Utrecht University

Gauss's method of least squares applied to geodetic measurements in Friesland

Bachelor Thesis, Mathematics

Tom Reijngoudt

Supervisor:
Jan Hogendijk

Utrecht, June 2018

Contents

Contents **iii**

1 Introduction **1**

2 The geodetic problem **3**

3 The equations **6**
3.1 The numerical equations for the most plausible errors . . . 7
3.1.1 Deriving the two equations of third kind . . . 8
3.1.2 Linearization of the logarithm functions . . . 9
3.2 Spherical excess . . . 10

4 Gauss's solution to the problem **11**
4.1 Solving the system of equations . . . 11
4.2 Correctness of the method and its result . . . 15
4.3 The idea behind the solution method in modern terms . . . 16
4.4 The normal distribution . . . 17

Bibliography **19**

Chapter 1

Introduction

Born on 30 April 1777 in Braunschweig, Carl Friedrich Gauss was a German mathematician who contributed to many different branches of mathematics. At the age of 24, when Gauss had just finished his doctoral degree, he became a well established name as an academic, due to his work on the prediction of the path of the dwarf planet Ceres [Kaufmann-Bühler, p. 45]. After its discovery on February 11, 1801 by Piazzi, Ceres was observed for 41 days after which it disappeared into the halo of the sun. Gauss discovered a method for computing the planet's orbit and successfully predicted where Ceres might be found after its exit from the halo.

Gauss continued working on planetary orbits and in 1809 published his *Theoria motus corporum coelestium in sectionibus conicis solem ambientium* (Theory of motion of the celestial bodies moving in conic sections around the Sun) [Gauss 1809, Gauss 1857]. In this work, Gauss used the technique of least squares and gave some arguments for this method which will be discussed in Ch. 4.4 below. Later on he extended on this in a more fundamental theoretical way in his two memoirs *Theoria Combinationis Observationum Erroribus Minimis Obnoxiae* (Theory Of The Combination Of Observations Least Subject To Errors) [Gauss 1823, 1823a] with the *Supplement* [Gauss 1828] which we will study in this thesis. Gauss wrote in Latin but the work was translated into German [Gauss 1887, pp. 1-91] and English [Gauss 1995].

The investigations in the *Supplement* were inspired by his work on geodesy, which mainly concerned the triangulation of the province of Hannover [Kaufmann-Bühler 1981, p. 95-109]. Using new and accurate observational tools, and the method of repetition, measurements became a lot more precise. However, Gauss realized that measurements can never produce perfect values, as he notes in the beginning of [Gauss 1823]: "However carefully one takes observation of the magnitude of objects in nature, the results are always subject to larger or smaller errors" [Gauss 1995, p. 3]. Throughout his life he studied the nature of these errors, and in doing so also discovered the 'Gaussian' normal distribution which plays an important role in probability theory.

In the *Supplement*, he gives an extension and application for (some of) the discussed least squares theory to geodesy. He starts with a very general problem setting, and his solutions are not reader-friendly because he omitted many steps which he found easy, and because his notation is difficult for a modern reader. When reading the *Supplement*, it is often not clear what Gauss is discussing or how he achieves his results. The translator Stewart also calls Gauss's style in the *Supplement* "oblique" and "particularly demanding" [Gauss 1995, p. 232]. Then Gauss illustrates his very general theories by a few concrete examples. The *Supplement* has not been discussed in much detail in the literature on the history of mathematics; the existing reports are brief and incomplete, and were written for specialists in statistics [Gauss 1995, pp. 232-235] or for geodesists, see for example [Jordan-Eggert 1948, p. 519-520].

In this thesis I have tried to illustrate a small part of the mathematics in the *Supplement* by turning Gauss's procedure around. I start with his first example, which consists of some geodetic measurements that had been made by G. Krayenhoff (1758-1840) in the Netherlands. From the work *Précis historique des opérations trigonométriques faites en Hollande*, Gauss took 27 measurements of angles between the cities Ballum, Harlingen, Leeuwarden, Sneek, Oldeholtpade, Drachten, Oosterwolde, Groningen and Dokkum. To prepare for this thesis, we emulated the field work done by Krayenhof by visiting most of these cities, to get an idea of the difficulty of making these observations in an precise manner. We climbed the tower of the “Oldehove” in Leeuwarden and inspected the location from which five measurements were made.

Because measurements done in the real world will always translate to some mathematical inaccuracy, Gauss wanted to test their accuracy, by comparing them with known mathematical properties. For instance the sum of the angles of a plane triangle is 180°, and the sum of all angles surrounding a point is 360°. Because such requirements are not exactly met by the measured angles, these angles need to be adjusted a bit. Krayenhoff had done this in rather arbitrary ways, but Gauss did this in a better way by means of least squares, which method was not known to Krayenhoff. I will explain the theory in the *Supplement* [Gauss 1828] only in the case of this example. Thus we will see only a small part of the beautiful mathematics in the *Supplement* but we can avoid many of the generalities which make Gauss’s own argument so difficult. In this way I try to make Gauss’s reasoning accessible to students who have finished their first year in mathematics. In an earlier work published in 1809, Gauss had given an interesting motivation for the method of least squares, see Chapter 4.4 below.

Gauss used his conclusions to criticize Krayenhoffs work and methods [Haasbroek 1972, p. 9]. This interesting criticism falls outside the scope of this thesis, compare [Haasbroek 1972, pp. 85, 109, 131].

Chapter 2

The geodetic problem

Here is the first example to which Gauss applies his theories. These are 27 measurements (out of more than 500) made by C.R.T. Krayenhof in 1807 and 1811 and published in his *Précis historique des opérations trigonométriques faites en Hollande* [Krayenhoff 815, pp. 77-81, 113-115] The measurements were made in nine localities, almost all in Friesland, on top of church towers, by a high-precision instrument called "repetition circle". The instrument was used to measure angles between towers of other cities seen in the distance, usually between 20 and 40 km away.

Figure 1: Painting of Krayenhoff [Haasbroek 1972, p. 7, 13] with his repeating circle on the left.

The towers in the distance were sighted by the two telescopes with a magnification of more than 20. They could rotate over circular scales that could be read off by small microscopes which can also be seen in the painting in Figure 1. Krayenhoff and his assistants read the individual angles with an accuracy of 5 arc seconds [Haasbroek 1972 p. 29]. A special part of the instrument (repetition circle) with a set of screws then made it possible to repeat the same measurement ca. 20 times. During this repetition, the sum of the measured angles was automatically indicated on the circular scale, for details see Haasbroek [1972, p. 20-24]. The average of the 20 angles was considered more accurate than the individual angle; it was computed in thousands of arc seconds, although the last two digits are not significant from a modern point of view. The purpose of all this was to draw a very accurate map of the Netherlands.

The high precision values make the mathematics more difficult. One meter seen from a distance of 30 kilometer corresponds to an angle of ca. 6 arc seconds, so the height of the towers in the distance cannot be ignored. If one climbs a tower of 30 meters high, the visible horizon sinks ca. 10 arc minutes, because of the curvature of the earth, which therefore cannot be ignored either.[1] Therefore one had to make a trigonometrical computation in order to "reduce the angle

[1] One meter seen at a distance of 30 km corresponds to ca. $(1/30000) * (180/\pi) * 60 * 60$ arc seconds. The

(between the two towers) to the horizon." This means: to find the angle between the perpendicular projections of the two arms of the angle between the two towers, on the plane of the mathematical horizon, that is the plane through the observer perpendicular to the direction of the zenith. In order to do this computation one first needed to measure with the same precision, using the repetition circle, the angles between the towers in the distance and the zenith. These zenith angles are usually a bit more than 90 degrees. We will spare the reader the details, for all trigonometrical formulas see [Haasbroek 1972, pp 51-56].

The "angles reduced to the horizon" are equal to the angles on the sphere between the great circle arcs from the point of observation to the two towers in the distance. From now on we will call these spherical angles the "observed angles" or "measured angles", even though they are products of a trigonometrical computation.

Here is a list of the 27 "measured angles" (that is, true spherical angles) which Gauss took from Krayenhoff [Gauss 1995, p. 150-151]. Figure 2 displays the network of these measured angles. Gauss calls each angle by a number $0 \ldots 26$, and we will also use these numbers in our notation. We use $w_i, 0 \leq i \leq 26$ for the measured values of these (spherical) angles. Gauss also lists the nine triangles which can be made from the measured angles, as shown in Figure 2.

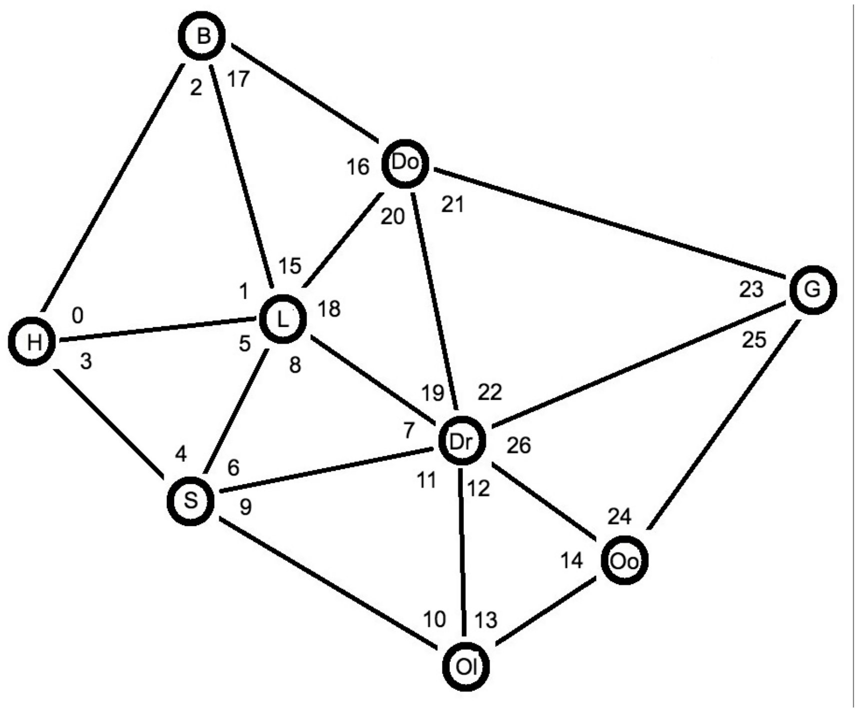

Figure 2. The network of 9 localities and 27 angles.

H = Harlingen	B = Ballum	Do = Dokkum
L = Leeuwarden	S = Sneek	Dr = Drachten
Ol = Oldeholtpade	Oo = Oosterwolde	G= Groningen

approximation formula for the depression of the visible horizon is $\sqrt{\frac{2h}{R}}$ radians, where h is the height of the tower, $R \approx 6371$ km the radius of the earth.

Triangle 1

w_0 = Harlingen . . . . 50°58′15.238″
w_1 = Leeuwarden . . . . 82°47′15.351″
w_2 = Ballum . . . . 46°14′27.202″

Triangle 2

w_3 = Harlingen . . . . 51°5′39.717″
w_4 = Sneek . . . . 70°48′33.445″
w_5 = Leeuwarden . . . . 58°5′48.707″

Triangle 3

w_6 = Sneek . . . . 49°30′40.051″
w_7 = Drachten . . . . 42°52′59.382″
w_8 = Leeuwarden . . . .87°36′21.057″

Triangle 4

w_9 = Sneek . . . . 45°36′7.492″
w_{10} = Oldeholtpade . . . . 67°52′0.048″
w_{11} = Drachten . . . . 66°31′56.513″

Triangle 5

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

Triangle 6

w_{15} = Leeuwarden . . . . 59°24′0.645″
w_{16} = Dokkum . . . . 76°34′9.021″
w_{17} = Ballum . . . . 44°1′51.040″

Triangle 7

w_{18} = Leeuwarden . . . . 72°6′32.043″
w_{19} = Drachten . . . .46°53′27.163″
w_{20} = Dokkum . . . . 61°0′4.494″

Triangle 8

w_{21} = Dokkum . . . .57°1′55.292″
w_{22} = Drachten . . . . 83°33′14.515″
w_{23} = Groningen . . . . 9°24′52.397″

Triangle 9

w_{24} = Oosterwolde . . . . 81°5417.447
w_{25} = Groningen . . . .31°5246.094
w_{26} = Drachten . . . . 66°1257.246

Gauss uses these 27 measured angles between 9 locations in order to give an application of the theory he discussed earlier in the *Supplement*. Gauss proceeds by setting up a system of equations, as discussed in the following section. The mathematical correct values of the angles should satisfy these equations but the measured values do not, and therefore Gauss computed adjustments of the measured values using the method of least squares. In this way he could also judge the reliability of the measurements. The reliability of measurements is relevant in this case because small deviations from the true angles can result in big changes over long distances.

Chapter 3

The equations

In the previous section we discussed the measured values of 27 spherical angles $w_0, w_1, \cdots w_{26}$ between 9 different localities, as in Figure 2 above. Gauss now sets up a system of equations. We will simplify his notations by the modern concept of functions.

There are certain mathematical properties the angles should satisfy, if the measurements could have been done perfectly. Thus, (1) the sum of the angles surrounding one angular point is 360 degrees, (2) the sum of the angles of a spherical triangle is 180 degrees plus the spherical excess, and (3) we will see that we can also derive another property of the angles from the sine rule in any triangle.

With these mathematical properties in mind, we look at the network of angles in Figure 2 above. We find two points completely surrounded by angles (leading to f_1 and f_2 below), we can make 9 triangles (this will lead to $f_3 \cdots f_{11}$ below), and as we will see, the sine rule can be exploited in two cases (leading to f_{12} and f_{13} below). We will use the notation $v_0, v_1, \ldots v_{26}$ for "perfect values" of the 27 angles, not subject to measurement errors.

Two angular points, namely Leeuwarden and Drachten, are completely surrounded by angles. This leads to two functions which must be zero at the perfect values v_i:

$$
\begin{aligned}
f_1(v_0, v_1, \ldots, v_{26}) &= v_1 + v_5 + v_8 + v_{15} + v_{18} - 360 \\
f_2(v_0, v_1, \ldots, v_{26}) &= v_7 + v_{11} + v_{12} + v_{19} + v_{22} + v_{26} - 360
\end{aligned}
$$

Angles 0, 1, 2 form a spherical triangle so the sum of them should be 180 degrees plus the spherical excess, if the measurements were done perfectly. So we now define $f_3(v_0, v_1, \ldots, v_{26}) = v_0 + v_1 + v_2 - (180 + s_1)$ where s_1 is the spherical excess in this triangle HLB in Figure 2. Then for "perfect" values v_0, v_1, v_2 of these angles, we have $f_3(v_0, v_1, \ldots, v_{26}) = 0$ but if we enter the observed values we find for $f_3(w_0, w_1, \ldots, w_{26})$ a number close to zero but not equal to it. We will give the numerical details below. In this way we can define nine functions for the triangles:

$$
\begin{aligned}
f_3(v_0, v_1, \ldots, v_{26}) &= v_0 + v_1 + v_2 - (180 + s_1) \\
f_4(v_0, v_1, \ldots, v_{26}) &= v_3 + v_4 + v_5 - (180 + s_2) \\
f_5(v_0, v_1, \ldots, v_{26}) &= v_6 + v_7 + v_8 - (180 + s_3) \\
f_6(v_0, v_1, \ldots, v_{26}) &= v_9 + v_{10} + v_{11} - (180 + s_4) \\
f_7(v_0, v_1, \ldots, v_{26}) &= v_{12} + v_{13} + v_{14} - (180 + s_5) \\
f_8(v_0, v_1, \ldots, v_{26}) &= v_{15} + v_{16} + v_{17} - (180 + s_6) \\
f_9(v_0, v_1, \ldots, v_{26}) &= v_{18} + v_{19} + v_{20} - (180 + s_7) \\
f_{10}(v_0, v_1, \ldots, v_{26}) &= v_{21} + v_{22} + v_{23} - (180 + s_8) \\
f_{11}(v_0, v_1, \ldots, v_{26}) &= v_{24} + v_{25} + v_{26} - (180 + s_9)
\end{aligned}
$$

The numbers $s_1, \ldots s_9$ are the spherical excesses of the nine triangles, as explained in paragraph 3.2 below.

Two extra equations can be derived from the sine rules, the derivation of which will be shown later. This will lead to two more functions which must be zero at the perfect values v_i:

$$
\begin{aligned}
& f_{12}(v_0, v_1, \ldots, v_{26}) \\
& = \log(\sin(v_0 - \frac{1}{3}s_1)) - \log(\sin(v_2 - \frac{1}{3}s_1)) - \log(\sin(v_3 - \frac{1}{3}s_2)) + \\
& \log(\sin(v_4 - \frac{1}{3}s_2)) - \log(\sin(v_6 - \frac{1}{3}s_3)) + \log(\sin(v_7 - \frac{1}{3}s_3) \\
& - \log(\sin(v_{16} - \frac{1}{3}s_6)) + \log(\sin(v_{17} - \frac{1}{3}s_6)) - \log(\sin(v_{19} - \frac{1}{3}s_7)) + \log(\sin(v_{20} - \frac{1}{3}s_7)),
\end{aligned}
$$

$$
\begin{aligned}
& f_{13}(v_0, v_1, \ldots, v_{26}) \\
& = \log(\sin(v_6 - \frac{1}{3}s_3)) - \log\sin(v_8 - \frac{1}{3}s_3)) - \log(\sin(v_9 - \frac{1}{3}s_4)) + \\
& \log(\sin(v_{10} - \frac{1}{3}s_4)) - \log(\sin(v_{13} - \frac{1}{3}s_5)) + \log(\sin(v_{14} - \frac{1}{3}s_5) \\
& + \log(\sin(v_{18} - \frac{1}{3}s_7)) - \log(\sin(v_{20} - \frac{1}{3}s_7)) + \log(\sin(v_{21} - \frac{1}{3}s_8)) - \log(\sin(v_{23} - \frac{1}{3}s_8)) - \\
& \log(\sin(v_{24} - \frac{1}{3}s_9)) + \log(\sin(v_{25} - \frac{1}{3}s_9)).
\end{aligned}
$$

The goal now is to find values $v_0, \cdots v_{26}$ such that $f_1(v_0, \cdots v_{26}) = f_2(v_0, \cdots v_{26}) = \cdots = f_{13}(v_0, \cdots v_{26}) = 0$ and the $v_0, \cdots v_{26}$ are "as close as possible" to the observed values $w_0, \cdots w_{26}$. Gauss calls the transition from w_i to v_i the "best adjustment" and $e_i = w_i - v_i$ the "most plausible error." With "as close as possible Gauss means that the sum of the squares of the differences $e_i = (w_i - v_i)$ is minimal. The reason for this last requirements is related to probability theory and I will discuss Gausss argumentation in Chapter 4.3 below. When substituting values $w_0, \cdots w_{26}$ into the 13 functions, the requirement $f_1 = f_2 = \cdots f_{13} = 0$ is not met, but all function values differ somewhat from 0. The reason for the existence of these deviations is that in practice it is not possible to measure an angle with mathematical perfection.

3.1 The numerical equations for the most plausible errors

As mentioned before Gauss specifies three kinds of equations, two equations for the sum of the angles surrounding Leeuwarden and Drachten, nine equations derived from the triangles, and two equations based on the sine rule. We now discuss the numerical implementation.

For the first kind of equations this is as follows. Take all the observed values w_i (i=1, 5, 8, 15, 18) of the angles with angular point Leeuwarden into consideration and sum these. We get $359°\ 59'\ 57.803''$ which falls 2.197 short of $360°$, the sum of the "perfect angles" v_i. So we know that the errors[1] $e_i = w_i - v_i$ of all these angles sum to -2.197. This can be done in a similar way for the angles of Drachten and we get the following two equations:

$$
\begin{aligned}
e_1 + e_5 + e_8 + e_{15} + e_{18} &= -2.197'' \\
e_7 + e_{11} + e_{12} + e_{19} + e_{22} + e_{26} &= -0.436''
\end{aligned}
$$

[1]Gauss did not use the notation $_i$ but he indicated the error in angle i by the number i inside a small circle.

We now take into consideration the triangle formed by the cities Harlingen, Leeuwarden and Ballum. The sum of the spherical angles w_0, w_1, w_2 in the triangle is $50^\circ\ 58'\ 15.238'' + 82^\circ\ 47'\ 15.351'' + 46^\circ\ 14'\ 27.202'' = 179^\circ\ 59'\ 57.791''$. Now because the measured angles are spherical angles, we also need to take the spherical excess of $1.749''$ (for this specific triangle) into consideration, see below Ch. 3.2 for its computation.

If the angles had been measured perfectly they should sum to $v_0 + v_1 + v_2 = 180^\circ + 1.749$. For the observed values w_i we find that $w_0 + w_1 + w_2 = 180^\circ - 2.209''$. If we write as above $w_i = v_i + e_i$ we find the first equation which Gauss uses as $e_0 + e_1 + e_2 = -3.958$". If this is done for all nine triangles we find the following:

$$\begin{aligned}
e_0 + e_1 + e_2 &= -3.958'' \\
e_3 + e_4 + e_5 &= +0.722'' \\
e_6 + e_7 + e_8 &= -0.753'' \\
e_9 + e_{10} + e_{11} &= +2.355'' \\
e_{12} + e_{13} + e_{14} &= -1.201'' \\
e_{15} + e_{16} + e_{17} &= -0.461'' \\
e_{18} + e_{19} + e_{20} &= +2.596'' \\
e_{21} + e_{22} + e_{23} &= +0.043'' \\
e_{24} + e_{25} + e_{26} &= -0.616''
\end{aligned}$$

Two more equations remain: these will be discussed in the next section.

3.1.1 Deriving the two equations of third kind

We consider the first equation of the third kind, and we will begin by explaining the principle. Suppose first that the network of triangles in Figure 2 is in a plane. Now we apply the sine rule in the triangles surrounding L=Leeuwarden. Call the sides of these triangles a=L-Harlingen, b = L-Sneek, c=L-Drachten, d=L-Dokkum, e=L-Ballum. In the first triangle, between Harlingen, Leeuwarden and Ballum, the sine rule gives $\frac{e}{\sin(v_0)} = \frac{a}{\sin(v_2)}$. Doing this for the other triangles we also derive $\frac{a}{sin(v_4)} = \frac{b}{\sin(v_3)}, \frac{b}{\sin(v_7)} = \frac{c}{\sin(v_6)}, \frac{c}{\sin(v_{20})} = \frac{d}{\sin(v_{19})}, \frac{d}{\sin(v_{17})} = \frac{e}{\sin(v_{16})}$. Multiplying these five equations gives:

$$\frac{a}{\sin(v_4)}\frac{b}{\sin(v_7)}\frac{c}{\sin(v_{20})}\frac{d}{\sin(v_{17})}\frac{e}{\sin(v_0)} = \frac{b}{\sin(v_3)}\frac{c}{\sin(v_6)}\frac{d}{\sin(v_{19})}\frac{e}{\sin(v_{16})}\frac{a}{\sin(v_2)},$$

which is equivalent to:

$$\frac{\sin(v_0)\sin(v_4)\sin(v_7)\sin(v_{17})\sin(v_{20})}{\sin(v_2)\sin(v_3)\sin(v_6)\sin(v_{16})\sin(v_{19})} = 1.$$

Now taking the logarithm:

$$\log(\sin(v_0)) - \log(\sin(v_2)) - \log(\sin(v_3)) + \log(\sin(v_4)) - \log(\sin(v_6)) + \log(\sin(v_7)) - \log(\sin(v_{16})) + \log(\sin(v_{17})) - \log(\sin(v_{19})) + \log(\sin(v_{20})) = 0$$

We now adapt the argument because the network is not on a plane but on a sphere. Gauss used Legendre's theorem for spherical triangles: If we have a spherical triangle with angular points A, B and C and spherical angles α, β, γ, and spherical excess s, then the angles in the plane triangle

ABC are in very good approximation $\alpha - \frac{1}{3}s$, $\beta - \frac{1}{3}s$, $\gamma - \frac{1}{3}s$, see for a proof [Legendre 1798].[2] Then we can use the same argument as above because we do not have to assume (to apply the sine rule) that all the triangles are in the same plane.

We get, by plugging in the spherical excesses, the equation which Gauss used:

$$\log(\sin(v^{(0)} - 0.583")) - \log(\sin(v^{(2)} - 0.583")) - \log(\sin(v^{(3)} - 0.382"))+$$
$$\log(\sin(v^{(4)} - 0.382")) - \log(\sin(v^{(6)} - 0.414")) + \log(\sin(v^{(7)} - 0.414"))$$
$$- \log(\sin(v^{(16)} - 0.389")) + \log(\sin(v^{(17)} - 0.389")) - \log(\sin(v^{(19)} - 0.368")) + \log(\sin(v^{(20)} - 0.368")) = 0$$

This is the equation $f_{12}(v_0, v_1, \ldots, v_{26}) = 0$ as above. If we do the same with all triangles surrounding Drachten, we then get the equation $f_{13}(v_0, v_1, \ldots, v_{26}) = 0$.

3.1.2 Linearization of the logarithm functions

The last two equations $f_{12} = 0$ and $f_{13} = 0$ were linearized by Gauss in order to obtain equations for the unknown errors which can easily be solved. Gauss only writes the result without giving any computation, so we will explain his procedure here. We write $w_i = v_i + e_i$ as above with w_i the observed value, v_i a "correct" value, and e_i the corresponding error. In this scenario Gauss assumes that the margins of the (unknown) errors e_i are sufficiently small to only consider the linear terms. He uses the following general form: $g(a - e) \approx g(a) - eg'(a)$. So we get the following linearization: $\log(\sin(v_i - \frac{1}{3}s)) = \log(\sin(w_i - \frac{1}{3}s) - e_i) \approx \log(\sin(a)) - e\frac{\cos(a)}{\sin(a)} = \log(\sin(a)) - e\cot(a)$ with $a = w_i - \frac{1}{3}s$ and $e = e_i$. Here log is the natural logarithm and a and e are in radians to make sure that the derivative of $\sin(x)$ is $\cos(x)$. However, Gauss works with logarithms with base 10 (not e) and with errors in arc seconds, not radians, and we note that ${}^{10}\log x = {}^{e}\log x / {}^{e}\log 10$ and 1 radian $= 180/\pi$ degrees $= 3{\cdot}60^3/\pi$ arc seconds. So if we have e in arcseconds, we have to divide by $3{\cdot}60^3/\pi$ to get e in radians. Therefore the linearisation becomes ${}^{10}\log(\sin(v_i - \frac{1}{3}s)) = {}^{10}\log(\sin(w_i - \frac{1}{3}s) - e_i)$ $\approx {}^{10}\log(\sin((w_i - \frac{1}{3}s)) - e\cot((w_i - \frac{1}{3}s)) \cdot \frac{1}{{}^{e}\log 10} \cdot (\pi/3 \cdot 60^3)$ if e is measured in arc seconds.

Here is the computation for the first term ${}^{10}\log(\sin(v^{(0)} - 0.583))$ of the equation $f_{12}(v_0, v_1, \ldots, v_{26}) = 0$. I print the formula in such a way that a modern student can easily program it in Mathematica (or plug it in in wolframalpha.com) and thus check the numbers in Gauss's equations.

$$\begin{aligned}
&{}^{10}\log(\sin((50 + 58/60 + 15.238/3600 - 0.583/3600)(\pi/180) - e_0)) \\
&\approx \log(\sin((50 + 58/60 + 15.238/3600 - 0.583/3600) * (\pi/180)))(1/\log 10) \\
&- (1/\log(10))\cot((50 + 58/60 + 15.238/3600 - 0.583/3600) * (\pi/180))(\pi/(3 * 60^3))e_0. \\
&= -0.109677114 - 1.70679696 \times 10^{-6} e_0.
\end{aligned}$$

Doing this for all the terms in the equation the constants add up to:

$$\begin{aligned}
&- 0.109677114 + 0.14131122 + 0.10891981 - 0.024830652 + 0.11888322 - 0.16716926 \\
&+ 0.0120431368 - 0.15798760 + 0.13664604 - 0.05817592588 \\
&= -0.00003712508.
\end{aligned}$$

Adding the linearizations of all the logarithms, and multiplying by -10^7, we obtain an equation, which we can write as follows, just like Gauss did:

$$\begin{aligned}
&17.068(e_0) - 20.174(e_2) - 16.993(e_3) + 7.328(e_4) - 17.976(e_6) + 22.672(e_7) - 5.028(e_{16}) \\
&+ 21.780(e_{17}) - 19.710(e_{19}) + 11.671(e_{20}) = -371.
\end{aligned}$$

[2] If we assume that the earth is a perfect sphere, the adaptation does not seem necessary because the sine rule changes from $\frac{e}{\sin(v_0)} = \frac{a}{sin(v_2)}$ to $\frac{\sin e}{\sin(v_0)} = \frac{\sin a}{\sin(v_2)}$ and so on, where the distances e, a have to be computed in radians rather than kilometers. This cannot have escaped Gauss, but he knew that Legendres theorem (possibly generalized) is also valid on more general curved surfaces. See [Gauss 1995, p. 147] "each of which must be reduced by one-third of the excess due to sphericity ore spheroidicity if they lie in a curved surface," and [Gauss 1828a, 144-149] which appeared in the same journal issue as [Gauss 1828].

If we linearize the terms in f_{13} in the same way, we obtain

$$17.976(e_6) - 0.880(e_8) - 20.617(e_9) + 8.564(e_{10}) - 19.082(e_{13}) + 4.375(e_{14}) + 6.798(e_{18}) - 11.671(e_{20}) + 13.657(e_{21}) - 25.620(e_{23}) - 2.995(e_{24}) + 33.854(e_{25}) = +370.$$

3.2 Spherical excess

The following is valid for spherical triangles, that is to say: triangles on a sphere whose sides are arcs of great circles. A great circle is a circle whose center is the same as the center of the sphere. The sum of the angles of a spherical triangle is (in radians) 2π plus the spherical excess, where the spherical excess is the surface area of the sphere divided by the square of the radius of the sphere. For a proof see, for example, [Molenbroek 1946, pp. 267-268]. The sphere in this case is the earth, with a radius of approximately 6371 kilometer; this value was known to Gauss. The distance between the cities was also known in good approximation.

An example of the spherical excess E of the triangle formed by Leeuwarden, Harlingen and Ballum will be calculated by $E = \frac{A}{R^2}$, where A= the area of the triangle and R the radius of the sphere. The area A can be approximated by Heron's formula $A = \sqrt{s(s-a)(s-b)(s-c)}$ where a, b, c are the sides of the triangle and $s = 1/2(a+b+c)$. For the area computation we assume the (small) spherical triangle to be plane. In the triangle given here as an example we can use the approximations, taken from a modern atlas, $a = 25$ km, $b = 35$ km and $c = 28$ km. So now $s = 44$, and thus $A = \sqrt{44(44-25)(44-35)(44-28)} \approx 347$ km^2. So $E \approx \frac{347}{6371^2}$ in radians and to express this in arc-seconds: $E \approx \frac{347}{6371^2} \times (180/\pi) \times 3600 \approx 1.763''$. Gauss uses 1.749. Our computation is approximate; a precise reconstruction of the computation which Gauss (or, more likely, his predecessor Kraijenhoff) used is outside the scope of this thesis.

Perhaps the reader will notice a circular reasoning here. The distance between the cities (taken from an old map) is used to find the spherical excess, which is then used to compute a new map, which can then be used to determine the distance between the cities again. However, the adjustments of the distances between the cities (between old and new map) was small, usually less than one hundred meters. Therefore the influence on the adjustment on the spherical excess can be ignored.

Chapter 4

Gauss's solution to the problem

4.1 Solving the system of equations

We will first give a practical description of the steps Gauss takes to arrive at his solution (see [Gauss 1995, pp. 149-158] combined with [Gauss 1995, pp. 107- 121]). The intuition behind this procedure will be commented on in Chapter 4.3. For simplicity we will explain the general procedure in a way inspired by our example, for $m = 13$ functions and $n + 1 = 27$ measured quantities. We start with our thirteen functions $f_1, f_2, \cdots f_{13}$, 27 measured angles w_i, $0 \leq i \leq 26$, and corresponding "most plausible" values v_i $0 \leq i \leq 26$ (to be determined later in this Chapter), as above. These "most plausible values" satisfy all equations $f_j = 0$, but the measured values do not.

The functions are all linear expressions in the variables plus a constant term, so we can rewrite the equations $f_1 = 0, \ldots f_{13} = 0$ which the "most plausible values v_i satisfy, as :

$$\begin{aligned} a_{1,0}v_0 + a_{1,1}v_1 + a_{1,2}v_2 + \cdots + a_{1,n}v_n &= s_1 \\ &\vdots \\ a_{m,0}v_0 + a_{m,1}v_1 + a_{m,2}v_2 + \cdots + a_{m,n}v_n &= s_m \end{aligned}$$

In these equations all coefficients $a_{i,j}$ and all constants s_i are known. The notation is modern, for Gausss notation see the end of this section.

Instead of the "most probable values" we can also substitute the measured value's w_i into the equation. This yields:

$$\begin{aligned} a_{1,0}w_0 + a_{1,1}w_1 + a_{1,2}w_2 + \cdots + a_{1,n}w_n &= c_1 + s_1 \\ &\vdots \\ a_{m,0}w_0 + a_{m,1}w_1 + a_{m,2}w_2 + \cdots + a_{m,n}w_n &= c_m + s_m \end{aligned}$$

Here the numbers $c_1, \ldots c_m$ are non-zero and they can be computed by plugging the measured values w_i into the equations.

The measured values w_i are not precisely the "most plausible" values v_i, so we set $w_i = v_i + e_i$ as above and subtract the first set of equations from the second. Thus we obtain:

$$\begin{aligned} &a_{1,0}e_0 + a_{1,1}e_1 + a_{1,2}e_2 + \cdots + a_{1,n}e_n = c_1 \\ &\vdots \\ &a_{m,0}e_0 + a_{m,1}e_1 + a_{m,2}e_2 + \cdots + a_{m,n}e_n = c_m \end{aligned}$$

In our situation this gives the equations we mentioned above: first the two equations for all the angles measured in Leeuwarden and in Drachten:

$$e_1 + e_5 + e_8 + e_{15} + e_{18} = -2.197'' \tag{4.1}$$
$$e_7 + e_{11} + e_{12} + e_{19} + e_{22} + e_{26} = -0.436'' \tag{4.2}$$

then the nine equation for the sums of the angles in the triangles:

$$e_0 + e_1 + e_2 = -3.958'' \tag{4.3}$$
$$e_3 + e_4 + e_5 = +0.722'' \tag{4.4}$$
$$e_6 + e_7 + e_8 = -0.753'' \tag{4.5}$$
$$e_9 + e_{10} + e_{11} = +2.355'' \tag{4.6}$$
$$e_{12} + e_{13} + e_{14} = -1.201'' \tag{4.7}$$
$$e_{15} + e_{16} + e_{17} = -0.461'' \tag{4.8}$$
$$e_{18} + e_{19} + e_{20} = +2.596'' \tag{4.9}$$
$$e_{21} + e_{22} + e_{23} = +0.043'' \tag{4.10}$$
$$e_{24} + e_{25} + e_{26} = -0.616'' \tag{4.11}$$

And finally the two equations of the third kind described earlier (for e_i in arc seconds):

$$17.068(e_0) - 20.174(e_2) - 16.993(e_3) + 7.328(e_4) - 17.976(e_6) + 22.672(e_7) - 5.028(e_{16}) + 21.780(e_{17}) - 19.710(e_{19}) + 11.671(e_{20}) = -371$$

$$17.976(e_6) - 0.880(e_8) - 20.617(e_9) + 8.564(e_{10}) - 19.082(e_{13}) + 4.375(e_{14}) + 6.798(e_{18}) - 11.671(e_{20}) + 13.657(e_{21}) - 25.620(e_{23}) - 2.995(e_{24}) + 33.854(e_{25}) = +370$$

We will now explain the solution in modern notation. We write $\vec{c} = (c_1, c_2, \cdots, c_{13})^t$, a known vector, and $\vec{e} = (e_1, e_2, \cdots, e_{13})^t$, the vector of the "most plausible" errors, which we want to find, and we use the matrix $Q = (a_{i,j})$. The letter ^{t} or ^{T} will indicate the transpose of a vector or a matrix. In the example we have

$$Q= \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1\ldots \\ \vdots & \vdots & \vdots & \ddots & \vdots & & & & & \\ 17.068 & 0 & -20.172 & -16.993 & 7.328 & 0 & -17.976 & 22.672 & 0 & \ldots \\ 0 & 0 & 0 & 0 & 0 & 0 & 17.976 & 0 & -0.880 & \ldots \end{bmatrix}$$

Now in modern terms Gauss wants to find a vector $\vec{e}$ such that $Q\vec{e} = \vec{c}$, where the c_i are the constants found earlier by evaluating the equations at the measured values of the angles. Furthermore he wants the solution $\vec{e}$ such that the sum of squares of the components (that is, the square of the length of the vector $\vec{e}$) is minimized. We show his general solution to this problem in our concrete example. In brief, Gauss first finds a vector $\vec{x}$ such that $QQ^T\vec{x} = \vec{c}$ and he then finds $\vec{e} = Q^T\vec{x}$. Here are the details.

First Gauss makes a computation which is equivalent to finding the square, invertible and symmetric $m \times m=$ matrix $QQ^T =$

$$\begin{bmatrix} \sum_{i=0}^{26} a_{1,i}a_{1,i} & \sum_{i=0}^{26} a_{1,i}a_{2,i} & \dots & \sum_{i=0}^{26} a_{1,i}a_{13,i} \\ \sum_{i=0}^{26} a_{2,i}a_{1,i} & \sum_{i=0}^{26} a_{2,i}a_{2,i} & \dots & \sum_{i=0}^{26} a_{2,i}a_{13,i} \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{i=0}^{26} a_{13,i}a_{1,i} & \sum_{i=0}^{26} a_{13,i}a_{2,i} & \dots & \sum_{i=0}^{26} a_{13,i}a_{13,i} \end{bmatrix}$$

In order to keep in touch with Gauss's argument we will explain here some of the notation which he used and which was frequently used by others after him (even by B.L. van der Waerden in *Mathematische Statistik* [Van der Waerden 1965, p. 126] who calls this notation "old-fashioned" but "very easy"). Here is Gauss notation $a, a', a'', b, b', b'', c, c', c''$ and its translation in our notation:

$a = a_{1,0}, a' = a_{1,1}, a'' = a_{1,2}, \dots, b = a_{2,0}, b' = a_{2,1}, b'' = a_{2,2} \dots c = a_{3,0}, c' = a_{3,1}, c'' = a_{3,2}$ ... etc. Then Gauss defines $[aa] = aa+a'a'+a''a''+\ldots = \sum_{i=0}^{26} a_{1,i}a_{1,i}$, $[bb] = bb+b'b'+b''b''+\ldots = \sum_{i=0}^{26} a_{2,i}a_{2,i}$, $[ab] = ab+a'b'+a''b''+\ldots = \sum_{i=0}^{26} a_{1,i}a_{2,i}$, $[ba] = ba+b'a'+b''a''+\ldots = \sum_{i=0}^{26} a_{2,i}a_{1,i}$, and so on. Clearly $[ab] = [ba]$.

In this notation of Gauss (who did not use the letter j in his mathematical notation), our matrix QQ^T would look as follows (m=13):

$$\begin{bmatrix} [aa] & [ab] & \dots & [an] \\ [ba] & [bb] & \dots & [bn] \\ \vdots & \vdots & \vdots & \vdots \\ [na] & [nb] & \dots & [nn] \end{bmatrix}$$

Gauss does not use the concept of matrix but he sets up the system of 13 equations which we can express as $QQ^T\vec{x} = \vec{c}$.[1]

In the example, Gauss calls the coordinates of our vector $\vec{x}$ the "correlates A,B, ..., N. He introduces these numbers as the solutions of the equations which we can write as $vecc = QQ^T\vec{x}$. The 13 equations appear as follows (* indicates errors in the translation [Gauss 1995, p. 155]; the

[1]Gauss tries to not assume that the functions f_i are linear, and he works as long as possible with derivatives. Instead of $f_1, f_2, f_3 \dots$ he uses X, Y, Z. In our notation, the elements of the matrix Q are his derivatives:

$$\frac{dX}{dv^{(0)}} = a_{1,0}, \frac{dX}{dv^{(1)}} = a_{1,1}, \cdots$$
$$\frac{dY}{dv^{(0)}} = a_{2,0}, \frac{dY}{dv^{(1)}} = a_{2,1}, \cdots$$
$$\frac{dZ}{dv^{(0)}} = a_{3,0} \frac{dZ}{dv^{(1)}} = a_{3,1}, \cdots$$
$$\vdots$$

equations are printed correctly in the original text [Gauss 1826, p 90])

$$
\begin{aligned}
-2.197 &= 5A + C + D + E + H + I + 5.917N \\
-0.436 &= 6B + E + F + G + I + K + L + 2.962M \\
-3.958 &= A + 3C - 3.106M \\
+0.722 &= A + 3D - 9.665M \\
-0.753 &= A + B + 3E + 4.696M + 17.096N \\
+2.355 &= B + 3F - 12.053N \\
-1.201 &= B + 3G - 14.707N \\
-0.461 &= A + 3H + 16.752M \\
+2.596 &= A + B + 3I - 8.039M - 4.874N \\
+0.043 &= B + 3K - 11.963N \\
-0.616 &= B + 3L + 30.859N \\
-371 &= 2.962B - 3.106C - 9.665D + 4.696E + 16.752H - 8.039I + 2902.27M - 459.33N \\
+370 &= +5.917A + 17.096E - 12.053F - 14.707G - 4.874I - 11.963K + 30.859L - 459.33M + 3385.96N.
\end{aligned}
$$

Gauss then solves this system of 13 linear equations in 13 unknowns "by elimination" and finds the following:

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

In terms of the Gaussian "correlates, the equation $\vec{e} = Q^T\vec{x}$ boils down to:

$$
\begin{aligned}
e_i = {} & a_{1,i}A + a_{2,i}B + a_{3,i}C + a_{4,i}D + a_{5,i}E + a_{6,i}F + a_{7,i}G + a_{8,i}H \\
& + a_{9,i}I + a_{10,i}K + a_{11,i}L + a_{12,i}M + a_{13,i}N.
\end{aligned}
$$

Doing this for all e_i we get:

$e_0 = C + 17.068N$	$e_1 = A + C$	$e_2 = C - 20.174M$
$e_3 = D - 16.993M$	$e_4 = D + 7.328M$	$e_5 = A + D$
$e_6 = E - 17.976M + 17.976N$	$e_7 = B + E + 22.672M$	$e_8 = A + E - 0.880N$
$e_9 = F - 20.617N$	$e_{10} = F + 8.564N$	$e_{11} = B + F$
$e_{12} = B + G$	$e_{13} = G - 19.082N$	$e_{14} = G + 4.375N$
$e_{15} = A + H$	$e_{16} = H - 5.028M$	$e_{17} = H + 21.780M$
$e_{18} = A + I + 6.798N$	$e_{19} = B + I - 19.710M$	$e_{20} = I + 11.671M - 11.671N$
$e_{21} = K + 13.657N$	$e_{22} = B + K$	$e_{23} = K - 25.620N$
$e_{24} = L - 2.995N$	$e_{25} = L + 33.854N$	$e_{26} = B + L$

Gauss only writes out the first four of these equations.

Using this Gauss now determines the "most plausible errors $e_0 \cdots e_{26}$ in arc seconds:

$e_0 = -3.108$	$e_1 = -1.832$	$e_2 = 0.981$
$e_3 = +1.952$	$e_4 = -0.719$	$e_5 = -0.512$
$e_6 = +3.648$	$e_7 = -3.221$	$e_8 = -1.180$
$e_9 = -1.116$	$e_{10} = +2.376$	$e_{11} = +1.096$
$e_{12} = +0.016$	$e_{13} = -2.013$	$e_{14} = +0.795$
$e_{15} = +0.061$	$e_{16} = +1.211$	$e_{17} = -1.732$
$e_{18} + 1.265$	$e_{19} = +2.959$	$e_{20} = -1.628$
$e_{21} = +2.211$	$e_{22} = +0.322$	$e_{23} = -2.489$
$e_{24} = -1.709$	$e_{25} = +2.701$	$e_{26} = -1.606$

Gauss notes that the sum of squares of these errors is 97.8845 and he then states that the "mean error" is $\sqrt{\frac{97.8845}{13}} = 2.7''$. This number is connected to the reliability of the measurements.

4.2 Correctness of the method and its result

In his article Gauss also shows that his method to determine the "the most plausible errors", produces the solution for the system for which the squares of the errors is minimized. Because the number of variables (the errors, so 27) is larger then the number of equations (we have 13 equations in this instance), there is an infinite amount of possible solutions to the system. Clearly, the vector $\vec{e} = Q^T\vec{x}$ which Gauss has found is a solution to the system $Q\vec{e} = \vec{c}$ because $Q\vec{e} = QQ^T\vec{x} = \vec{c}$, and that is how x was computed. He proceeds as follows (we render his argument in our modernized notation but keep close to his line of reasoning).
Let $\vec{E} = (E_0, E_1, E_2, \ldots)$ be another solution to the system $Q\vec{E} = \vec{c}$. We write out the equations:

$$\begin{aligned} c_1 &= a_{1,0}E_0 + a_{1,1}E_1 + a_{1,2}E_2 + \cdots \\ c_2 &= a_{2,0}E_0 + a_{2,1}E_1 + a_{2,2}E_2 + \cdots \\ c_3 &= a_{3,0}E_0 + a_{3,1}E_1 + a_{3,2}E_2 + \cdots \\ &\qquad\qquad\qquad\qquad\qquad\qquad \vdots \end{aligned}$$

If we multiply the equations in turn by the 13 "correlates" $A, B, C, \cdots$, add the equations and change the order of the terms we get:

$$\begin{aligned} &a_{1,0}AE_0 + a_{2,0}BE_0 + a_{3,0}CE_0 + \cdots \\ &+ a_{1,1}AE_1 + a_{2,1}BE_1 + a_{3,1}CE_1 \cdots \\ &+ a_{1,2}AE_2 + a_{2,2}BE_2 + a_{3,2}CE_2 + \cdots \\ &\vdots \\ &= c_1A + c_2B + c_3C + \cdots \end{aligned}$$

Furthermore we also have:

$$e_0 = a_{1,0}A + a2, 0B + a_{3,0}C + \ldots$$

and indeed

$$e_i = a_{1,i}A + a2, iB + a_{3,i}C + \ldots$$

.

So now rewriting the former equation for $c_1A + c_2B + c_3C + \ldots$, we get:

$$e_0E_0 + e_1E_1 + e_2E_2 + \cdots = c_1A + c_2B + c_3C + \ldots$$

This same procedure can of course be done for $(E_0, E_1, \ldots) = (e_0, e_1, \ldots)$ (because this is also a solution of the system), here we get:

$$e_0e_0 + e_1e_1 + e_2e_2 + \ldots = c_1A + c_2B + c_3C + \ldots$$

From this Gauss concludes the following:

$$\begin{aligned} & e_0e_0 + e_1e_1 + e_2e_2 + \ldots + (E_0 - e_0)^2 + (E_1 - e_1)^2 + (E_2 - e_2)^2 + \ldots \\ & = E_0E_0 + E_1E_1 + E_2E_2 + \cdots + 2(e_0^2 + e_1^2 + e_2^2 + \cdots) - 2(e_0E_0 + e_1E_1 + e_2E_2 + \cdots \\ & = c_1A + c_2B + c_3C + \cdots = E_0^2 + E_1^2 + E_2^2 + \cdots \end{aligned}$$

And from this we see that for every other solution $E_0, E_1, E_2, \ldots$ the sum of the E_i^2 must be larger then $e_0^2 + e_1^2 + e_2^2 + \ldots$. Gauss concludes that the solution $e_0, e_1, e_2, \ldots$ which he has found are indeed the "most plausible errors" for our measured values w_i.

In modern terms, we can describe this procedure in an even simpler way as follows. Assume $\vec{E} = (E_0, E_1 \cdots E_{26})^t$ is an arbitrary solution of $Q\vec{E} = \vec{c}$. If we denote the standard inner product as $< x, y >$, we can now write: $< \vec{e}, \vec{E} >=< Q^T\vec{x}, E >=< \vec{x}, Q\vec{E} >=< \vec{x}, \vec{c} >$. If we set $\vec{E} = \vec{e}$ we get also $< \vec{e}, \vec{e} >=< \vec{x}, \vec{c} >$. From this it follows that $< \vec{E}, \vec{E} > - < \vec{e}, \vec{e} > - < \vec{E - e}, \vec{E - e} >=< \vec{E}, \vec{E} > - < \vec{e}, \vec{e} > - < \vec{E}, \vec{E} > +2 < \vec{E}, \vec{e} > - < \vec{e}, \vec{e} >= 2 < \vec{E}, \vec{e} > -2 < \vec{e}, \vec{e} >= 2 < \vec{x}, \vec{c} > -2 < \vec{x}, \vec{c} >= 0$. Now it is clear that $< \vec{E}, \vec{E} >=< \vec{e}, \vec{e} > + < \vec{E - e}, \vec{E - e} >$, but $< \vec{E}, \vec{E} >$ is the sum of the squares of the components. Thus we conclude that $\vec{E}$ is a solution with minimal sum of squares only if $\vec{E} - \vec{e} = 0$, which proves that Gauss has found the unique vector $\vec{e}$ such that the sum of the squares of the components is minimal.

4.3 The idea behind the solution method in modern terms

One of the difficulties in working on this thesis was to get a sense of how Gauss found his solution method, or even to get a feel for the reasoning behind it. Gauss' reasoning as described in the *Supplement* is not very mind-full of the understanding of the not so advanced reader. The literature [Gauss 1995] gives a few comments on the parts where the method is described but in my experience this did not give much clarifications, if at all.

In his method (explained in the case of the example), Gauss started with 27 measured angles w_i. Mathematically, the angles should satisfy 13 equations but because of the errors of measurement, the measured values w_i of the angles do not satisfy the equations. Gauss then wants to find the solution angles v_i to the equations such that such the sum of squares of the measured angles minus the solution angles is minimal i.e. $\min\{\sum_{i=1}^{26}(w_i - v_i)^2\} = \min\{\sum_{i=1}^{26} e_i^2\}$. He considers v_i the "most plausible" values of the angles; and $w_i - v_i = e_i$ the "most plausible" errors of measurement.

The idea behind Gauss' method of solution can be understood in terms of Lagrange multipliers (he doesn't mention that he uses this principle, but it is clear that he does). Gauss wants to minimize $\{\sum_{i=1}^{26} e_i^2\}$ subject to 13 conditions which one can express as $g_1(e_0, e_1, \ldots e_{26}) = c_1, \ldots g_{13}(e_0, e_1, \ldots e_{26}) = c_{13}$. According to the principle of Lagrange multipliers, the extremum is found in a point $(e_0, e_1, \ldots e_{26})$ such that $\nabla(e_0^2 + e_1^2 + \ldots + e_{26}^2) = (2e_0, 2e_1, \ldots 2e_{26}) = \sum_{k=0}^{26} \lambda_k \nabla g_k$. Here λ_k is (an as yet unknown) number, called Lagrange multiplier, and ∇g_k is the vector of partial derivatives of the function g_k, that is the vector of coefficients of the linear function g_k. Gauss uses $\frac{\lambda_1}{2} = A$, $\frac{\lambda_2}{2} = B, \cdots$, until $\frac{\lambda_{13}}{2} = N$. He calls the numbers $A, B, \ldots N$ the "correlates." By taking the first, second, etc. coordinate in the equation $(e_0, e_1, \ldots e_{26}) = \sum_{k=0}^{26} \frac{1}{2}\lambda_k \nabla g_k$, we can express the e_i in terms of the correlates. This produces $e_0 = C + 17.068N$ and $e_1 = A + C$, etc., as in the end of the previous section. If we call, as above, Q the matrix with in the k-th row the coefficients of g_k, and $\vec{x} = (A, B, \ldots, N)$ we can write $\vec{e} = (e_0, e_1, \ldots e_{26}) = \sum_{k=0}^{26} \frac{1}{2}\lambda_k \nabla g_k$ as $\vec{e} = Q^T\vec{x}$ as above.

Now to find the $A, B, C, etc.$ we know that $\vec{x}$ should comply to: $\vec{c} = Q\vec{e} = QQ^T\vec{x}$.

Since $\vec{c}$ is known and QQ^T can easily be computed, Gauss can find his correlates, that is $\vec{x}$ by solving the system of 13 equations in the 13 unknowns.

4.4 The normal distribution

This paragraph is dedicated to Gauss's motivation for using the least squares method. He starts the article [Gauss 1823] [Gauss 1995, p. 1-10] by discussing different kinds of errors one can encounter when doing measurements and makes a distinction between "random errors" and " constant errors", meaning systematic errors. He then argues why investigating these random errors gives an idea of the accuracy of the measurements. An argument supporting this claim might go as follows. When doing measurements there are bound the be some inaccuracies, intuitively it is clear that some inaccuracies have a higher probability of happening then others. For example when measuring the length of a piece of wood, lets say with a straightedge, it is obviously more likely that the measurement will have an inaccuracy of 1 centimeter then 1 kilometer. Naturally Gauss understood this very well, but to perform mathematics with this abstract notion in mind a more precise way of determining probabilities of certain inaccuracies to occur is needed. Gauss starts with the following assumptions: (1) small errors are more likely more likely to occur then large errors, (2) given a collection of measurements of the same quantity the most likely value to be observed is their average, that is their arithmetic mean. Another assumption is, (3) it is equally likely that the inaccuracy in the measurement, relative to the true value, is a positive amount or a negative amount with the same absolute value [Gauss 1995, p. 7]. But in [Gauss 1823] he gives no clear motivation for using (least) squares of the errors [Gauss 1995, p. 9-10].

In an earlier work on motion of the planets [Gauss 1809] Gauss had used these assumptions to determine a function which corresponds to the behavior $(1), (2), (3)$. In modern day mathematics this function would be known as a probability density function. In this derivation he needs to solve a differential equation. He proceeds as follows, see the German translation in [Gauss 1887, pp. 97-104] and the English translation in [Gauss 1857, pp. 253-261], see also [Sheynin 1972, pp.30-31.] We will amplify Gauss's account a bit, because it may be interesting for the modern readers to see Gauss's own argument for the normal 'Gaussian' distribution.
Let p be the true value of a measured quantity, and assume we are given n independent observations $M_1, \cdots M_n$. Suppose that $\phi(x)$ is the probability density function of the random error (which function Gauss is deriving here). Gauss assumes that this is a differentiable function, and he says that the probability that the error lies between a and b with $a < b$ is the integral $\int_a^b \phi(x)\,dx$. Now combining the assumption that small errors are more likely then large ones (1) and the assumption that negative errors are equally likely as positive ones with the same absolute value (3), it follows that $\phi(x)$ is maximized at $x = 0$ and $\phi(x) = \phi(-x)$.Now $M_i - p$ is the error of the i-th measurement. Because all the M_i are independent measurements he concludes that the joint density function Φ of all n combined errors is given by:

$$\Phi = \phi(M_1 - p)\phi(M_2 - p)\cdots\phi(M_n - p).$$

Gauss denotes the average of the measurements by

$$\hat{M} = \frac{M_1 + M_2 + \cdots + M_n}{n}$$

He now considers Φ as a function of p for given M_i, and he assumes that it is most likely that the true value p is equal to the average $\hat{M}$. This means that the choice of $p = \hat{M}$ maximizes the function Φ, so $\Phi'(p) = 0$ in $p = \hat{M}$.

Taking the logarithm of the equation $\Phi(p) = \phi(M_1 - p)\phi(M_2 - p)\cdots\phi(M_n - p)$ on both sides and differentiating we get the following:

$$\frac{\Phi'(p)}{\Phi(p)} = \frac{\phi'(M_1 - p)}{\phi(M_1 - p)} + \cdots \frac{\phi(M_n - p)}{\phi(M_n - p)} = 0$$

If we define[2] the function f as: $f(x) = \frac{\phi'(x)}{\phi(x)}$ the former equation can be expressed as:

$$f(M_1 - \hat{M}) + f(M_2 - \hat{M}) + \cdots f(M_n - \hat{M}) = 0.$$

We note $\phi(-x) = \phi(x)$, so $\phi'(x) = -\phi'(x)$ and $f(-x) = -f(x)$.

Now Gauss argues in a very clever way. Because we are not considering any specific case the measurements M_i can assume any arbitrary value, and in particular if M and N are arbitrary real numbers we can set $M_2 = M_3 = \cdots M_n = M - nN$ and $M_1 = M$. From this we obtain $\hat{M} = M - (n-1)N$ and substituting this into the former equation, we get:

$$f((n-1)N) + (n-1)f(-N) = 0 \text{ so } f((n-1)N) = (n-1)f(N).$$

Gauss concludes that $f(x) = kx$ for some real number k,[3] and from this it follows that $\frac{\phi'(x)}{\phi(x)} = kx$. Integrating with respect to x we have:

$$\log(\phi(x)) = \frac{k}{2}x^2 + c \text{ or } \phi(x) = Ce(\frac{kx^2}{2})$$

Now $\phi(x)$ reaches its maximum at $x = 0$ only when the constant $\frac{k}{2}$ is negative so we can set $k/2 = -h^2$. Now we have:

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

So finally, it follows that:

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

Now suppose that, as in the previous chapter, there are n quantities whose "true values" are $v_1, \ldots v_n$, and whose measured values are $w_1, w_2, \ldots w_n$, and also suppose that the measurements are independent and with the same accuracy, so the probability density for each error $(v_i - w_i)$ is the same function ϕ with the same h. Then the joint density function of these errors is

$$\begin{aligned}\Phi &= \phi(w_1 - v_1)\phi(w_2 - v_2)\cdots\phi(w_n - v_n) \\ &= (\frac{h}{\sqrt{\pi}})^n \exp(-\sum_{k=1}^{n} h^2(w_k - v_k)^2)\end{aligned}$$

For Gauss, the most likely errors $w_1 - v_1, \ldots w_n - v_n$ are the errors which maximize the joint density function Φ. Which is equivalent to minimizing the term $\sum_{k=0}^{n}(w_k - v_k)^2$, so, in words, minimizing the sum of the squares of the errors.

[2] Gauss denotes our function f by ϕ'; with the accent he means something like ϕ_1 rather than the derivative. Still, his notation will confuse many modern readers.

[3] We can see that this is true because if s and t are natural numbers, also by choosing $n = s+1$, $f((s/t)N) = f(s \cdot (N/t)) = sf(N/t)$ and then by choosing $t = s$, $f(N) = f(t/t)N = tf(N/t)$ so for all positive rational numbers a $f(aN) = af(N)$, and by a continuity argument we can see that this is also valid for all positive irrational numbers a because a is the limit of a series of rational numbers.

Bibliography

[Gauss 1809] = C.F. Gauss, *Theoria Motus Corporum Coelestium in sectionibus conicis solem ambientium.* Hamburg 1809. Available online on https://www.e-rara.ch/zut/content/titleinfo/142787

[Gauss 1823a] = Theoria combinationis observationum erroribus minimis obnoxiae, pars prior, in: *Commentationes societatis regiae scientarum Gottinginensis recentiores* 5, 1823. Commentationes classis mathematicae pp. 33-62. Available online
https://gdz.sub.uni-goettingen.de/id/PPN35283028X_0005_2NS
Also printed in [Gauss 1873, pp. 1-26].

[Gauss 1823b] = Theoria combinationis observationum erroribus minimis obnoxiae, pars posterior, in: *Commentationes societatis regiae scientarum Gottinginensis recentiores* 5, 1823. Commentationes classis mathematicae pp. 63-90.
Available online https://gdz.sub.uni-goettingen.de/id/PPN35283028X_0005_NS
Also printed in [Gauss 1873, pp. 27-54].

[Gauss 1828] = Supplementum theoriae combinationis observationum erroribus minimis obnoxiae, in: *Commentationes societatis regiae scientarum Gottinginensis recentiores* 6, 1828. Commentationes classis mathematicae pp. 57-98.
Available online https://gdz.sub.uni-goettingen.de/id/PPN35283028X_0006_2NS
Also printed in [Gauss 1873, pp. 55-93].

[Gauss 1828a] = Disquitiones generales circa superficies curvas, in: *Commentationes societatis regiae scientarum Gottinginensis recentiores* 6, 1828. Commentationes classis mathematicae pp. 99-149.

[Gauss 1857] = C.F. Gauss, *Theory of the motion of heavenly bodies moving about the sun in conic sections,* English translation by C.H. Davis, Boston 1857.
Available online https://archive.org/details/theoryofmotionof00gausuoft

[Gauss 1873] = Gauss, *Werke*, vol. 4. Gottingen, Konigliche Gesellschaft der Wissenschaften, 1873. Available online https://archive.org/details/werkecarlf04gausrich

[Gauss 1887] = *Abhandlungen zur Methode der kleinsten Quadrate von C.F. Gauss, in deutscher Sprache herausgegeben von A. Börsch und P. Simon.* Berlin 1887.
Available online https://archive.org/details/abhandlungenmet00gausrich

[Gauss 1995] = *C.F. Gauss, theory of the combination of observations least subject to errors, part one, part two, supplement.* Translated by G.W. Stewart. Philadelphia 1995.

[Haasbroek 1972] = N.D. Haasbroek, *Investigation of the accuracy of Krayenhoff's triangulation (1802-1811) in Belgium, The Netherlands and a part of North Western Germany*, Delft 1972. Available online at www.ncgeo.nl, Publicaties, Green series.

[Kaufmann-Buhler], W. Kaufmann-Buhler, *Gauss: A biographical study.* Berlin 1981.

[Jordan-Eggert 1948] = W. Jordan, *Handbuch der Vermessungskunde, Erster Band*, bearbeitet von O. Eggert, Stuttgart 1948.

[Krayenhoff 1815] = *Précis Historique des opérations géodésiques et astronomiques faites en Hollande*. La Haye 1815. Available online:
https://books.google.nl/books?id=ORAOAAAAQAAJ

[Legendre 1798] = A.-M. Legendre, *Méthode pour déterminer la longueur exacte du quart du méridien d'après les observations faites pour la mesure de l'arc compris entre Dunkerque et Barcelone*, in: J.B.J. Delambre, *Méthodes analytiques pour la détermination d'un arc du méridien*, Paris, an VII (1798). Available online: https://www.e-rara.ch/zut/content/pageview/580461

[Molenbroek 1946] = P. Molenbroek, *Leerboek der Stereometrie*, Groningen 1946.

[Sheynin 1972] = O. B. Sheynin, C.F. Gauss and the Theory of Errors, *Archive for History of Exact Sciences* 20 (1979), pp. 21-72.

[Van der Waerden 1965] = B.L. van der Waerden, *Mathematische Statistik*, zweite Auflage, Berlin etc. 1965.