous principle the choice of $\mathbf{N}$ becomes something of an art. For example, Johan Tobias Mayer, in his 1750 work on the libration of the moon [32], was faced with the problem of combining twenty-seven observations in three unknowns. His procedure was to divide the equations into three sets of nine and average within sets to give three linear equations in three unknowns. In matrix terms this amounts to ordering the equations into groups and taking

$$\mathbf{N} = \begin{pmatrix} \mathbf{1} & 0 & 0 \\ 0 & \mathbf{1} & 0 \\ 0 & 0 & \mathbf{1} \end{pmatrix},$$

where each vector $\mathbf{1}$ has nine components that are all one. The arbitrariness comes in choosing how to group the equations.

As another example, in 1785 Laplace [22] combined a set of observations by taking $\mathbf{N}$ to be a matrix whose elements were 1, 0, or $-1$. He choose $\mathbf{N}$ to give a set of "good" equations for the problem at hand. According to Stigler, this method competed with the method of least squares until the middle of the nineteenth century simply on grounds of computational utility — the formation of $\mathbf{N}^{\mathrm{T}}\mathbf{A}$ requires no multiplications.

An important step in providing an objective criterion for combining observations had already been taken by Boskovich [1], [40], [31] in 1757 (for more details see [2], [42]). He regarded the residual vector $\mathbf{r} = \mathbf{A}\mathbf{x} + \mathbf{l}$ as a function of $\mathbf{x}$ and required that $\mathbf{x}$ be chosen so that the sum of the components of $\mathbf{r}$ be zero and the sum of the absolute values of the components be a minimum. In other words he cast the problem in the form

$$\begin{aligned} &\text{minimize} && \mathbf{1}^{\mathrm{T}}|\mathbf{A}\mathbf{x} + \mathbf{l}| \\ &\text{subject to} && \mathbf{1}^{\mathrm{T}}(\mathbf{A}\mathbf{x} + \mathbf{l}) = 0, \end{aligned} \tag{2}$$

where **1** is the vector whose components are all one. Since Boscovich had only to fit a straight line, he was able to use the geometry of the problem to derive an effective algorithm for solving (2). Laplace [23] took up Boscovich's method in 1789 and gave an analytic derivation. It was this paper that spurred Gauss to defend probability calculus against Laplace.

Conceptually, Boscovich's criterion is an advance over ad-hoc choices of $\mathbf{N}$, and the criterion itself is reasonable. Without the side condition that the residuals sum to zero, it is $\ell_1$ regression, which is recommended today as being resistant to outliers in the data. Computationally, however, it is a step backward from the simple algorithms of Mayer and Laplace. Boscovich was able to come up with an algorithm only because of the specific form of his problem. Even today, computer programs for $\ell_1$ approximation are written only by experts, and they are computationally expensive. With the tools available at the time, Boscovich's method could be nothing more that a promising curiosity.

The connection of probability theory with a special problem in the combination of observations was made by Laplace in 1774 [21]. The paper, entitled *Mémoire sur la probabilité des causes par les évènements,* begins with a principle of inverse probability. It is effectively Bayes's theorem with a uniform prior. Laplace gives no proof.

Laplace then proceeds to four applications. The one that concerns us here is the estimation of the true value represented by three observations $l_1$, $l_2$, and $l_3$. If the true value is $-x$ and $v_i = x + l_i$ $(i = 1, 2, 3)$, then the $v_i$ are errors. Laplace assumes a common distribution $\varphi$ for the $v_i$ and applies his principle to show (in modern notation)

$$\Pr(x|l_1, l_2, l_3) \propto \varphi(x + l_1)\varphi(x + l_2)\varphi(x + l_3).$$

Laplace's next problem is how to use this posterior distribution to estimate $x$. He proposes two estimates. The first is the median of the posterior distribution. The second is the value that minimizes the mean absolute error; that is, choose $x$ so that

$$\int |y + x| \Pr(y|l_1, l_2, l_3)\, dy$$

is minimized. Laplace goes on to establish the beautiful result that the two choices are the same.

The reader is referred to Stigler [42, pp. 109 ff.] for the difficulties Laplace encountered when he attempted to derive and use a specific distribution function

$\varphi(x)$. For our purposes, we need only note that the memoir contains a principle of inverse probability and, once again, the mean absolute error as an optimizing criterion. Both ideas will recur later.

### Gauss and Laplace

Laplace's work described above was the beginning of a game of intellectual leapfrog between Gauss and Laplace that spanned several decades, and it is not easy to untangle their relative contributions. The problem is complicated by the fact that the two men are at extremes stylistically. Laplace is slapdash and lacks rigor, even by the standards of the time, while Gauss is reserved, often to the point of obscurity. Neither is easy to read. They also have different strong points: Laplace was the more acute probabilist, while Gauss had a deeper understanding of the machinery of least squares. By all accounts, relations between the two were cordial.

### The Theoria Motus

Gauss's response to Boscovich as presented by Laplace was to develop the general theory of least squares that he finally presented in 1809 in the *Theoria Motus* [3]. In outline, Gauss's theory consists of the following elements.

1. A nonlinear model with a general distribution of the errors.
2. A Bayesian argument with uniform prior to justify maximum likelihood estimates of the parameters in the model.
3. An argument to justify the assumption of normally distributed errors.
4. The introduction of the linear model and the derivation normal equations.
5. A proof that the normal equations can be solved.
6. The use of Gaussian elimination to calculate the relative precision of the estimates.

Let's take a brief look at each item in turn.

**The nonlinear model.** Gauss was concerned with estimation in astronomy and geodesy, both of which give rise to nonlinear problems. There are two ways of dealing with nonlinearity. The first is to linearize the problem and analyze the resulting model—the usual practice. The alternative is to live with the nonlinearities as long as possible. Gauss chose the latter option in the *Theoria Motus*, and the choice accounts for its stately progression from the general to the specific.

Gauss's nonlinear model may be written $\mathbf{v} = \mathbf{f}(\mathbf{x}) + \mathbf{l}$, where $\mathbf{l}$ is a vector of observations, $\mathbf{x}$ is a vector of unknown parameters, and $\mathbf{f}$ is a nonlinear function, explicitly assumed to be nondegenerate. The components of $\mathbf{v}$ are the errors, which are assumed to be independent and to have a common distribution. This is the first appearance of what we now call the (nonlinear) Gaussian regression model.

**Maximum likelihood estimation.** The development here is almost modern. Gauss assumes that in the absence of other information all values of the unknowns are equally likely. He then uses a version of Bayes's theorem with a uniform prior to show that the most probable estimate is the one that maximizes the likelihood of the observations. By setting derivatives to zero, he derives a system of maximum likelihood equations that the estimates must satisfy. He calls these estimates the *most probable values* of the unknowns.

Here we must ask how much Gauss was influenced by Laplace's 1774 paper. The answer, I think, is unknowable. If Gauss had studied Laplace's paper he would have had all that was necessary for this part of his theory. But Gauss was quite capable of working things out for himself.

**Normality.** To progress further, Gauss must commit himself to a distribution of the errors. He reasons as follows. The mean is generally agreed to furnish a good approximation to the most probable value of a set of measurements taken under uniform conditions. But if the mean is to be equal to the most probable value, then (Gauss shows) the distribution of the errors $\Delta$ must be proportional to $e^{-h^2\Delta^2}$; i.e., the errors must be normally distributed. Therefore, the normal distribution is to be assumed in the general case.

Gauss's reasoning has been called circular; but, as the quotation on page 208 shows, he was following the common scientific technique of looking at special cases to determine necessary hypotheses. If you believe, as Gauss evidently did at the time, that the mean is good in the sense that it approximates the probable value of the error, then Gauss's proof compels you to conclude that errors in observations are approximately normal. To be sure, the argument is not very convincing, since it is not clear how one might come by the supporting belief. But it is not circular; and what's more it hit the mark. In many instances errors in observations are approximately normal, and the hypothesis of normality continues to play an important role in modern regression theory.

Gauss goes on to identify the parameter $h$ with the precision of the observations. Although the reciprocal of $h^2$ is twice the variance, it would be wrong to impute the introduction of variance and standard deviation to Gauss at this time. His argument here is that the size of a centered interval containing the error with fixed probability is inversely proportional to $h$. Only later in the *Theoria Combinationis* was Gauss to recognize the importance of second moments.

Gauss now returns to the maximum likelihood estimate and shows that if the errors are normal and of equal precision, then the maximum likelihood estimate minimizes the residual sum of squares, thus completing his justification of the principle of least squares.

**The linear model.** Only at this point does Gauss pass to a linear model

$$\mathbf{v} = \mathbf{A}\mathbf{x} + \mathbf{l},$$

which he obtains by linearizing about rough approximations to the true values of $\mathbf{x}$. He derives the normal equations and points out that they can be solved by the "usual method of elimination" (*eliminatio vulgaris*). It is ironic that the *Theoria Motus* should have become the principle reference for the algorithm of Gaussian elimination. As we will see, the method is implicit in the way Gauss derives formulas for the precision of his estimates, but he gives the explicit algorithm in a later work [4].

**The existence of estimates.** With typical caution, Gauss verifies that the normal equations always have a solution. He appeals to a "theory of elimination" for a condition under which the normal equations fail to have a solution and shows that the condition implies the original problem is underdetermined.

**The precision of the estimates.** Gauss next turns to computing the precision of his estimates. He begins with the case of one unknown, and argues that the precision of the posterior distribution should be taken as the precision of the most probable value. The argument leans heavily on the assumption of normality.

Gauss's problem is now to elevate this result into general formulas for the precision of the most probable values of the unknowns, and he solves it by Gaussian elimination. We will present the details later (see also [41]). For now note that Gauss decomposed the residual sum of squares into a sum of functions, the second of which is independent of the first unknown, the third of which is independent of the first two unknowns, and so on to the last, which depends only on the last unknown. Assuming normality, he was then able to integrate out all the unknowns except the last and deduce its distribution. He then goes on to show that the precisions of all the unknowns can be found by inverting the normal equations. Although today we can establish this result with a few of lines of matrix algebra, at the time it was a virtuoso derivation of one of the fundamental results of regression theory.

There are other juicy nuggets scattered throughout the work. Gauss sketches what is today called the Gauss–Newton method for solving nonlinear least squares problems (Art. 180). In Art. 186, he considers the possibility of minimizing the sum of the errors taken to an even power greater than two, and he rejects it because of computational difficulties. He points out that for an infinite power this procedure leads to $\ell_\infty$ approximation. With explicit reference to Boscovich and Laplace, he considers $\ell_1$ approximation in general, and gives the modern characterization of the solution. This work also contains the famous *principium nostrum*.

## Laplace and the Central Limit Theorem

Laplace [24] published his central limit theorem in 1809. In 1810 he applied it to justify the principle of least squares [25]. Here we follow the version in his *Théorie Analytique des Probabilités* [26], which was published in 1812.

Laplace considered the cases of one and two unknowns, treating each separately. We will examine his treatment of two unknowns, since only at this generality can we appreciate both the magnitude and limitations of Laplace's achievement. We begin with an exposition in matrix notation of what Laplace did.

Laplace starts with the linear Gaussian model

$$\mathbf{v} = \mathbf{A}\mathbf{x} + \mathbf{l}.$$

The errors $\mathbf{v}$ are assumed to be independent and have a common distribution that is symmetric about the origin and has finite support. He then chooses two combinations, which we will represent by a matrix $\mathbf{N}$ having two columns, and lets $\hat{\mathbf{x}}$ be the solution of the square system

$$0 = \mathbf{N}^{\mathrm{T}}\mathbf{A}\hat{\mathbf{x}} + \mathbf{N}^{\mathrm{T}}\mathbf{l}.$$

If we let $\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}$ be the error in the estimate $\hat{\mathbf{x}}$, then clearly

$$\mathbf{g} \equiv \mathbf{N}^{\mathrm{T}}\mathbf{v} = \mathbf{N}^{\mathrm{T}}\mathbf{A}\mathbf{e}.$$

Laplace now proceeds in three steps.

1. He applies a bivariate version of his proof of the central limit theorem to the relation $\mathbf{g} = \mathbf{N}^{\mathrm{T}}\mathbf{v}$ to show that $\mathbf{g}$ is asymptotically normally distributed with variance proportional to $\mathbf{N}^{\mathrm{T}}\mathbf{N}$.

2. From the relation $\mathbf{g} = \mathbf{N}^{\mathrm{T}}\mathbf{A}\mathbf{e}$ he deduces that $\mathbf{e}$ is asymptotically normally distributed with variance proportional to

$$\mathbf{B}_{\mathbf{N}} = (\mathbf{N}^{\mathrm{T}}\mathbf{A})^{-1}(\mathbf{N}^{\mathrm{T}}\mathbf{N})(\mathbf{N}^{\mathrm{T}}\mathbf{A})^{-1}.$$

3. With the distribution of the two components $e_1$ and $e_2$ of $\mathbf{e}$ in hand, he asks which choice of $\mathbf{N}$ will minimize the mean *absolute* error of the components of $\mathbf{e}$. Since the mean absolute error of a normal variate is an increasing function of its variance, $\mathbf{N}$ must be chosen to minimize the diagonal elements of $\mathbf{B}_{\mathbf{N}}$. From this Laplace concludes that $\mathbf{N} \propto \mathbf{A}$; i.e., that the values of $\hat{\mathbf{x}}$ are the ones that would result from the principle of least squares.[18]

This is remarkable stuff. The bivariate case of the central limit theorem is not a mere variant of the univariate case, and Laplace's proof—what it lacks in rigor it makes up in chutzpa—is a testament to his analytical powers. In explaining why he minimizes the mean absolute error, Laplace fully orchestrates one of the themes of his 1774 paper by comparing the absolute value of an error to a loss in a game of chance: if you can't win, you should at least cut your average loss.

On the other hand, our exposition in terms of matrices hides the fact that Laplace worked in terms of scalar formulas that do not suggest ready generaliza-

[18]In the univariate case Laplace also argues that the probability of finding the error—call it $e$—in a fixed interval symmetric about zero increases as its variance $\beta$ decreases. Consequently, the probability that $e$ remains near zero increases as $\beta$ decreases, clearly a good reason for minimizing $\beta$. The focus on intervals, which is equally applicable to the individual unknowns of the multivariate case, echoes Gauss's treatment of the precision of a normal random variable in the *Theoria Motus*.

tions. For example, Laplace writes the distribution of $\mathbf{g}$ in the form[19]

$$\frac{1}{\frac{4k''\pi}{k}.a^2\sqrt{E}}.e^{-\frac{k}{4k''a^2}.\frac{l^2.S.n^{(i)2}-2ll'.S.m^{(i)}n^{(i)}+l'^2.S.m^{(i)2}}{E}}$$

where

$$E = S.m^{(i)2}.S.n^{(i)2} - (S.m^{(i)}S.n^{(i)})^2.$$

Here $l$ and $l'$ are the components of $\mathbf{g}$ and the numbers $S.m^{(i)2}$, $S.n^{(i)2}$, and $S.m^{(i)}S.n^{(i)}$ are the elements of $\mathbf{N}^{\mathrm{T}}\mathbf{N}$. With hindsight we can recognize the quadratic form $\mathbf{g}^{\mathrm{T}}(\mathbf{N}^{\mathrm{T}}\mathbf{N})\mathbf{g}$ and the determinant $\det(\mathbf{N}^{\mathrm{T}}\mathbf{N})$ in these expressions, but to Laplace they are only the result of ingenious substitutions. Again, Laplace passes from the distribution of $\mathbf{g}$ to that of $\mathbf{e}$ by an explicit change of variables in a double integral, a process that does not generalize without additional mathematical apparatus. Finally, having arrived at a formula for the elements of $(\mathbf{N}^{\mathrm{T}}\mathbf{A})^{-1}(\mathbf{N}^{\mathrm{T}}\mathbf{N})(\mathbf{N}^{\mathrm{T}}\mathbf{A})^{-1}$, Laplace pulls out his famous *"il est facile de voir"* and asserts without proof that the choice $\mathbf{N} \propto \mathbf{A}$ minimizes the diagonals. The general result is not trivial, even when cast in terms of matrices.[20]

The limitations of Laplace's approach should not be allowed to obscure the magnitude of his contribution: the failure is one of technique, not of inspiration. We can point to four significant ideas.

1. The bivariate central limit theorem.
2. The notion of a combination of equations as a linear combination whose coefficients can be optimized.
3. The introduction of the absolute error as a loss function and hence its use as an optimizing criterion.
4. The realization that the first three ideas justify the principle of least squares when the number of observations is large.

Gauss would draw heavily on items two and three in his final treatment of least squares.

[19]This is not exactly the modern form. The constant $a$ represents the interval of support for the distribution of the errors, and major surgery is required to get rid of it.

[20]Gauss in a letter to Olbers dated 1819 [34] questions whether Laplace's approach generalizes. On Laplace's own claim to generality, Todhunter says [44, Art. 1010], "This assertion, however, seems very far from being obvious."

## The Theoria Combinationis Observationum

In 1821 Gauss [6] announced a new justification of the principle of least squares to the Royal Scientific Society of Gottingen. In it he gave a concise summary of the state of affairs after Laplace's justification.

> From the foregoing we see that the two justifications [Gauss (1809), Laplace (1812)] each leave something to be desired. The first depends entirely on the hypothetical form of the probability of the error; as soon as that form is rejected, the values of the unknowns produced by the method of least squares are no more the most probable values than is the arithmetic mean in the simplest case mentioned above. The second justification leaves us entirely in the dark about what to do when the number of observations is not large. In this case the method of least squares no longer has the status of a law ordained by the probability calculus but has only the simplicity of the operations it entails to recommend it.

As an astronomer and geodesist, Gauss would have been especially concerned with small numbers of observations. In astronomy systems with small numbers of observations are common; in geodesy the number of observations may not greatly exceed the number of conditional equations, which amounts to the same thing.

The nub of Gauss's argument was that among all linear combinations of the observations, the least squares estimate has minimum variance—or, as Gauss would have said, mean error to be feared. This result, and much more, appeared in 1823 in a memoir divided into two parts: the *Pars Prior* [8] and the *Pars Posterior* [9]. In 1828 Gauss published the *Supplementum* [11], a separate memoir devoted to least squares in geodesy.

The purpose of the present translation is to allow readers to follow Gauss for themselves, and there is no need for a blow-by-blow description of the *Theoria Combinationis.* However, some of the topics covered by Gauss deserve further amplification. We will treat five: the precision of observations, the combination of observations, the inversion of linear systems, Gaussian elimination, and the generalized minimum variance theorem.

## The Precision of Observations

Gauss's point of departure is the use of the second moment of a distribution of errors to measure their weight. In one sense this can be seen as Laplace's mean error, but with the square of the error, rather than its absolute value, as the loss function, and Gauss presents it as such. But the larger part of the *Pars Prior* is devoted to giving this number a life of its own: it says things about the distribution itself—in particular, things that suggest that large mean square errors are to be avoided.

Gauss begins with a careful consideration of errors in observations (Arts. 1–3). He distinguishes between systematic and random errors and urges the observer to eliminate the former. He then (Arts. 4–6) introduces the distribution function, its mean or "constant part," and finally its second moment, which he claims is the best measure of the uncertainty of the observations. Gauss's second moment or "mean error to be feared" is not quite our variance, since he does not necessarily subtract the mean before computing it. However, the distinction is moot in Gauss's application, since he explicitly assumes that the errors in his observations have no constant part.

As we have indicated, Gauss's mean error to be feared is defined in analogy with Laplace's mean error. In a lengthy passage, Gauss repeats Laplace's argument that an error can be considered to be the outcome of a game of chance with an associated loss. But where Laplace takes the loss to be proportional to the absolute value of the error, Gauss recognizes that the loss can be represented by an arbitrary positive function of the error, and he chooses the square for its simplicity.

Gauss expressed himself on this point at least three times. In a letter to Olbers, dated February 1819 [34], Gauss was clearly struggling to come to terms with Laplace's mean error. When he covers the same ground in the *Anzeige* (which was probably written after the *Pars Prior*), he is evenhanded, even genial. The present passage, on the other hand, has a slightly shrill tone and ends with a flamboyant piece of puffery. Possibly the contrast is due to Gauss's difficulties with Latin, which, he felt, distorted the natural flow of thought [39, §2.6.1].

In any event, Gauss clearly recognized that he was introducing something new, something that required more than rhetoric to justify it. He thus moves on to examine the consequences of his choice. Recall that in the *Theoria Motus* Gauss had shown that the parameter $h$ in the normal distribution was related to what we now call the interpercentile range. His concern now is to relate his mean error to interpercentile ranges, and in particular with the interquartile range, also called the probable error (see [39, §4.4]). Since the relation depends on the distribution, the most he can hope for in general are inequalities. After treating three specific distributions — uniform, triangular, and normal — he proves a general Chebyshev-like inequality under the assumption that the distribution is unimodal (Art. 10). The proof is Gauss at his grimmest — a forced march from hypothesis to conclusion — though an aptly drawn picture will reveal what he is about.

Gauss then states and sketches a proof of the central theorem of the calculus of expectations (Arts. 12–13). If $f$ is a function of a random variable $x$ with distribution function $\varphi$, then the expectation of $f(x)$ is given by

$$\mathbf{E}[f(x)] = \int f(x)\varphi(x)\,dx.$$

Gauss restricts himself to "rational" functions, and it requires great generosity on the part of the reader to conclude that he actually proves anything. However, he uses the result only in simple cases where, as he remarks, it is easily established.

Gauss next treats convergence of the sample mean and variance (Arts. 15–16). His approach is to calculate the variances of the sample mean and variance and show that they approach zero as the sample size increases. From this he concludes that the probability of the sample deviating from its true value approaches zero. Strictly speaking, Gauss's Chebyshev-like inequality is not sufficient to establish this conclusion: the mean and the mode of the sample variance are not the same. But he is close, and more important he gets the correct rates of convergence.

In the *Theoria Motus* Gauss gave formulas for the weights of his least squares estimates. However, the formulas assume that the weight of the errors in the observations is known. In an 1816 paper, *Bestimmung der Genauigkeit der Beobachtungen* [5], Gauss considered the problem of estimating the weight of a normally distributed variable from a finite sample, and, as we have just seen, he treated the general case in the *Pars Prior*.[21] Toward the end of the *Pars Posterior* (Arts. 37–38), he comments that one seldom if ever has a supply of pure errors. He therefore considers the problem of estimating the precision of the observations from the residual sum of squares, which he denotes by $N$.

Gauss first criticizes the practice of estimating precision from the quantity $N/\pi$, where $\pi$ is the number of observations. Since $N$ minimizes sum of squares,

[21] Gauss's mean error to be feared is completely absent from the *Bestimmung*. The weight in question is a parameter in a specific distribution, namely the normal. This suggests that Gauss formulated his second justification of the principle of least squares sometime between the writing of the *Bestimmung* and his letter to Olbers cited on page 221 — say conservatively between 1815 and 1819.

it must be less than the value obtained by substituting the true values of the parameters, which would, of course, give the exact errors. Thus the use of $N/\pi$ exaggerates the precision of the measurements. Gauss then goes on to derive the unbiased formula $N/(\pi - \rho)$, where $\rho$ is the number of parameters. Finally (Art. 39), he derives an expression for the variance of $N$, to be used in assessing the accuracy of the of the estimate of the weights of the errors.

This part of Gauss's work on least squares has been less than fully appreciated. One reason is that the mean and variance of a distribution along with the calculus of expectations are the staples of introductory mathematical statistics courses, which makes it hard to see their formulation as anything but a trivial exercise. Gauss, of course, owes a debt to Laplace's mean error; but he took the raw idea, gave it a new twist, and expanded it into a body of useful theorems and techniques. It is appropriate to conclude with a list of what was new in Gauss's treatment of random errors.

1. The careful distinction between systematic and random errors.
2. The use of the first moment of a distribution to measure its center.
3. The use of the second moment to measure precision.
4. A Chebyshev-like inequality.
5. The correct formula for the expectation of a function of a random variable.
6. The rate of convergence of the sample mean and variance.
7. The correct formula for estimating the precision of observations from the residual sum of squares.

### The Combination of Observations

The climax of the *Pars Prior* is Gauss's proof of his minimum variance theorem (Arts. 19–21). Unlike Laplace, he was able to deal with the linear model in its full generality. To see how he was able to do it, it will be worthwhile to recast his development in terms of matrices.

After a linearization, Gauss arrives at the model

$$\mathbf{v} = \mathbf{A}\mathbf{x} + \mathbf{l}, \tag{3}$$

where $\mathbf{A}$ is an $\pi \times \rho$ matrix. The vector $\mathbf{v}$ is the vector of errors, and $\mathbf{l}$ is the vector of measured quantities.

Gauss takes his key from Laplace in seeking to optimize a combination. Specifically, he seeks to estimate the first component $x_1$ of $\mathbf{x}$ in the form $\boldsymbol{\kappa}^{\mathrm{T}}\mathbf{x}$. He notes (later—here we depart slightly from his order of presentation) that if

$$\boldsymbol{\kappa}^{\mathrm{T}}(\mathbf{A}\mathbf{x}+\mathbf{l}) \equiv x_1 - k, \qquad \text{identically in } \mathbf{x}, \tag{4}$$

then $k$ is an estimate of $x_1$ whose mean square error is $m^2\|\boldsymbol{\kappa}\|^2$, where $m$ is the mean error of the components of $\mathbf{v}$ and $\|\cdot\|$ is the usual Euclidean norm. The problem of finding the best estimate is thus one of finding the vector $\boldsymbol{\kappa}$ of least norm satisfying (4).

To solve this problem Gauss introduces the components of

$$\boldsymbol{\xi} = \mathbf{A}^{\mathrm{T}}\mathbf{v},$$

which can be written as a function of $\mathbf{x}$ in the form

$$\boldsymbol{\xi} = \mathbf{B}\mathbf{x} + \mathbf{m}, \tag{5}$$

where $\mathbf{B} = \mathbf{A}^{\mathrm{T}}\mathbf{A}$ and $\mathbf{m} = \mathbf{A}^{\mathrm{T}}\mathbf{l}$. Solving this equation for $x_1$, Gauss obtains the general relation

$$x_1 = n_1 + \mathbf{c}_1^{\mathrm{T}}\boldsymbol{\xi}, \tag{6}$$

which uniquely defines the vector $\mathbf{c}_1$. If[22]

$$\mathbf{a}_1^{(\dagger)} = \mathbf{A}\mathbf{c}_1,$$

then (5) and (6) imply that

$$\mathbf{a}_1^{(\dagger)\mathrm{T}}\mathbf{v} \equiv x_1 - n_1.$$

This equation shows that $\mathbf{a}_1^{(\dagger)}$ is a candidate for the minimizing vector. If $\boldsymbol{\kappa}$ is any other such vector, i.e., a vector satisfying (6), then

$$(\boldsymbol{\kappa} - \mathbf{a}_1^{(\dagger)})^{\mathrm{T}}\mathbf{v} \equiv n_1 - k.$$

Since $n_1 - k$ is independent of $\mathbf{x}$,

$$(\boldsymbol{\kappa} - \mathbf{a}_1^{(\dagger)})^{\mathrm{T}}\mathbf{A} = 0.$$

[22]The vector $\mathbf{a}^{(\dagger)\mathrm{T}}$ is the first row of the pseudo-inverse $\mathbf{A}^{\dagger}$ of $\mathbf{A}$ [35].

But $\mathbf{a}_1^{(\dagger)} = \mathbf{A}\mathbf{c}_1$. Hence

$$(\boldsymbol{\kappa} - \mathbf{a}_1^{(\dagger)})^{\mathrm{T}}\mathbf{a}_1^{(\dagger)} = 0,$$

and it follows that

$$\|\boldsymbol{\kappa}\|^2 = \|\mathbf{a}_1^{(\dagger)}\|^2 + \|\boldsymbol{\kappa} - \mathbf{a}_1^{(\dagger)}\|^2,$$

where $\|\cdot\|$ is the usual Euclidean norm. This establishes the minimality of $\|\mathbf{a}_1^{(\dagger)}\|^2$. Gauss shows that $\|\mathbf{a}_1^{(\dagger)}\|^2$ is equal to $c_{11}$ — the first component of the vector $\mathbf{c}_1$. He calls the estimate $k$ *the most reliable value*[23] of $x_1$.

To establish the connection with least squares, Gauss first points out that the general solution is obtained by solving the system (5) with $\boldsymbol{\xi} = 0$ for $\mathbf{x}$. But the vector $2\boldsymbol{\xi}$ is the gradient of the residual sum of squares with respect to $\mathbf{x}$. Hence this choice of $\mathbf{x}$ minimizes the residual sum of squares.

Next Gauss writes down the general relation

$$\mathbf{x} = \mathbf{n} + \mathbf{C}\boldsymbol{\xi}, \tag{7}$$

and concludes, as he had already done in the *Theoria Motus,* that the mean square errors in the components of $\mathbf{x}$ are proportional to the diagonal elements of $\mathbf{C}$.

The idea of searching for optimal linear combinations is due to Laplace. But it is important to recognize that Laplace and Gauss sought different kinds of combinations. Laplace's coefficients were to be used to combine the equations to form a square system — the $\mathbf{N}^{\mathrm{T}}\mathbf{A}$ of our earlier discussions. Gauss's $\boldsymbol{\kappa}$ is a combination of observations that estimates the first unknown. In a very real sense, Gauss's predecessors had been combining equations, while Gauss was the first to combine observations.[24]

## The Inversion of Linear Systems

In deriving his results, Gauss proceeds with great economy: the introduction of matrices in the above description does not shorten the argument; it merely saves the trouble of writing out Gauss's equations in scalar form. There are two closely related reasons for Gauss's success in dealing with his general linear model. First, Gauss chose his notation carefully so that relations between entire sets equations

[23]Actually, *valor maxime plausible.* The term used here is a translation of the German *sicherst wert,* which is what Gauss called it in German.

[24]For a single unknown, of course, the distinction vanishes.

stand out. For example, in the original the equations (3), (5), and (7) take the forms

$$\left.\begin{aligned} v &= ax + by + cz + \text{etc.} + l \\ v' &= a'x + b'y + c'z + \text{etc.} + l' \\ v'' &= a''x + b''y + c''z + \text{etc.} + l'' \text{ etc.} \end{aligned}\right\} \text{(I)}$$

$$\left.\begin{aligned} \xi &= x\Sigma aa + y\Sigma ab + z\Sigma ac + \text{etc.} + \Sigma al \\ \eta &= x\Sigma ab + y\Sigma bb + z\Sigma bc + \text{etc.} + \Sigma bl \\ \zeta &= x\Sigma ac + y\Sigma bc + z\Sigma cc + \text{etc.} + \Sigma cl \text{ etc.} \end{aligned}\right\} \text{(III)}$$

and

$$\left.\begin{aligned} x &= A + [\alpha\alpha]\xi + [\alpha\beta]\eta + [\alpha\gamma]\zeta + \text{etc.} \\ y &= B + [\beta\alpha]\xi + [\beta\beta]\eta + [\beta\gamma]\zeta + \text{etc.} \\ z &= C + [\gamma\alpha]\xi + [\gamma\beta]\eta + [\gamma\gamma]\zeta + \text{etc.} \\ &\text{etc.} \end{aligned}\right\} \text{(VII)}$$

The use of $\Sigma aa$, $\Sigma ab$, etc. to denote inner products among the coefficients in (I) is a typical nicety. In general, a careful attention to notation is a hallmark of the *Theoria Combinationis.*

But good notation alone is not enough. A more basic reason for Gauss's success is that he had a clear notion of the operation of inversion, as the passage from (III) to (VII) demonstrates. Gauss uses the phrase *eliminatio indefinite* for this operation. It refers to the process of obtaining the coefficients of the inverse of a linear system, either conceptually as above or numerically. An example is the definition of $\mathbf{c}_1$ in (6). The process is contrasted with *eliminatio definite*, which refers to the solution of a linear system by elimination. Thus, although Gauss did not possess the concept of a matrix, he had at hand one of the most important matrix operations — inversion — and he used it to good effect.

Throughout the *Theoria Combinationis* Gauss shows his ease with ideas of linear algebra. For example, he proves (*Supplementum,* Art. 21) that the inverse of a symmetric system of equations is symmetric. In Art. 4 of the *Supplementum* he makes a statement that is equivalent to saying that the null space of the transpose of a matrix is the orthogonal complement of its column space. However, Gauss's strongest point in this regard was his ability to devise elegant, efficient algorithms to implement the procedures of least squares.

### Gaussian Elimination and Numerical Linear Algebra

Gauss introduced Gaussian elimination in the *Theoria Motus* as a theoretical tool to show that the weights of his most probable values were inversely proportional to the diagonals of $\mathbf{C}$, which was to be computed by "common elimination" (*eliminatio vulgaris*). Nonetheless, Gauss clearly knew the computational significance of his formulas, since he says he will describe a better algorithm on another occasion. The other occasion was his *Disquisitio de elementis ellipticis Palladis* [4], where he gave the algorithm described above with detailed formulas and a nontrivial example. The formulas amount to what we would today call the outer-product form of Gaussian elimination.

Gauss described his algorithm in terms of quadratic forms—a practice to be followed by many nineteenth-century mathematicians, who developed most of the canonical matrix decomposition as transformations of bilinear and quadratic forms.[25] Gauss's quadratic form was the residual sum of squares, which he denoted by $\Omega$ and regarded as a function of $\mathbf{x}$. His derivation is as follows. Set

$$u_1 = \tfrac{1}{2}\tfrac{d\Omega}{dx_1} = s_1 + r_{11}x_1 + r_{12}x_2 + \cdots + r_{1p}x_p.$$

Then

$$\Omega_1 = \Omega - \tfrac{u_1^2}{r_{11}}$$

is easily seen to be independent of $x_1$. Repeat the reduction by writing

$$u_2 = \tfrac{1}{2}\tfrac{d\Omega_1}{dx_2} = s_2 + r_{22}x_2 + r_{23}x_3 + \cdots + r_{2p}x_p,$$

so that

$$\Omega_2 = \Omega_1 - \tfrac{u_2^2}{r_{22}}$$

is independent of $x_1$ and $x_2$. Continuing in this manner he arrived at a decomposition

$$\Omega = \tfrac{u_1^2}{r_{11}} + \tfrac{u_2^2}{r_{22}} + \cdots + \tfrac{u_p^2}{r_{pp}} + M, \tag{8}$$

[25] Gauss's reduction was anticipated in 1759 by Lagrange [19], whose concern was not with linear equations, but with sufficient conditions for a minimum of a functions. In modern terminology, he shows that if Gaussian elimination is performed on a symmetric matrix and the pivot elements remain positive, then the matrix is positive definite. Lagrange, it is true, treated only the $2 \times 2$ and $3 \times 3$ cases; but he explicitly points out that it is clear how to proceed with matrices of higher order. There is no indication that this work of Lagrange influenced Gauss, or anyone else.

in which $u_i$ is independent of $x_1, \ldots, x_{i-1}$ and $M$ is independent of $\mathbf{x}$. It follows that $\Omega$ is minimized when

$$u_1 = u_2 = \cdots = u_\rho = 0, \tag{9}$$

and the minimum value is $M$. The values of the $x_i$ can be determined by solving the triangular system of equations (9) in reverse order.

Although Gauss always presented his algorithm as a decomposition of a sum of squares, it will be useful to recast it in matrix form. In the notation used here, we can write $\Omega$ in the form

$$\Omega = (\mathbf{x}^{\mathrm{T}}\ 1)\begin{pmatrix} \mathbf{B} & \mathbf{m} \\ \mathbf{m}^{\mathrm{T}} & \mathbf{l}^{\mathrm{T}}\mathbf{l} \end{pmatrix}\begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix}.$$

If Gaussian elimination is performed on the matrix of this quadratic form, we get an upper triangular matrix

$$\begin{pmatrix} \mathbf{R} & \mathbf{s} \\ 0 & M \end{pmatrix},$$

which satisfies

$$\begin{pmatrix} \mathbf{B} & \mathbf{m} \\ \mathbf{m}^{\mathrm{T}} & \mathbf{l}^{\mathrm{T}}\mathbf{l} \end{pmatrix} = \begin{pmatrix} \mathbf{R}^{\mathrm{T}} & 0 \\ \mathbf{s}^{\mathrm{T}} & M \end{pmatrix}\begin{pmatrix} \mathbf{D}^{-1} & 0 \\ 0 & M^{-1} \end{pmatrix}\begin{pmatrix} \mathbf{R} & \mathbf{s} \\ 0 & M \end{pmatrix},$$

where

$$\mathbf{D} = \operatorname{diag}(r_{11}, r_{22}, \ldots, r_{\rho\rho}).$$

From this it is easily seen that Gauss's functions $u_i$ are the components of the vector

$$\mathbf{u} = \mathbf{R}\mathbf{x} + \mathbf{s},$$

and that the equations (9) are equivalent to the triangular system

$$\mathbf{R}\mathbf{x} = -\mathbf{s}.$$

This is the state of the algorithm before the *Pars Posterior*. It provides for the efficient solution of the normal equations and gives the variance of $x_\rho$ as well as the residual sum of squares. The matrix form of the algorithm makes it clear that by working with the inhomogeneous quadratic form $\Omega$ Gauss was effectively decomposing the augmented cross-product matrix—a standard practice in present-day regression computations.

Gauss begins Art. 31 of the *Pars Posterior* with a brief description of the algorithm. He then points out that the general elimination to find all the variances requires a great deal more work. Since the variance of the last unknown in the decomposition (8) can be found by inspection (Gauss goes on), it has become common practice for calculators to repeat the algorithm with other quantities in the last position.[26] He then says he will derive a better method (in words that echo his diary entry; see page 209).

The better method (Arts. 32–33) amounts to writing $\mathbf{C} = \mathbf{B}^{-1}$ in the form $\mathbf{T}^{\mathrm{T}}\mathbf{D}^{-1}\mathbf{T}$, where $\mathbf{T} = \mathbf{D}\mathbf{R}^{-\mathrm{T}}$. Thus if we know $\mathbf{T}$, we can compute the diagonal elements of $\mathbf{C}$ by taking weight sums of squares of the components of the rows of $\mathbf{T}$.

Unfortunately, Gauss cannot invert $\mathbf{R}$ to get $\mathbf{T}$ without an explicit system of equations and he must proceed indirectly. He begins with the equation

$$\mathbf{u} = \mathbf{R} + \mathbf{s},$$

and by considering the differential of one of the forms of $\Omega$ derived earlier, he arrives at

$$\boldsymbol{\xi} = \mathbf{R}^{\mathrm{T}}\mathbf{D}^{-1}\mathbf{u}. \tag{10}$$

He now defines $\mathbf{T}$ by inverting the relation to get

$$\mathbf{u} = \mathbf{T}\boldsymbol{\xi}. \tag{11}$$

Finally, by considering the differential of another form of $\Omega$, he gets

$$\mathbf{x} = \mathbf{T}^{\mathrm{T}}\mathbf{D}^{-\mathrm{T}}\mathbf{u} + \hat{\mathbf{x}}, \tag{12}$$

where the components of $\hat{\mathbf{x}}$ are the most reliable values of the unknowns. Substituting (11) into (12) gives the desired relation.

It remains to compute $\mathbf{T}$. By playing (10) against (11), Gauss arrives at two algorithms, one that generates $\mathbf{T}$ by rows and one that generates it by columns. He notes that the second is preferable if one only needs the weights of a few of the first unknowns.

Gauss does not mention that one can compute $\mathbf{T}$ by rows from the other end, which gives a very inexpensive algorithm for computing the weights of the last few

[26] In fact, Laplace recommends a similar, though more sophisticated, procedure in the first supplement to the third edition of his *Théorie Analytique des Probabilités* [27].

unknowns. However, his final result along these lines works equally well (Art. 34). Specifically, he proves the following theorem. If the linear function $\mathbf{t}^{\mathrm{T}}\mathbf{x} + k$ is written in the form $\mathbf{v}^{\mathrm{T}}\mathbf{u} + K$, then $K$ is the least squares estimate of $k$ and the variance of the estimate is proportional to $\mathbf{v}^{\mathrm{T}}\mathbf{D}\mathbf{v}$. He also shows that the vector of coefficients $\mathbf{v}$ satisfies the triangular system $\mathbf{R}^{\mathrm{T}}\mathbf{v} = \mathbf{t}$, from which the components of $\mathbf{v}$ can be determined by back substitution. Thus, the variance of an arbitrary linear function can be computed by solving a triangular system and computing a weighted sum of squares.

An important theme in modern numerical linear algebra is the use a single decomposition as a tool to perform a variety of computations. In this respect Gauss was thoroughly modern. The elements of $\mathbf{R}$, which are the final product of his elimination algorithm, are the basis for all the computations described here. With the introduction in the *Supplementum* (Art. 13) of the inner product forms of the elimination algorithm, Gauss will have said just about everything there is to say on algorithms for solving dense, positive-definite systems.

Do these algorithms represent an improvement over the practices of the time? If we assume that general elimination was done by what now would be called Gauss–Jordan elimination, the calculator would perform roughly $\frac{1}{2}\rho^3$ additions and multiplications to solve a system and $\frac{5}{6}\rho^3$ to invert it. The operation count for Gaussian elimination is $\frac{1}{6}\rho^3$ — twice that if all the variances are required. Thus Gauss's procedures reduce the amount of work by a factor of two and a half to three. Of course the overhead for forming the normal equations, which may dominate, is the same in both cases.

Gauss partly anticipated another theme of modern matrix computations: the updating of previously computed quantities when the problem is changed. Specifically, he considers two problems. The first is to determine new estimates and variances from old ones when a new equation is added. The second is to do the same thing when the weight of an equation is changed. However, Gauss falls short of modern practice here. Today we update entire decompositions, which allows the updating to continue indefinitely. Gauss, however, updates only the solutions and must repeat the entire calculation if more than one change is made.

Finally, Gauss may have been the first to solve linear systems iteratively. In Arts. 18–20 of the *Supplementum,* he introduces what we today would call a block Gauss–Seidel iteration. The idea is to partition the equations (actually constraints) into groups and adjust the estimates cyclically to satisfy one group after another. Gauss does not prove the convergence of his method, and he suggests that making a good choice of groups is something of an art. Although he does not mention it, he must surely have known that after the first round of adjustments the results of the eliminations used to solve the reduced systems can be reused to reduce computations in subsequent iterations.

Gauss's influence on future generations of computers was mixed. Gauss was probably the one who popularized the idea of reducing quadratic forms. In its various generalizations, it lead to the discovery of many matrix decompositions before the widespread use of matrices themselves.[27] For example, Jacobi [17] decomposed a biquadratic form a la Gauss — the nonsymmetric variant of Gaussian elimination.

Gauss's method for solving positive-definite systems had lasting influence. Gaussian elimination in Gauss's notation continued to appear in astronomical and geodetic textbooks well into the twentieth century. Thereafter, his notation fell into eclipse, and the name Gaussian elimination began to be applied to any reduction of a system, symmetric or nonsymmetric, to triangular form followed by back substitution. Other names — e.g., Cholesky, Crout, Doolittle — became attached to specific variants, some of which had been anticipated by Gauss himself. The subject is complicated by the fact that the method, taken generally, was often rediscovered and presented as new [16, p. 141].

The Gauss–Seidel iterative method and its variants has also had a long and fruitful career. Because of its low storage requirements and its repetitive nature it came into its own in the early days of the digital computer and has not yet been entirely superseded.

The fact that Gauss treated his reduced form as we would a matrix decomposition is due more to his computational parsimony than to any conscious realization of a general decompositional approach. The use of matrix decompositions emerged gradually in the 1950s and 1960s, partly as a result of the recasting of algorithms, previously described by tableaus, in terms of matrices and partly because the computer made it practical to calculate a wider variety of decompositions. Similarly, the notion of updating arose in connection with the simplex algorithm for linear programming, with no reference to Gauss.

[27]Kline [18, Ch. 33] gives a concise survey.

### The Generalized Minimum Variance Theorem

Gauss called the least squares estimates of his unknowns the *valores maxime plausibles,* which I have translated by the phrase "most reliable values." They are praiseworthy and reliable because they have minimum variances. Now if $t = \mathbf{t}^{\mathrm{T}}\mathbf{x} + k$ is any linear function of the unknowns, it is natural to ask for its most reliable value. Today we know that it is the value obtained by evaluating $t$ at the most reliable values of the unknowns $\mathbf{x}$ (a value which Gauss denoted by $K$), and it has been asserted that this result is due to Gauss himself [37]. In this section we will attempt to untangle what Gauss knew about the estimation of linear functions.

Gauss certainly did not prove the generalization in the *Pars Posterior*. He is initially quite careful to call the number $K$ the value of $t$ at the most the value of $\mathbf{x}$ (*valor ipsius t, e valoribus maxime plausibilibus ipsarum x, y, z, etc. prodiens*). Later he refers to it simply as the most reliable value, thereby conferring optimality by fiat. But nowhere does he give a proof, though he does calculate the variance of $K$. We must look to the *Supplementum* for further work on the optimality of estimates.

The *Supplementum* is a separate memoir devoted to the application of least squares to geodesy. Here the problem is not to estimate unknowns in an overdetermined set of equations but to adjust data to satisfy constraints. For example, the measured angles of a triangle must be adjusted to sum to 180 degrees (with a correction for spherical excess). This memoir is given less than its due in the secondary literature, partly because Gauss's commentators tend to run out of steam at this point, but also because Gauss's oblique style is particularly demanding here. Nonetheless, it contains some of Gauss's finest work on least squares.

Gauss sets the scene by reviewing the problem treated in the previous memoir (Art. 1). He then points out that in some applications there are no unknowns to be determined; instead the data must satisfy certain conditional equations.

Gauss then states that the new problem can be reduced to the old. The exposition is not very clear, but Gauss seems to mean the following.[28]

Let there be $\pi$ measurements and $\sigma$ conditional equations. Let $\mathbf{v}$ be the vector of measurements and let the conditional equations be written in the form

[28]Some confirmation of this interpretation can be found in a letter dated February 1824 from Gauss to Gerling [36, No. 165]. The procedure sketched there leads to the normal equations of the system (13).

$\mathbf{a}_0 + \mathbf{A}^{\mathrm{T}}\mathbf{v} = 0$. Choose a set of $\pi - \sigma$ of the measurements, say the first, as unknowns, and partition $\mathbf{A}^{\mathrm{T}} = (\mathbf{A}_1^{\mathrm{T}} \ \mathbf{A}_2^{\mathrm{T}})$, where $\mathbf{A}_2$ is square and assumed to be nonsingular. Partition $\mathbf{v}^{\mathrm{T}} = (\mathbf{v}_1^{\mathrm{T}} \ \mathbf{v}_2^{\mathrm{T}})$ conformally. Then Gauss's equivalent procedure amounts to solving the least squares problem of minimizing the sum of squares of the components of the vector

$$\begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 + \mathbf{A}_2^{-\mathrm{T}}\mathbf{a}_0 \end{pmatrix} - \begin{pmatrix} \mathbf{I} \\ -\mathbf{A}_2^{-\mathrm{T}}\mathbf{A}_1^{\mathrm{T}} \end{pmatrix} \hat{\mathbf{v}}_1 \tag{13}$$

and setting $\hat{\mathbf{v}}_2 = -\mathbf{A}_2(\mathbf{a}_0 + \mathbf{A}_1^{\mathrm{T}}\hat{\mathbf{v}}_1)$, so that $\hat{\mathbf{v}}$ satisfies the constraints.

There are two reasons why Gauss's meaning is unclear. In the first place his procedure would seem to depend on the choice of $\mathbf{A}_1$; however, it is easy to see that the problem is equivalent to that of minimizing the sum of squares of the components of $\hat{\mathbf{v}} - \mathbf{v}$ subject to $\mathbf{a}_0 + \mathbf{A}^{\mathrm{T}}\hat{\mathbf{v}} = 0$, which is independent of the choice of $\mathbf{A}_1$. But Gauss does not say this.

A more important reason is that Gauss has not yet stated an optimality criterion for the problem of adjusting observations. His equivalent least squares procedure will give an adjustment of the data that satisfies the conditional equations, but no reason is given for preferring this adjustment over another. The complete answer only emerges in Art. 9. In the meantime, the reader is forced to shift gears in the next article as Gauss appears to change the problem.

Gauss states his problem as follows (Art. 2). Let $u = \kappa + \mathbf{l}^{\mathrm{T}}\mathbf{v}$ be a linear function[29] of $\mathbf{v}$. If $\mathbf{v}$ is in error by $\mathbf{e}$ then the value of $u$ at $\mathbf{v}$ will be in error by $\mathbf{l}^{\mathrm{T}}\mathbf{e}$; and if the components of $\mathbf{e}$ have mean error[30] $m$, then the error in $u$ has mean error $m\|\mathbf{l}\|$.

If nothing further is known about $\mathbf{v}$, we must accept $u(\mathbf{v})$ as an estimate of the true value of $u$. But if we know that $\mathbf{v}$ satisfies a system of linear constraints

$$\mathbf{a}(\mathbf{v}) \equiv \mathbf{a}_0 + \mathbf{A}^{\mathrm{T}}\mathbf{v} = 0, \tag{14}$$

then we may estimate the true value of $u$ by any function $\hat{u} = u_0 + \hat{\mathbf{l}}^{\mathrm{T}}\mathbf{v}$ with the property that $u(\mathbf{v}) - \hat{u}(\mathbf{v}) = 0$ whenever $\mathbf{v}$ satisfies (14). Since the mean error in

[29]As always Gauss starts from a nonlinear problem and linearizes. For example, here the function is $u$ and the components of $\mathbf{l}$, written $l$, $l'$, $l''$, etc., are its derivatives.

[30]Gauss allows for errors with different mean errors and adjusts his formulas by inserting suitable weights. For simplicity we will stick with the case where all the errors have the same mean error $m$.

$\hat{u}(\mathbf{v})$ is $m\|\hat{\mathbf{l}}\|$, the problem of finding a best estimate of the true value of $u$ is one of choosing $\hat{\mathbf{l}}$ so that $\|\hat{\mathbf{l}}\|$ is minimized.

There is no need to follow Gauss's solution of this problem (Arts. 4–6), which is in some sense a way station. For Gauss goes on (Arts. 8–9) to show that there are corrections $\boldsymbol{\epsilon}$ of the observations $\mathbf{v}$ such that $\mathbf{a}(\mathbf{v}-\boldsymbol{\epsilon}) = 0$ and such that for any function $u$ we have $u(\mathbf{v}-\boldsymbol{\epsilon}) = \hat{u}(\mathbf{v})$. Moreover, among all corrections causing $\mathbf{v}$ to satisfy the conditional equations, $\boldsymbol{\epsilon}$ has minimum norm. Gauss calculates the corrections in the form $\mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{a}(\mathbf{v})$, the inverse of $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ being introduced as usual by inverting a system of equations.

To put this remarkable result more concisely, Gauss has shown that if $\hat{\mathbf{v}}$ is the vector satisfying the conditional equations that is nearest $\mathbf{v}$ in the least squares sense then the most reliable estimate of any linear function of $\mathbf{v}$ is its value at $\hat{\mathbf{v}}$. Thus, for the problems of adjusting data, Gauss proved a general minimum variance theorem that applies to all functions of the data.

Gauss is also very close to generalizing the original minimum variance theorem of the *Pars Prior* and proving that the most reliable estimate of any linear function of the unknowns is its value at the most reliable values of the unknowns. All he has to do is reverse what he did in the first article of the *Supplementum* and show that for every ordinary least squares problem there is an equivalent constrained least squares problem of the kind treated in the *Supplementum*. In retrospect it is easy to see how to do this. The observations must be constrained to lie in the column space of the least squares matrix.[31]

Whether Gauss made this connection is not easily decided. He never explicitly states that the results of the *Supplementum* imply that the most reliable value of a linear function of the unknowns is its value at the most reliable values of the unknowns. But in the Notice to the *Supplementum* he says rather cryptically:

> The solution itself, which is self-contained and independent of the earlier memoir, is in turn so adaptable that in many cases it is advantageous to use it even when the basic assumptions of the first method hold.

Perhaps Gauss was referring only to systems that can be easily formulated in both

[31]Specifically, if the columns of $\mathbf{U}$ form a basis for the orthogonal complement of the column space of the least matrix $\mathbf{A}$, then the true values of the observations must satisfy $\mathbf{U}^{\mathrm{T}}\mathbf{v} = 0$, in which case the adjusted values $\hat{\mathbf{v}}$ are just the values obtained by ordinary least squares. The most reliable value of the function $t = \mathbf{t}^{\mathrm{T}}\mathbf{x} + k \equiv \mathbf{t}^{\mathrm{T}}\mathbf{A}^{\dagger}\mathbf{v} + k$ will then be $\hat{t} = \mathbf{t}\mathbf{A}^{\dagger}\hat{\mathbf{v}} + k = \mathbf{t}\hat{\mathbf{x}} + k$.

ways. Arguably, he was aware of the general equivalence of the two problems and therefore was in possession of the generalized minimum variance theorem. What is certain is that Gauss can be as enigmatic to us as he was to his contemporaries.

*This page intentionally left blank*

# References


[1] Roger Joseph Boscovich. De litteraria expeditione per pontificiam ditionem, et synopsis amplioris operis, ac habentur plura ejus ex exemplaria etiam sensorum impressa. *Bononiensi Scientarum et Artum Instituto Atque Academia Commentarii*, IV:353–396, 1757. Cited in [2].

[2] C. Eisenhart. Boscovich and the combination of observations. In L. L. Whyte, editor, *Roger Joseph Boscovich*, pages 200–212. Fordham University Press, New York, 1961.

[3] Carl Friedrich Gauss. *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium.* Perthes and Besser, Hamburg, 1809. Cited and reprinted in [14, v. 7, pp. 1–261]. English translation by C. H. Davis [13]. French and German translations of Book II, Part 3 in [12, 15].

[4] Carl Friedrich Gauss. Disquisitio de elementis ellipticis Palladis. *Commentatines societatis regiae scientarium Gottingensis recentiores*, 1, 1810. Cited and reprinted in [14, v. 6, pp. 1–64]. French translation of §§13–14 in [12]. German translation of §§10–15 in [15].

[5] Carl Friedrich Gauss. Bestimmung der Genauigkeit der Beobachtungen. *Zeitschrift für Astronomie*, 1, 1816. Cited and reprinted in [14, v. 4, pp. 109–117]. French translation in [12].

[6] Carl Friedrich Gauss. Anzeige: Theoria combinationis observationum erroribus minimis obnoxiae: Pars prior. *Göttingische gelehrte Anzeigen*, 33:321–327, 1821. Cited and reprinted in [14, v. 4, pp. 95–100].

[7] Carl Friedrich Gauss. Anzeige: Theoria combinationis observationum erroribus minimis obnoxiae: Pars posterior. *Göttingische gelehrte Anzeigen*, 32:313–318, 1823. Cited and reprinted in [14, v. 4, pp. 100–104].

[8] Carl Friedrich Gauss. Theoria combinationis observationum erroribus minimis obnoxiae: Pars prior. *Commentatines societatis regiae scientarium Gottingensis recentiores*, 5, 1823. Cited and reprinted in [14, v. 4, pp. 1–26]. French and German translations in [12, 15].

[9] Carl Friedrich Gauss. Theoria combinationis observationum erroribus minimis obnoxiae: Pars posterior. *Commentatines societatis regiae scientarium Gottingensis recentiores*, 5, 1823. Cited and reprinted in [14, v. 4, pp. 27–53]. French and German translations in [12, 15].

[10] Carl Friedrich Gauss. Anzeige: Supplementum theoriae combinationis observationum erroribus minimis obnoxiae. *Göttingische gelehrte Anzeigen*, 153:1521–1527, 1826. Cited and reprinted in [14, v. 4, pp. 104–108].

[11] Carl Friedrich Gauss. Supplementum theoriae combinationis observationum erroribus minimis obnoxiae. *Commentatines societatis regiae scientarium Gottingensis recentiores*, 6, 1828. Cited and reprinted in [14, v. 4, pp. 55–93]. French and German translations in [12, 15].

[12] Carl Friedrich Gauss. *Méthode des Moindres Carres*. Ballet–Bachelier, Paris, 1855. Translation by J. Bertrand of various works of Gauss on least squares.

[13] Carl Friedrich Gauss. *Theory of the Motion of the Heavenly Bodies Moving about the Sun in Conic Sections.* Little, Brown, and Company, 1857. Translation by Charles Henry Davis of *Theoria Motus* [3]. Reprinted by Dover, New York, 1963.

[14] Carl Friedrich Gauss. *Werke.* Königlichen Gesellschaft der Wissenschaften zu Göttingen, 1870–1928.

[15] Carl Friedrich Gauss. *Abhandlungen zur Methode der kleinsten Quadrate.* P. Stankeiwica', Berlin, 1887. Translation by A. Borsch and P. Simon of various works of Gauss on Least Squares.

[16] A. S. Householder. *The Theory of Matrices in Numerical Analysis.* Dover Publishing, New York, 1964. Originally published by Ginn Blaisdell.

[17] C. G. J. Jacobi. Über eine elementare Transformation eines in Buzug jedes von zwei Variablen-Systemen linearen und homogenen Ausdrucks. *Journal für die reine und angewandte Mathematik*, 53:265–270, 1857, posthumous.

[18] M. Kline. *Mathematical Thought from Ancient to Modern Times.* Oxford University Press, New York, 1972.

[19] J.-L. Lagrange. Researches sur la métode de maximis et minimis. *Miscellanea Taurinensi*, 1, 1759. Cited and reprinted in [20, v. 1, pp. 1–16].

[20] J.-L. Lagrange. *Œvres de Langrange.* Gauthier–Villars, Paris, 1867–1892.

[21] P. S. Laplace. Mémoir sur la probabilité des causes par les évènements. *Memoires de l'Academie des Sciences de Paris*, 6, 1774. Cited and reprinted in [28, v. 8, pp. 25–65].

[22] P. S. Laplace. Théorie de Jupiter et de Saturne. *Memoires de l'Academie des Sciences de Paris*, 1785. Cited and reprinted in [28, v. 11, pp. 93–207].

[23] P. S. Laplace. Sur quelques points du systéme du monde. *Memoires de l'Academie des Sciences de Paris*, 1789. Cited and reprinted in [28, v. 11, pp. 475–558].

[24] P. S. Laplace. Mémoire sur les approximations des formules qui sont functions de trés grands nombres et sur leur application aux probabilités. *Memoires de l'Academie des Sciences de Paris*, 10, 1809. Cited and reprinted in [28, v. 12, pp. 300–353]. According to Todhunter [44, §914] the publication date was 1810.

[25] P. S. Laplace. Mémoire sur les intégrales définies et sur application aux probabilités. *Memoires de l'Academie des Sciences de Paris*, 11, 1810. Cited and reprinted in [28, v. 12, pp. 355–412]. According to Todhunter [44, §914] the publication date was 1811.

[26] P. S. Laplace. *Théorie Analytique des Probabilités.* Courcier, Paris, 1812. Facsimile edition by Impression Anastaltique Culture et Civilisation, Brussels, 1967.

[27] P. S. Laplace. *Théorie Analytique des Probabilités.* Courcier, Paris, third edition, 1820. Reprinted in [28, v. 7].

[28] P. S. Laplace. *Œvres Compeétes.* Gauthier–Villars, Paris, 1878–1912.

[29] A. M. Legendre. *Nouvelles méthodes pour la détermination des orbites des comètes.* Courcier, Paris, 1805. Cited in [42], where the appendix on least squares is reproduced.

[30] Christopher Maire and Roger Joseph Boscovich. *De Litteraria Expeditione per Pontificiam ditionem ad dimetiendas duas Meridiani gradus, et corrigendam mappam geographicam, jussu, et auspiciis Benedicti XIV Pont. Max. suscepta.* Rome, 1755. Cited in [2]. French translation [31].

[31] Christopher Maire and Roger Joseph Boscovich. *Voyage Astronomique et Gógraphique dans l'État de lÉglise, entrepris par l'Ordre et sous les Auspices du Pape Benoøît XIV, pour mesure deux degrés du méridien, et corriger la Carte de l'État ecclesiastique.* Rome, 1770. Cited in [2]. This translation of [30] also contains a supplement by Boscovich to [40], in which he gives his algorithm for $\ell 1$ approximation.

[32] T. Mayer. Abhandlung über de Umwalzung des Monds um seine Axe und die scheinbare Bexegung der Mondsflectem. *Kosmographische Nachrichted un Sammulungen auf das Jahr 1748*, pages 52–183, 1748. Cited in [42].

[33] L. S. Multhaus. Olbers, Heinrich William Mathias. In C. C. Gillispe, editor, *Dictionary of Scientific Biography, X*, pages 197–199. Charles Scribner's Sons, New York, 1976.

[34] R. L. Plackett. The discovery of the method of least squares. *Biometrika*, 59:239–251, 1972.

[35] D. W. Robinson. Gauss and generalized inverses. *Historia Mathematica*, 7:118–125, 1980.

[36] C. Schaefer, editor. *Briefwechsel zwischen Carl Friedrich Gauss und Christian Ludwig Gerling.* Elsner, Berlin, 1927.

[37] H. L. Seal. The historical development of the Gauss linear model. *Biometrika*, 54:1–24, 1967.

[38] O. B. Sheynin. Laplace's theory of errors. *Archive for History of Exact Sciences*, 17:1–61, 1977.

[39] O. B. Sheynin. C. F. Gauss and the theory of errors. *Archive for History of Exact Sciences*, 20:21–72, 1979.

[40] Benedict Stay. *Philosophiae Recentioris, a Benedicto Stay in Romano Archigynasis Publico Eloquentare Professore, Versibus Traditae, Libri X, cum Adnotationibus et Supplementis P. Rogerii Josephi Boscovich S.J.* Rome, 1760. Cited in [2].

[41] G. W. Stewart. Gauss, statistics, and Gaussian elimination. Technical Report CS-TR-3307, Department of Computer Science, University of Maryland, 1994. To appear in *Journal of Computation and Graphical Statistics.*

[42] S. M. Stigler. *The History of Statistics.* Belknap Press, Cambridge, Massachusetts, 1986.

[43] S. M. Stigler. Gauss and the invention of least squares. *Annals of Statistics*, 9:465–474, 1991.

[44] I. Todhunter. *A History of the Mathematical Theory of Probability from the Time of Pascal to that of Laplace.* G. E. Stechert, New York, 1865. Reprint 1931.