

# The Method of Gauss in 1799

Aivars Celmiņš

*Abstract.* It has been suggested that Gauss used the method of least squares on a data set published in 1799. The data set and its adjustment are reexamined, and it is concluded that the result of Gauss cannot be obtained by the least-squares method nor by any other approach mentioned by Gauss.

*Key words and phrases:* Least squares, priority claim, Legendre and Gauss, nonlinear constraints, adjustment principles.

## 1. INTRODUCTION

The controversy about the discovery of the method of least squares is described in detail by Plackett (1972), Sprott (1978), Stigler (1977, 1981) and Stewart (in Gauss, 1995). It might be summarized as follows. Legendre (1752–1833) published in 1805 a memoir, *Nouvelles méthodes pour la détermination des comètes*, in which he introduced and named the method of least squares. Gauss (1777–1855) published in 1809 a book, *Theoria motus corporum coelestium in sectionibus conicis solem ambientium* (Gauss, 1809 or 1963), where he discussed the method of least squares and, mentioning Legendre's work, stated that he himself had used the method since 1795. Legendre was offended by Gauss's statement, and in a letter to Gauss with compliments about the new book, he indicated that claims of priority should not be made without proof by previous publications. Gauss did not have such a publication, but was convinced that the idea of least-squares adjustment is so simple that many people must have used the method even before Gauss. In the following years, Gauss tried to produce evidence for his claim but had only little success. His own computational notes were lost; his diary entry of 1798 where he indicates work on a probability theory different from Laplace's is ambiguous; and his colleagues apparently did not remember discussions with Gauss or did not want to be involved in the dispute. After repeated prodding, only the astronomer Olbers included in a paper in 1816 a footnote asserting that Gauss had shown him the method of least squares in 1802. (Bessel published a similar note in a report in 1832.) Eventually Gauss gave up the search but did not retract

his claim. In 1820 Legendre published a supplement to his 1805 memoir with an appendix where he publicly attacked Gauss's claims of priority. The controversy continued, and in 1831 Schumacher wrote to Gauss about a publication of 1799 that contains data and adjustment results by Gauss. Schumacher suggested repeating the calculations and thereby demonstrating that the method of least squares was indeed used by Gauss in 1799. Gauss's answer was that he was well aware of the data but would not permit a recalculation, and that he furthermore opposed any more public testimony on his behalf; his word should be enough, and a testimony would only suggest that he could not be trusted. Nevertheless, the priority continued to be on Gauss's mind. In 1840 he expressed disappointment for not having found evidence of applications of the least-squares method among the papers of the deceased astronomer Tobias Mayer. Earlier Gauss had named Mayer as someone who most certainly must have used the method even before Gauss.

Schumacher's suggestion to repeat Gauss's calculations was taken up by Stigler (1981). He obtained the data in question and tried least-squares adjustments on them. He could not reproduce Gauss's results and hypothesized that Gauss might have used a constraint that is more accurate than the linearized one-term expansion of the constraint equation which was used by Stigler. In the present paper, we review the adjustment and conclude that the results published by Gauss certainly are not obtained by a minimization of observational errors in a least-squares sense even when the exact constraints are used. This raises the intriguing question, what method or principle did Gauss use? In his note, Gauss called the calculation procedure "my method" and announced a publication of it. However, the method was never published. Therefore, by reconstructing Gauss's calculations, one

---

Aivars Celmiņš is a staff member, U.S. Army Research Laboratory, Aberdeen Proving Ground, Maryland 21005-5067 (e-mail: celms@arl.mil).

could obtain an indication of what numerical approach was considered by Gauss appropriate for the particular adjustment problem in 1799.

## 2. THE ADJUSTMENT PROBLEM

The data in question are from the measurement of a meridian arc of the Earth. In 1770 the French Academy of Sciences was directed to work out a system of units that would be suitable for the whole world. The Academy proposed to define a new length unit *metre* as one  $10^7$ th part of the quadrant of the meridian arc of the Earth. To determine the length of the arc, measurements were made along a meridian from Dunkirk to Barcelona. The measurements consisted of astronomical determinations of latitudes and land surveying between the latitude observations. The results of the measurements are listed in Table 1 (see also Plackett, 1972, and Stigler, 1981). Originally, the data were published in *Allgemeine Geographische Ephemeriden* (1799, 4, page XXXV). Gauss reported his results in the same publication, page 378, and added a comment in the Corrections to Volume 4 of the *Allgemeine Geographische Ephemeriden* (1800, page 193). The originally published data contained a printer's error, and Gauss asserted that he had used his method (which was not explained) on both sets, with and without the error. His values for the ellipticity  $f$  of the meridian ellipse and the length  $Q$  of the quadrant are as follows:

- data without error,  $f = 1/187$  and  $Q = 2,565,006$  modules;
- data with printer's error,  $f = 1/50$ .

(One module =  $1/1000$  league  $\approx 3.898$  metres.) Gauss also reported in his comment that the ellipticity found by French surveyors was  $f = 1/150$  and that the difference between his and the French result was "not important in this case." The method

TABLE 1  
Original data

No.	Location	$S_i$ , modules	$\delta_i$ , degrees	$\Phi_i$
1	Barcelona to Carcassone	52,749.48	1.852 66	42° 17' 20"
2	Carcassone to Evaux	84,424.55	2.963 36	44° 41' 48"
3	Evaux to Pantheon	76,145.74	2.668 68	47° 30' 46"
4	Pantheon to Dunkirk	62,472.59	2.189 10	49° 56' 30"
	Totals	275,792.36	9.673 80	

Note:  $S_i$  are the distances between the indicated locations,  $\delta_i$  are the corresponding differences in latitudes, and  $\Phi_i$  are the latitudes of the midpoints of the distances. The distance  $S_3$  between Evaux and Pantheon was due to a printer's error originally given as 76,545.74 modules.

used by the French is not known. Using the exact constraint equations and a simultaneous adjustment of all observations one obtains the following least-squares results:

- data without error,  $f = 1/152$  and  $Q = 2,564,897$  modules;
- data with printer's error,  $f = 1/79$  and  $Q = 2,568,230$  modules.

The presently agreed ellipticity of the International Ellipsoid of the Earth is  $1/297$ .

Assuming that the meridian is an ellipse and that the latitude is defined by the elevation angle of the normal to the ellipse, one has the following relation between an arc length  $S$  and the latitudes  $\Lambda_S$  and  $\Lambda_E$  of its end points (see the Appendix):

$$(2.1) \quad S = A \int_{\Lambda_S}^{\Lambda_E} (1 - B \sin^2 \phi)^{-3/2} d\phi,$$

where  $A$  and  $B$  are constants. These constants can be determined by a least-squares adjustment of the data listed in Table 1 with (2.1) as constraint. After determination of the values of  $A$  and  $B$ , the ellipticity and the length of the quadrant can be computed as follows. Let  $a$  and  $b$  be the semimajor and semiminor axis of the ellipse, respectively, and let  $e$  be its eccentricity. Then

$$(2.2) \quad A = \frac{b^2}{a} = a(1 - e^2)$$

and

$$(2.3) \quad B = \frac{a^2 - b^2}{a^2} = e^2.$$

The ellipticity is defined by

$$(2.4) \quad f = \frac{a - b}{a} = 1 - \sqrt{1 - e^2} = 1 - \sqrt{1 - B}.$$

The length  $Q$  of the quadrant is given by the integral

$$(2.5) \quad Q = A \int_0^{\pi/2} (1 - B \sin^2 \phi)^{-3/2} d\phi.$$

The integral in (2.1) cannot be evaluated in closed form. Therefore one might use an approximate expression for the numerical treatment of the adjustment problem. Because in our case  $B \ll 1$  and  $\Lambda_E - \Lambda_S < 3^\circ$  a good approximation of  $S$  is

$$(2.6) \quad S \approx (\Lambda_E - \Lambda_S) A \left[ 1 + \frac{3}{2} B \sin^2 \left( \frac{\Lambda_S + \Lambda_E}{2} \right) \right].$$

This one-term approximation of (2.1) is also suggested by the form in which the data were published in Table 1. The table entries are the arc lengths  $S$ ; the differences  $\delta = \Lambda_E - \Lambda_S$  between

the latitudes of the end points; and the midpoint latitudes  $\Phi = (\Lambda_S + \Lambda_E)/2$ . If one treats the midpoint latitudes  $\Phi$  as fixed parameters, then (2.6) is a linear constraint equation for the observations  $S$  and  $\delta$ . Most likely this linearized form of the constraint (2.1) was used by everyone working on the problem in the 1790s. The most general least-squares constraint equation considered by Gauss in *Theoria motus* has the form  $S = f(A, B, \dots)$ , where  $S$  is an adjustable observable and  $A, B, \dots$  are free parameters. Equations (2.1) and (2.6) both have this form if the arc length is the regressand. If  $\delta = \Lambda_E - \Lambda_S$  is treated as the regressand, then (2.6) can be brought into such a form by solving it for  $\delta$ .

The exact constraint (2.1) can be approximated also by more sophisticated formulas than (2.6). Therefore, a duplication of Gauss's calculations is complicated by the necessity to guess which constraint form he could have used. Stigler (1981) has found by numerical experimentation that Gauss did not use the linearized form (2.6) of the constraint with  $S$  as regressand in a least-squares adjustment and suggested that Gauss had a better approximation to the exact constraint. If this were true, then an adjustment based on the exact constraint (2.1) should be closer to Gauss's solution than to an adjustment with the approximate constraint (2.6). We shall test this property of the solution by computing several variants of adjustments based on exact constraints. We need several variants because even with a given constraint equation, one can adjust, for instance, only the surveyed arc lengths  $S$ , or only the observed latitudes  $\Lambda$  or both with appropriate weights.

Also missing are estimates of data accuracies that might have been used by Gauss for the computation of adjustment weights. (Gauss mentions weighted adjustment in *Theoria motus*.) In particular, one would normally assume that the standard deviations of the arc lengths  $S_i$  are proportional to  $\sqrt{S_i}$ , but we are not at all sure that Gauss made such an assumption. Moreover, if one simultaneously adjusts the arc length  $S_i$  as well as the latitudes  $\Lambda_j$ , then one needs prior estimates of the standard deviations of all data. Fortunately, assumptions about data accuracies are not essential for the present investigation because they do not greatly influence the values of the fitted constants  $A$  and  $B$ .

### 3. PROBLEM FORMULATIONS

We describe in this section three formulations of the adjustment problem that were used in our calculations. The corresponding numerical solutions

can be easily obtained with any software for weighted adjustment that allows general constraint equations. For the present paper we used the utility routines described in Celmiņš (1979). Those routines solve constrained least-squares problems that are defined as follows:

Minimize

$$(3.1a) \quad W = \sum_{i=1}^s c_i^T P_i^{-1} c_i$$

subject to

$$(3.1b) \quad F_i(X_i + c_i; T) = 0, \quad i = 1, \dots, s,$$

where  $X_i$  are observed vectors with  $\dim(X_i) = n_i$ ,  $c_i$  are the corresponding least-squares corrections,  $P_i$  are estimated variance-covariance matrices of the observations  $X_i$ ,  $T$  is a free model parameter vector with  $\dim(T) = p$  and the  $F_i$  are constraint functions with  $\dim(F_i) = r_i$ . The unknowns of the problem are the corrections  $c_i$  of the observations  $X_i$  and the parameter vector  $T$ . It is assumed that the constraint or model functions  $F_i$  are twice differentiable with respect to all their arguments and that

$$(3.2) \quad \sum r_i - \sum n_i < p < \sum r_i.$$

If the constraint functions  $F_i$  are scalar ( $r_i \equiv 1$ ) then the utility routine COLSAC (Celmiņš, 1979) can be used. If the constraints  $F_i = 0$  contain sets of simultaneous equations for the  $c_i$  ( $r_i > 1$ ), then the more complicated routine COLSMU must be used.

We have tried adjustments of the arc lengths as well as of the latitude observations. The adjustment of the latter can be somewhat simplified by expressing the constraints in terms of the endpoint observations  $\Lambda_i$  themselves rather than in terms of the differences  $\delta_i$  and midpoint latitudes  $\Phi_i$  that are given in Table 1 because the differences and midpoint values are interdependent. (They are constrained by the condition that adjacent arcs must have common endpoints after adjustment.) We therefore reconstructed the observed endpoint latitudes from the data in Table 1. The result is shown in Table 2, which also contains a priori estimates of the standard deviations of the observations. Such estimates are necessary for the joint adjustment of arc lengths and latitudes. We obtained the estimates with the help of preliminary adjustments as follows.

To obtain estimates for the standard deviations  $e_{S_i}$  of the arc length measurements, we assumed that they are proportional to  $\sqrt{S_i}$ , and that the latitude observations are free of errors. We solved the corresponding adjustment problem with the

TABLE 2  
Reconstructed latitude data

No.	Location	Latitude, degrees $\Lambda_i$	Arc, modules	
			$S_i$	$e_{S_i}$
1	Barcelona	41.362 42	52, 749.48	22.16
2	Carcassone	43.215 08		
3	Evaux	46.178 44	84,424.55	28.03
4	Pantheon	48.847 12	76,145.74	26.62
5	Dunkirk	51.036 22	62,472.59	24.11

Note: The estimated standard errors of the latitudes are  $e_\Lambda = 5.005 \cdot 10^{-4}$  degrees. The estimated standard error for the erroneous distance  $S_3 = 76,545.74$  is  $e_{S_3} = 26.69$  modules.

help of the utility routine COLSAC (see Case 1 for details) using the correct data set from Table 1 and the exact constraint (2.1). The solution provided the proportionality factor for the standard deviations of  $S_i$ :

$$(3.3) \quad e_{S_i} = 0.0965\sqrt{S_i} \quad \text{modules.}$$

The corresponding values of  $e_{S_i}$  are listed in Table 2. One notices that the estimated standard deviations of all four arc measurements are similar because the differences among the square roots of the arc lengths are small.

To obtain an estimate for the standard deviation  $e_\Lambda$  of the astronomic observations of the latitudes we assumed that all latitude observations are equally accurate and that the measurements  $S_i$  of the arc lengths are free of errors. We then solved the adjustment problem with the  $\Lambda_i$  from Table 2 as observations, the exact equation (2.1) as constraint and the correct distances  $S_i$  as regressor variables. The problem was numerically solved using the utility program COLSMU (see Case 2, described below). The program produced for the standard deviation of the latitude observations the estimate

$$(3.4) \quad e_\Lambda = 5.005 \cdot 10^{-4} \text{ degrees} = 1.802''.$$

We now describe the adjustment processes for which we distinguish three cases.

**CASE 1. Adjustment of arc lengths.** In this case the adjustable data are the surveyed arc lengths  $S_i$  whereas the latitude observations  $\Lambda_j$  are treated as fixed nonadjustable constants. In terms of the prob-

lem formulation (3.1) we have, therefore, the data (regressand variables)

$$(3.5) \quad X_i = S_i, \quad i = 1, 2, 3, 4,$$

with variance estimates from (3.3),

$$(3.6) \quad P_i = e_{S_i}^2, \quad i = 1, 2, 3, 4.$$

From the exact relation (2.1) we obtain the following constraint equations for  $i = 1, 2, 3, 4$ :

$$(3.7) \quad \begin{aligned} &F_i(S_i + c_{S_i}; A, B) \\ &= S_i + c_{S_i} \\ &- A \int_{\Lambda_i}^{\Lambda_{i+1}} (1 - B \sin^2 \phi)^{-3/2} d\phi = 0, \end{aligned}$$

where the  $\Lambda_j$  are fixed constants (regressor variables). Corresponding linearized constraints are, for  $i = 1, 2, 3, 4$ ,

$$(3.8) \quad \begin{aligned} &L_i(S_i + c_{S_i}; A, B) \\ &= S_i + c_{S_i} - A \delta_i \left( 1 + \frac{3}{2} B \sin^2 \Phi_i \right) \\ &= 0, \end{aligned}$$

where  $\delta_i = \Lambda_{i+1} - \Lambda_i$ , and  $\Phi_i = (\Lambda_{i+1} + \Lambda_i)/2$  are fixed constants (regressor variables). The parameters of the adjustment problem are the free constants  $A$  and  $B$ . Condition (3.2) is satisfied with  $\sum r_i = 4$ ,  $\sum n_i = 4$  and  $p = 2$ . Because the constraints are scalar, this problem can be solved using the utility program COLSAC with either the exact constraints (3.7) or the linearized constraints (3.8).

**CASE 2. Adjustment of latitudes.** In this case, we adjust the latitude observations  $\Lambda_j$  and treat the surveyed arc lengths  $S_i$  as fixed numbers. Because the  $\Lambda_j$  enter the model equation (2.1) as limits of the arc-length integrals, the same latitude observation generally appears in two constraint equations corresponding to adjacent arcs. Particularly, the adjustable latitude observations  $\Lambda_2$ ,  $\Lambda_3$  and  $\Lambda_4$  appear each in two of the four constraint equations, and the constraints must be treated as a single equation system  $F_1 = 0$  of four simultaneous equations.

To cast the adjustment problem into the form (3.1), we define the adjustable data (the regressand variables) as a single vector  $X_1$  of observations. That is, in (3.1a),  $s = 1$  and the data vector is

$$(3.9) \quad X_1^T = (\Lambda_1, \Lambda_2, \Lambda_3, \Lambda_4, \Lambda_5).$$

The data variance matrix  $P_1$  is diagonal ( $5 \times 5$ )-matrix with the diagonal elements  $e_\Lambda^2$ . The single

constraint function  $F_1(X_1 + c_1; A, B)$  has four components  $f_i$ . If the exact relation (2.1) is used, then the components  $f_i = 0$  of the constraint equation  $F_1 = 0$  are as follows:

$$(3.10) \quad \begin{aligned} f_i &= S_i - A \\ &\cdot \int_{\Lambda_i + c_{\Lambda, i}}^{\Lambda_{i+1} + c_{\Lambda, i+1}} (1 - B \sin^2 \phi)^{-3/2} d\phi \\ &= 0, \quad i = 1, 2, 3, 4. \end{aligned}$$

In linearized form, the constraint equation has the components

$$(3.11) \quad \begin{aligned} l_i &= S_i - (\Lambda_{i+1} + c_{\Lambda, i+1} - \Lambda_i - c_{\Lambda, i}) \\ &\cdot A \left( 1 + \frac{3}{2} B \sin^2 \Phi_i \right) = 0, \end{aligned}$$

for  $i = 1, 2, 3, 4$ . The arc lengths  $S_i$  and the midpoint latitudes  $\Phi_i$  (in the linearized constraints (3.11)) are assumed to be fixed nonadjustable constants (regressor variables). Condition (3.2) is satisfied with  $r_1 = 4$ ,  $n_1 = 5$  and  $p = 2$ . This type of problem (with constraints in form of simultaneous equations) can be solved using the utility program COLSMU.

**CASE 3. Adjustment of arc lengths and latitudes.** In this case, all observations, the surveyed arc lengths  $S_i$  as well as the latitude observations  $\Lambda_j$ , are adjusted simultaneously. The problem can be solved by treating the arc lengths  $S_i$  in the constraint equations (3.10) or (3.11) as adjustable observations and using the utility program COLSMU for constraints in the form of simultaneous equations. Then the corresponding vector of observations  $X_1$  would have nine components (five  $\Lambda_j$  and four  $S_i$ ). The variance matrix  $P_1$  of the single observation vector would be a diagonal  $(9 \times 9)$ -matrix, and the constraint  $F_1 = 0$  would again be a system of four simultaneous equations. However, the numerical treatment and the coding of the problem can be simplified by introducing nonessential parameters (Celmiņš, 1982) that render the problem separable and transform the constraints into a set of nine independent scalar equations. Let the added parameters be  $\Theta_1, \dots, \Theta_5$  so that the augmented parameter vector  $T$  has seven components:

$$(3.12) \quad T^T = (A, B, \Theta_1, \Theta_2, \Theta_3, \Theta_4, \Theta_5).$$

The data are nine scalars:

$$(3.13a) \quad \begin{aligned} X_i &= S_i, \quad i = 1, 2, 3, 4, \\ X_i &= \Lambda_{i-4}, \quad i = 5, 6, 7, 8, 9. \end{aligned}$$

The corresponding variances are

$$(3.13b) \quad \begin{aligned} P_i &= e_{S_i}^2, \quad i = 1, 2, 3, 4, \\ P_i &= e_{\Lambda}^2, \quad i = 5, 6, 7, 8, 9. \end{aligned}$$

The constraints are nine scalar equations. The first four equations are, if the exact relation (2.1) is used,

$$(3.14) \quad \begin{aligned} F_i &= S_i + c_{S_i} \\ &- A \int_{\Theta_i}^{\Theta_{i+1}} (1 - B \sin^2 \phi)^{-3/2} d\phi \\ &= 0, \quad i = 1, 2, 3, 4. \end{aligned}$$

The next five constraint equations are new and define the nonessential parameters:

$$(3.15) \quad \begin{aligned} F_i &= \Lambda_{i-4} + c_{\Lambda, i-4} - \Theta_{i-4} = 0, \\ &i = 5, 6, 7, 8, 9. \end{aligned}$$

The linearized form of the constraint equations (3.14) is

$$(3.16) \quad \begin{aligned} L_i &= S_i + c_{S_i} - (\Theta_{i+1} - \Theta_i) A \\ &\cdot \left( 1 + \frac{3}{2} B \sin^2 \Phi_i \right) = 0, \\ &i = 1, 2, 3, 4, \end{aligned}$$

where the midpoint latitudes  $\Phi_i$  are assumed to be fixed parameters. The remaining constraints (3.15) for  $i = 5, 6, 7, 8, 9$  are already linear and need not be simplified.

Because the constraints (3.14), (3.15) and (3.16) are scalar, the exact as well as the linearized adjustment problem can be solved with the help of the utility routine COLSAC. Condition (3.2) is satisfied with  $\Sigma r_i = 9$ ,  $\Sigma n_i = 9$  and  $p = 7$ .

#### 4. LEAST-SQUARES RESULTS

The results of adjustments using the correct data set are listed in Table 3 and shown in Figures 1 and 2. Table 3 lists six adjustment results (Cases 1, 2 and 3 with exact and linearized constraints, respectively) giving the values of the arc length  $Q$ ; the inverse ellipticity  $1/f$ ; their corresponding estimated standard deviations  $e_Q$  and  $e_{1/f}$ , respectively; and an estimate of the correlation coefficient between  $Q$  and  $1/f$ . For the adjustments involving only the arc length  $S_i$  (Case 1) we assumed that their standard deviations are proportional to  $\sqrt{S_i}$ . The adjustments of the  $\Lambda_i$  only (Case 2) were made assuming that the standard deviations of the data are all equal. The adjustment weights in Case 3 were estimated as explained Section 3.

Figure 1 shows the results of adjustments with the linearized constraints, that is, results which we expect to be close to Gauss's result. The figure displays the values of  $Q$  and  $1/f$  and error ellipses

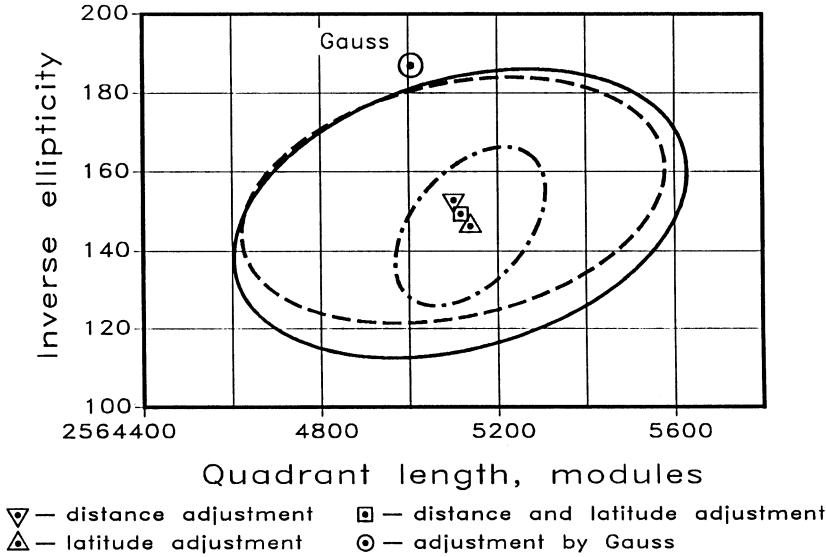


FIG. 1. Linearized adjustments of correct data.

TABLE 3  
Least-squares results for correct data set

Constraint	Observ.	Q	$e_Q$	1/f	$e_{1/f}$	Correlation
Exact	S	2,564,891	501	155.3	31.3	0.338477
	S and $\Lambda$	4,897	544	151.8	36.8	0.380900
	$\Lambda$	4,909	198	148.7	20.2	0.576044
Linear	S	5,100	479	152.7	31.3	0.174293
	S and $\Lambda$	5,116	513	149.3	36.8	0.192755
	$\Lambda$	5,138	170	146.2	20.2	0.299832

representing one standard deviation. The dashed curve corresponds to distance adjustments, the dot-dash curve corresponds to latitude adjustments, and the solid curve represents the standard deviation in the case where all data are adjusted simultaneously. Gauss's result is about one standard deviation apart from our result. The difference is not important statistically, but it indicates that the values reported by Gauss are not obtained with the linearized constraint. Based on this, Stigler (1981) hypothesized that Gauss might have used a better approximation than (2.6) or (3.8) to the exact constraint (2.1). (Stigler also considered the possibility

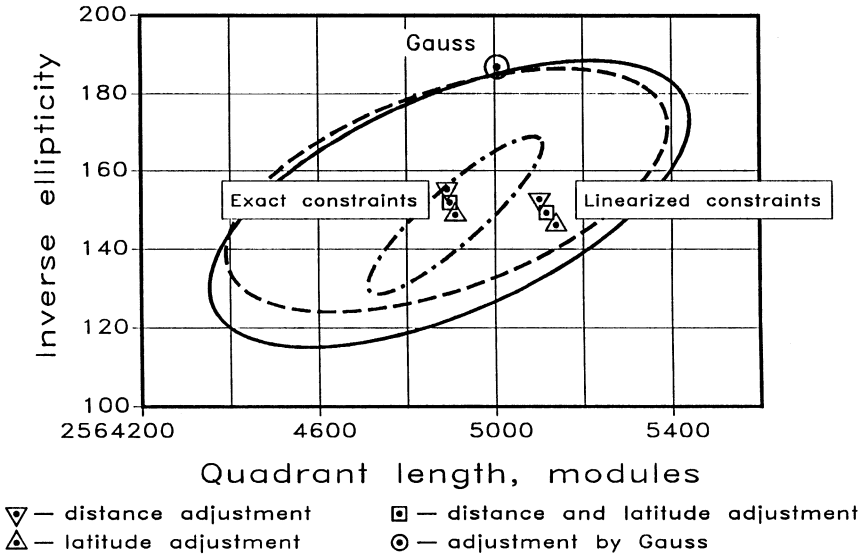


FIG. 2. Adjustments of correct data.

that the differences between his and Gauss's results could have been caused by rounding errors, but found that trigonometric and logarithmic tables available to Gauss were sufficiently accurate to prevent such large differences in the results.) To test Stigler's hypothesis, we also calculated the adjustments using exact constraints instead of the linearized ones. (The integrals in the exact constraints were numerically evaluated using a Romberg algorithm; see Acton, 1990.) The results are shown in Figure 2 in the same form as in Figure 1, and for comparison the results from Figure 1 are also displayed. The solution of Gauss is about equally far removed from the exact solution as from the solution using the linearized constraint but is not located between them. This clearly shows that the values of  $Q$  and  $1/f$  that were reported by Gauss are not obtained from a least-squares adjustment with improved constraints.

Next we consider the data set that contains the printer's error. The adjustment results are listed in Table 4 and displayed in Figure 3. One observes that differences among the six results are larger than in Figure 2 but the overall situation is about the same as shown in Figure 2. In this case, Gauss did not report a value for the quadrant length  $Q$  and we can only compare the line  $1/f = 50$  (Gauss's value) with our results. The line is well below any of our results.

Figure 4 is a combined display of all least-squares results. The error ellipses correspond to one standard deviation, as before, and are for the simultaneous adjustments of all data. The figure shows that the quadrant length and the inverse ellipticity

TABLE 4  
Least-squares results for data set with error

Constraint	Observ.	$Q$	$e_Q$	$1/f$	$e_{1/f}$	Correlation
Exact	$S$	2,568,000	491	87.8	9.9	0.281788
	$S$ and $\Lambda$	8,230	528	79.3	10.0	0.303993
	$\Lambda$	8,523	183	72.7	4.8	0.457031
Linear	$S$	8,682	472	85.0	9.9	-0.043977
	$S$ and $\Lambda$	9,067	507	76.6	9.9	-0.116073
	$\Lambda$	9,525	171	70.0	4.8	-0.320064

ties that were reported by Gauss are not obtained by least-squares adjustments. A printer's error in the results reported by Gauss is not likely, because Gauss's ellipticity values were published twice, in two different issues of the journal. Thus we are left with the question whether Gauss used a different adjustment principle or made an arithmetical error.

## 5. ADJUSTMENTS USING DIFFERENT PRINCIPLES

Candidates for adjustment principles that might have been used by Gauss are the minimization of the sum of  $n$ th powers of the absolute values of residuals, Boscovich's (1711–1787) method and a minimization of the maximum deviation. (Boscovich's method consisted of a minimization of the sum of absolute values of the residuals under the condition that the sum of the residuals should be equal to zero.) Gauss discusses all these methods in Article 186 of *Theoria motus*: the minimization of the sum of even powers of residuals, Boscovich's method and the minimization of very large powers of residuals that produces a minimax solution. He

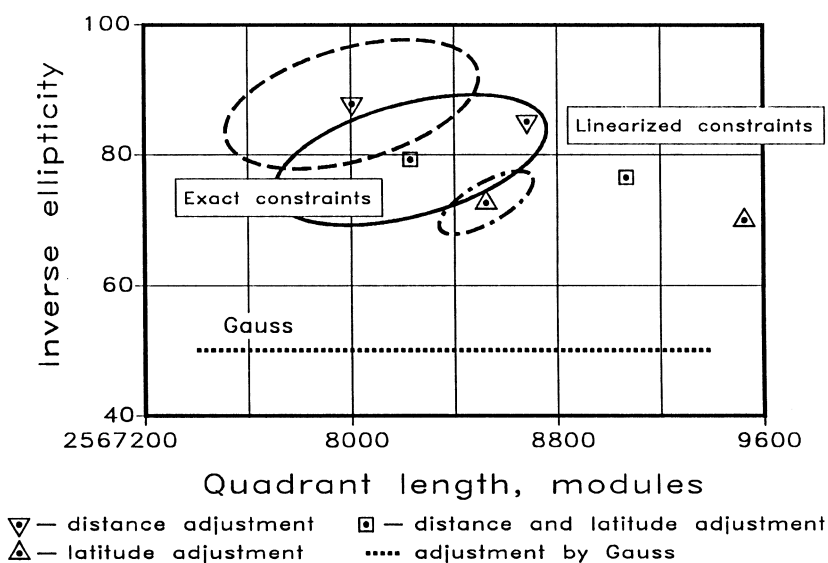


FIG. 3. Adjustments of data with error.

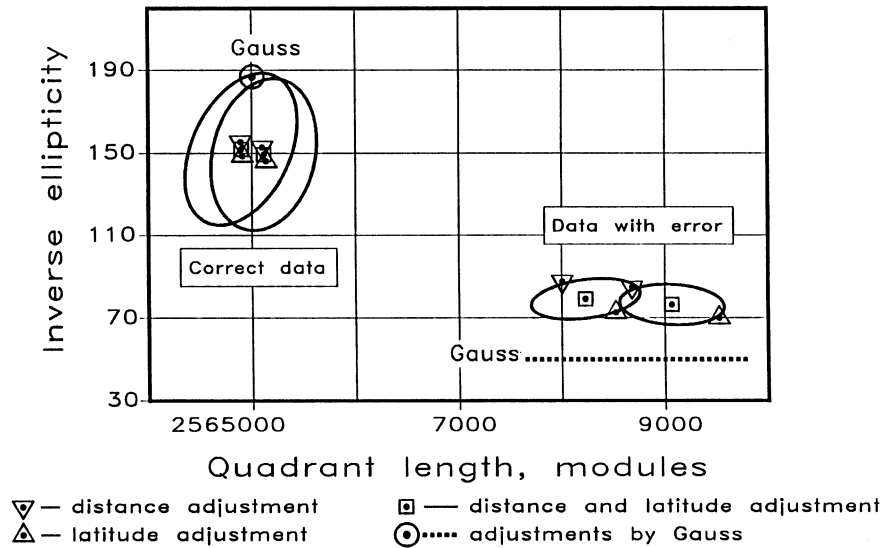


FIG. 4. All least-squares adjustments.

suggests using least squares on grounds of numerical expediency.

A further method that might have been used by Gauss is suggested by Sheynin (1993). In that method, a least-squares technique is not used to minimize a norm of observational errors but to minimize an objective function in the parameter space. The method has an ad hoc nature and it is not considered by Gauss in *Theoria motus*, but Sheynin asserts that the method has been widely used in land surveying during the past two centuries.

To get an idea about the range of results that can be obtained with these different adjustment methods, we carried out a number of adjustments of the arc length measurements  $S_i$ , equally weighted, and using the linearized constraint (3.8). The solutions were obtained by a numerical search for the minimum of the respective objective function. Some typical results are listed in Tables 5 and 6 and dis-

played in Figure 5. (A comparison of the corresponding entries in Tables 3 and 5, or Tables 4 and 5, respectively, shows that the weighting of the  $S_i$  inversely proportional to  $\sqrt{S_i}$  indeed makes little difference in the least-squares results.) The results for the correct data show that the quality and consistency of the data are so good that the adjustment method does not matter: all methods produce very similar results, and all are different from Gauss's result. On the other hand, adjustments of the data with the printer's error produce parameters that vary over a large range as the power  $n$  of the residuals varies between unity and infinity. The Boscovich method produces solutions that are in both cases close to the corresponding least-squares solutions. One also notices that the minimization of the sum of the absolute values of the residuals without additional conditions produces exactly the same result for both data sets. That is, the error in the value of  $S_3$  does not affect the result. Gauss

TABLE 5  
Various adjustments of correct distance set

Constraint	Objective	Q	1/f
Exact	Boscovich	2,564,907	158.2
	$\Sigma  c_S/e_S $	4,684	173.6
	$\Sigma  c_S/e_S ^{1.5}$	4,804	159.2
	$\Sigma  c_S/e_S ^2$	4,875	151.3
	$\Sigma  c_S/e_S ^4$	4,995	150.0
	$\Sigma  c_S/e_S ^\infty$	5,108	152.7
Linear	Boscovich	2,565,108	155.7
	$\Sigma  c_S/e_S $	4,853	171.1
	$\Sigma  c_S/e_S ^{1.5}$	5,004	156.7
	$\Sigma  c_S/e_S ^2$	5,096	148.7
	$\Sigma  c_S/e_S ^4$	5,218	147.3
	$\Sigma  c_S/e_S ^\infty$	5,323	150.0

TABLE 6  
Various adjustments of distance set with error

Constraint	Objective	Q	1/f
Exact	Boscovich	2,567,820	74.9
	$\Sigma  c_S/e_S $	4,684	173.6
	$\Sigma  c_S/e_S ^{1.5}$	6,858	89.7
	$\Sigma  c_S/e_S ^2$	8,137	79.3
	$\Sigma  c_S/e_S ^4$	2,570,072	93.0
	$\Sigma  c_S/e_S ^\infty$	1,205	105.8
Linear	Boscovich	2,568,760	72.4
	$\Sigma  c_S/e_S $	4,853	171.1
	$\Sigma  c_S/e_S ^{1.5}$	7,512	87.0
	$\Sigma  c_S/e_S ^2$	8,974	76.5
	$\Sigma  c_S/e_S ^4$	2,570,670	90.2
	$\Sigma  c_S/e_S ^\infty$	1,663	103.1





A relation different from (2.1) is obtained if the latitude is differently defined, for instance, as the elevation angle of the plumb line to a solid ellipsoid, or as the elevation angle of the ray from the center of the ellipsoid. Corresponding constraint equations are derived in the Appendix, which also contains a transformation formula for the values of the ellipticities in the three cases. Let  $f_C$ ,  $f_N$  and  $f_P$  be the ellipticities that correspond to latitude definitions in terms of the center ray, the normal to the ellipsoid, and the plumb line to the ellipsoid, respectively. Then one obtains, with a linearized unweighted least-squares adjustment of the arc lengths  $S_i$ , the results in Table 7 [see Tables 5 and 6 and (A.33)–(A.35)]. If one uses the center-ray definition of the latitudes, then the inverse ellipticity  $1/f_C$  is less than  $1/f_N$  for both data sets. If one uses the plumb-line definition, then  $1/f_P$  is larger than  $1/f_N$  for both data sets. Gauss's value is higher than  $1/f_N$  for the correct data and lower for the erroneous data set. Hence a change of the definition of latitudes that reduces the difference between Gauss's values and our  $1/f_N$  for one data set increases the difference for the other data set. Therefore, neither of the two alternative definitions considered can explain the discrepancies.

An error in the simplification of the exact constraint equation (2.1) cannot be excluded, except for the reason that the equation and corresponding analyses are so simple that it is difficult to make an error.

Arithmetical errors seem at first unlikely because both results by Gauss are erroneous, suggesting at least two errors. However, this need not be the case. The calculations by Gauss were done manually, writing down intermediate results, such as the values of trigonometric functions and logarithms. (The usefulness of writing down intermediate results is emphasized in several places in *Theoria motus*.) Then, as the problem was solved again with a corrected value of the distance  $S_3$ , only those parts had to be recalculated that directly involved the new datum. An error in a quantity that was not recalculated would influence both results. We tested this possibility by assuming that one of the four values of  $\sin \Phi_i$  was in error. By a proper choice of the value of  $\sin \Phi_3$  we obtained  $1/f = 187$  for the correct data set and a corresponding  $1/f = 64$  for

the data set with printer's error. This does not exactly duplicate Gauss's result, but it shows that a single error can indeed increase the ellipticity in one case and reduce it for the other data set.

We conclude from these considerations that the results published by Gauss likely contain arithmetical errors. Hence Gauss's publication neither supports nor falsifies his claim that he used the method of least squares before 1800. As Gauss suggested, we have to trust his word.

## APPENDIX: DERIVATION OF CONSTRAINT EQUATIONS

Astronomical latitude observations measure the elevation angle of the local plumb line with respect to the equatorial plane. Usually, it is assumed that the observed plumb line approximately coincides with the local normal to the ellipsoid of the Earth. Alternatively, one might assume that the observed plumb line coincides with the local plumb line of a homogeneous ellipsoid of the Earth, or, because the ellipticity of the Earth is very small, coincides with the ray to the center of the ellipsoid. Each of these assumptions yields a different relation between the measurement of the lengths of an arc and the corresponding difference of latitudes of the endpoints of the arc. This appendix provides a derivation of these relations.

Let  $a$  be the semimajor and  $b$  the semiminor axis of an ellipse. Then the ellipse can be represented by the following set of equations:

$$(A.1) \quad \begin{aligned} x &= a \cos \phi, \\ y &= b \sin \phi, \end{aligned}$$

where  $\phi$  is the elevation angle of a ray from the center of the ellipse. The arc length element of the ellipse is

$$(A.2) \quad \begin{aligned} ds &= \left( \left( \frac{\partial x}{\partial \phi} \right)^2 + \left( \frac{\partial y}{\partial \phi} \right)^2 \right)^{1/2} d\phi \\ &= (a^2 \sin^2 \phi + b^2 \cos^2 \phi)^{1/2} d\phi. \end{aligned}$$

Let  $e$  be the eccentricity of the ellipse, defined by

$$(A.3) \quad e^2 = \frac{a^2 - b^2}{a^2}.$$

Then (A.2) can be expressed in the form

$$(A.4) \quad ds = a\sqrt{1 - e^2} \left( 1 + \frac{e^2}{1 - e^2} \sin^2 \phi \right)^{1/2} d\phi.$$

Now let us assume that the latitudes are defined by the angle  $\phi$ ; that is, the plumb line coincides with the ray to the center of the ellipsoid, and that arc

TABLE 7

Data set	$1/f_C$	$1/f_N$	$1/f_P$	Gauss
Correct distances	49.2	148.7	168.6	187
Distances with error	25.2	76.5	86.8	50

lengths  $\Delta S$  and corresponding latitude differences  $\Delta \phi$  have been observed. Because for the ellipsoid of the Earth  $e^2 \ll 1$ , (A.4) can be expressed in the following linearized form:

$$(A.5) \quad \Delta S = \Delta \phi a \sqrt{1 - e^2} \left(1 + \frac{1}{2} e^2 \sin^2 \Phi\right),$$

where  $\Phi$  is the midpoint latitude of the observed arc  $\Delta \phi$ .

We now differently define the latitude  $\lambda$  as the elevation angle of the normal to the ellipse. (This is the usual definition.) The unit normal vector to the ellipse is

$$(A.6) \quad n = \frac{1}{(a^2 \sin^2 \phi + b^2 \cos^2 \phi)^{1/2}} \begin{pmatrix} b \cos \phi \\ a \sin \phi \end{pmatrix}.$$

Hence

$$(A.7) \quad \tan \lambda = \frac{a}{b} \tan \phi,$$

$$(A.8) \quad \sin^2 \phi = \frac{(1 - e^2) \sin^2 \lambda}{1 - e^2 \sin^2 \lambda}$$

and

$$(A.9) \quad d\phi = \frac{b \cos^2 \phi}{a \cos^2 \lambda} d\lambda = \frac{\sqrt{1 - e^2}}{1 - e^2 \sin^2 \lambda} d\lambda.$$

Substituting expressions (A.8) and (A.9) into (A.4), one obtains

$$(A.10) \quad ds = a(1 - e^2) \frac{1}{(1 - e^2 \sin^2 \lambda)^{3/2}} d\lambda.$$

An integration of (A.10) yields (2.1). A linearization for small  $e^2$  produces

$$(A.11) \quad \Delta S = \Delta \lambda a(1 - e^2) \left(1 + \frac{3}{2} e^2 \sin^2 \Lambda\right),$$

where  $\Lambda$  is the midpoint latitude of the observed arc  $\Delta \lambda$  [cf. (2.6)].

Next, we define the latitude as the angle of elevation of the plumb line to a homogeneous rotational ellipsoid with density  $\rho$ . The gravity potential of such an ellipsoid is, assuming that the semiminor axis  $b$  is the rotation axis (see Kellogg, 1953, page 194),

$$(A.12) \quad U = D - \frac{1}{\alpha} x^2 - \frac{1}{\beta} y^2,$$

where

$$(A.13) \quad D = \pi \rho a^2 b \int_0^\infty \frac{dt}{(a^2 + t)(b^2 + t)^{1/2}},$$

$$(A.14) \quad \frac{1}{\alpha} = \pi \rho a^2 b \int_0^\infty \frac{dt}{(a^2 + t)^2 (b^2 + t)^{1/2}}$$

and

$$(A.15) \quad \frac{1}{\beta} = \pi \rho a^2 b \int_0^\infty \frac{dt}{(a^2 + t)(b^2 + t)^{3/2}}.$$

The components of the gravitational force are

$$(A.16) \quad \begin{aligned} f_x &= \frac{\partial U}{\partial x} = -\frac{2x}{\alpha}, \\ f_y &= \frac{\partial U}{\partial y} = -\frac{2y}{\beta}. \end{aligned}$$

Hence the elevation angle  $\psi$  of the plumb line and the angle  $\phi$  are related by

$$(A.17) \quad \tan \psi = \frac{f_y}{f_x} = \frac{\alpha}{\beta} \tan \phi.$$

We define  $\varepsilon^2$  by

$$(A.18) \quad \varepsilon^2 = \frac{\alpha^2 - \beta^2}{\alpha^2}.$$

Then

$$(A.19) \quad \sin^2 \phi = \frac{(1 - \varepsilon^2) \sin^2 \psi}{1 - \varepsilon^2 \sin^2 \psi}$$

and

$$(A.20) \quad d\phi = \frac{\sqrt{1 - \varepsilon^2}}{1 - \varepsilon^2 \sin^2 \psi} d\psi.$$

Substituting expressions (A.19) and (A.20) into (A.4) one obtains

$$(A.21) \quad ds = a((1 - e^2)(1 - \varepsilon^2))^{1/2} \cdot \frac{1 + (e^2 - \varepsilon^2)/(1 - e^2) \sin^2 \psi}{(1 - \varepsilon^2 \sin^2 \psi)^{3/2}} d\psi.$$

To compare this expression with the previous ones, we want to express  $\varepsilon$  in terms of  $e$ . To that end we first derive formulas for  $\alpha$  and  $\beta$  in terms of  $e^2$ , assuming as before that  $e^2 \ll 1$ . From (A.14) we obtain, using definition (A.3),

$$(A.22) \quad \begin{aligned} \frac{1}{\alpha} &= \pi \rho a^2 b \cdot \int_0^\infty \frac{dt}{(a^2 + t)^{5/2} [1 - e^2 a^2 / (a^2 + t)]^{1/2}} \\ &= \pi \rho a^2 b \left( \int_0^\infty \frac{dt}{(a^2 + t)^{5/2}} + \frac{1}{2} e^2 a^2 \cdot \int_0^\infty \frac{dt}{(a^2 + t)^{7/2}} + O(e^4) \right). \end{aligned}$$

After integration, one obtains

$$(A.23) \quad \frac{1}{\alpha} = \pi\rho \frac{b}{a} \frac{1}{3} \left( 1 + \frac{3}{10}e^2 + O(e^4) \right).$$

The formula for  $1/\beta$  is (A.15) or

$$(A.24) \quad \begin{aligned} \frac{1}{\beta} &= \pi\rho a^2 b \\ &\cdot \int_0^\infty \frac{dt}{(a^2 + t)^{5/2} [1 - e^2 a^2 / (a^2 + t)]^{3/2}} \\ &= \pi\rho a^2 b \left( \int_0^\infty \frac{dt}{(a^2 + t)^{5/2}} + \frac{3}{2} e^2 a^2 \right. \\ &\quad \cdot \left. \int_0^\infty \frac{dt}{(a^2 + t)^{7/2}} + O(e^4) \right). \end{aligned}$$

After integration it yields

$$(A.25) \quad \frac{1}{\beta} = \pi\rho \frac{b}{a} \frac{1}{3} \left( 1 + \frac{9}{10}e^2 + O(e^4) \right).$$

Hence

$$(A.26) \quad \begin{aligned} \frac{\beta}{\alpha} &= \frac{1 + (3/10)e^2 + O(e^4)}{1 + (9/10)e^2 + O(e^4)} \\ &= 1 - \frac{3}{5}e^2 + O(e^4), \end{aligned}$$

$$(A.27) \quad \varepsilon^2 = 1 - \left( \frac{\beta}{\alpha} \right)^2 = \frac{6}{5}e^2 + O(e^4)$$

and

$$(A.28) \quad \frac{e^2 - \varepsilon^2}{1 - e^2} = -\frac{1}{5}e^2 + O(e^4).$$

Substituting expressions (A.27) and (A.28) into (A.21) we obtain

$$(A.29) \quad \begin{aligned} ds &= a[(1 - e^2)(1 - \varepsilon^2)]^{1/2} \\ &\cdot \frac{(1 - (1/5)e^2 \sin^2 \psi + O(e^4))^{1/2}}{(1 - (6/5)\varepsilon^2 \sin^2 \psi + O(e^4))^{3/2}} d\psi \\ &= a[(1 - e^2)(1 - \varepsilon^2)]^{1/2} \\ &\cdot \left( 1 + \frac{17}{10}e^2 \sin^2 \psi + O(e^4) \right) d\psi. \end{aligned}$$

The linearized form of (A.29) is

$$(A.30) \quad \begin{aligned} \Delta S &= \Delta\psi a[(1 - e^2)(1 - \varepsilon^2)]^{1/2} \\ &\cdot \left( 1 + \frac{17}{10}e^2 \sin^2 \psi \right), \end{aligned}$$

TABLE 8

Definition of latitude	Linearized constraint formula	Eccentricity
Elevation angle of center ray	$\Delta S = \Delta\phi A(1 + \frac{1}{2}e_C^2 \sin^2 \Phi)$	$e_C^2 = 2C$
Elevation angle of normal	$\Delta S = \Delta\lambda A(1 + \frac{3}{2}e_N^2 \sin^2 \Lambda)$	$e_N^2 = \frac{2}{3}C$
Elevation angle of plumb line	$\Delta S = \Delta\psi A(1 + \frac{17}{10}e_P^2 \sin^2 \Psi)$	$e_P^2 = \frac{10}{17}C$

where  $\psi$  is the midpoint latitude of the observed arc  $\Delta\psi$ .

To summarize, we have in all three cases a linearized constraint equation of the form

$$(A.31) \quad \Delta S = \Delta\lambda A(1 + C \sin^2 \Lambda),$$

whereby the meaning of the observed  $\Delta\lambda$  and  $\Lambda$  depends on the definition of the latitude. An adjustment provides numerical values for the constants  $A$  and  $C$ . These values depend on the data but are independent of the definition of the latitude. On the other hand, the formula for the computation of the eccentricity of the fitted ellipse from the value of  $C$  depends on the definition. The different formulas are summarized in Table 8. To derive relations among the ellipticities  $f$  that correspond to difference eccentricities  $e$ , let us assume that the square  $e^2$  of an eccentricity is related to the square  $e_N^2$  of the “normal” eccentricity by

$$(A.32) \quad e^2 = \gamma e_N^2.$$

Then one can derive from (2.4) the following relationship between the corresponding ellipticities:

$$(A.33) \quad f = \gamma f_N \frac{2 - f_N}{1 + \sqrt{1 - \gamma f_N(2 - f_N)}}.$$

The value of  $\gamma$  for the center ray is

$$(A.34) \quad \gamma_C = 3.$$

and the value of  $\gamma$  for the plumb line is

$$(A.35) \quad \gamma_P = \frac{15}{17}.$$

Equation (A.33) is used in Section 6 with the corresponding constraints  $\gamma_C$  and  $\gamma_P$  to compute the inverse ellipticities  $1/f_C$  and  $1/f_P$ , respectively.

## REFERENCES

- ACTON, F. S. (1990). *Numerical Methods That Work*. Math. Assoc. Amer., Washington, DC.
- CELMINŚ, A. (1979). A manual for general least squares model fitting, Ballistic Research Laboratory Report ARBRL-TR-02167, Aberdeen Proving Ground, MD. (AD #B040229L.)

- CELMINŠ, A. (1982). Least squares optimization with implicit model equations. In *Mathematical Programming with Data Perturbations II* (A. V. Fiacco, ed.) 131–152. Dekker, New York.
- GAUSS, C. F. (1809). *Theoria motus corporum coelestium in sectionibus conicis solem ambientium*. Perthes und Besser, Hamburg.
- GAUSS, K. F. (1963). *Theory of the Motion of the Heavenly Bodies Moving about the Sun in Conic Sections*. Dover, New York. [Reprint of translation of Gauss's *Theoria motus* by C. H. Davis, originally published (1857) by Little, Brown, Boston.]
- GAUSS, C. F. (1995). *Theoria combinationis observationum erroribus minimis obnoxiae (1821–1826)*. SIAM, Philadelphia. (Reprint with translation and afterword by G. W. Stewart.)
- KELLOGG, O. D. (1953). *Foundations of Potential Theory*. Dover, New York. [Reprint; originally published (1929) by Springer, Berlin.]
- PLACKETT, R. L. (1972). The discovery of the method of least squares. *Biometrika* **59** 239–251.
- SHEYNNIN, O. (1993). On the history of the principle of least squares. *Arch. Hist. Exact Sci.* **46** 39–54.
- SPROTT, D. A. (1978). Gauss's contributions to statistics. *Hist. Math.* **5** 183–203.
- STIGLER, S. M. (1977). An attack on Gauss, published by Legendre in 1820. *Hist. Math.* **4** 31–35.
- STIGLER, S. M. (1981). Gauss and the invention of least squares. *Ann. Statist.* **9** 465–474.