200 years of least squares method

▌Elemente der Mathematik

200 years of least squares method

A. Abdulle and G. Wanner

A. Abdulle received his Ph.D. in mathematics from the University of Geneva in 2001 under the direction of G. Wanner, and has actually a post-doc position at Princeton University.

G. Wanner received his Ph.D. in mathematics from the University of Innsbruck in 1965 under the direction of W. Gröbner. Since 1973 he is teaching at the University of Geneva.

Both authors are specialists in the numerical solution of differential equations, in particular Runge-Kutta and Chebyshev methods for stiff and parabolic problems, and share their enthusiasm for the three uppercase M's: Mathematics, Music and Mountains.

1 The discovery of Ceres

The discovery of *Ceres*, the first of the asteroids between Mars and Jupiter, was *the* great scientific event at the beginning of the 19th century with important consequences for the later development of science, despite the fact that this feeble light-point has hardly been seen ever since by non-specialists, and thus had absolutely no immediate 'practical' importance.

1.1 The rule of Titius-Bode

Das Daseyn dieses Planeten scheint insbesondere aus einem merkwürdigen Verhältniss zu folgen ... Sollte der Urheber der Welt diesen Raum leer gelassen haben?

(J.E. Bode, *Anleitung zur Kenntniss des gestirnten Himmels*, 6. Aufl., Berlin 1792, quoted in *Hegels Werke* 5, Anmerkungen p. 810)

Vor genau 200 Jahren begann Gauss, anlässlich seiner spektakulären Wiederentdeckung der Ceres, mit der systematischen Benützung seiner Methode der kleinsten Fehlerquadrate. Diese Methode ist aus dem heutigen wissenschaftlichen Rechnen nicht mehr wegzudenken. Dies bietet eine gute Gelegenheit für eine Retrospektive, einschliesslich Gauss' Berechnungen und der Querelen mit Hegel und Legendre. Die Arbeit schliesst mit einigen kleinen heutigen Anwendungsbeispielen, nämlich der Berechnung der Position der Kamera eines Photos, der „Korrektur" von Bildern Leonardo da Vincis und der Vorhersage eines Gletscherabbruchs.

The astronomer priests of Babylon discovered seven distinguished celestial bodies: first the Sun and the Moon, then Venus (1600 B.C.), finally Mars, Mercury, Jupiter, and Saturn. Soon, days were alternatively consecrated to these divinities (Sunday, Monday, Mardi, Mercredi, Jeudi, Vendredi, Saturday) and since thousands of years all human activity on the earth pulsates in this 7-days rhythm. All this time, nothing was added to these Babylonian Gods, until Sir William Herschel, a German organist and amateur astronomer living in England, discovered the 13th of march 1781 a new planet, through a huge telescope of his own construction. Herschel wanted to name 'his' new planet *Georgium sidus* (George's star), in devotion to the British King, but Bode's proposition *Uranus* (in Greek mythology the father of Saturnus) was felt less patriotic and became generally accepted.

The discovery of Uranus also revived the discussions about the formula of Johann Daniel Tietz (Titius) and Johann Elert Bode, which stated that the semi-major axes of the orbits of the planets were given by the rule

$$0.4, \quad 0.4 + 0.3 = 0.7, \quad 0.4 + 2 \cdot 0.3 = 1 \text{ (the earth)}, \quad \ldots \quad 0.4 + 2^{n-2} \cdot 0.3, \ldots .$$

This, for $n = 2, 3, 4, 6, 7$, determines quite precisely the orbits of the known planets, and still worked rather well, with $n = 8$, for the orbit of Uranus. However, the number $n = 5$ was missing, which suggested that the 'Creator of this world' has certainly not left empty this gap (see citation). To prove or disprove this conjecture became then a major scientific challenge of the time.

1.2 The thesis of Hegel

> Sehen Sie sich doch nur bei den heutigen Philosophen um, bei Schelling, Hegel, Nees von Esenbeck und Consorten, stehen Ihnen nicht die Haare bei ihren Definitionen zu Berge?
>
> (Brief von Gauss an Schumacher, 1.11.1844, *Werke* 12, p. 62)

Also one of the most influential philosophers, Georg Wilhelm Friedrich Hegel, took part in these discussions and submitted in 1801 his thesis (*Dissertatio philosophica de orbitis planetarum*, Ienae MDCCCI, *Werke* 5, pp. 221–253) at the University of Jena. He starts by 'proving' the laws of Kepler without any need of mathematics or physics, and, in the last part, turns his attention to Bode's rule. This latter had of course no philosophical contents. Now, we have to look up Plato's *Timei* and find the magic numbers 1, 2, 3, 4, 9, 8, 27, where we are allowed to replace the 8 by a 16 (!)[1]. Then we take the third roots of the fourth powers of these numbers, still replacing without hesitation ('ponamus') the number 1 by $\sqrt[3]{3}$, and we obtain the sequence[2]

$$1.4 \quad 2.56 \quad 4.37 \quad 6.34 \quad 18.75 \quad 40.34 \quad 81$$

in which in fact 'between the fourth and fifth position is a lot of space[3]; hence there is *no planet missing* in this gap.

---

1) '16 enim pro 8 quem legimus ponere liceat'.

2) Most of these roots are wrong in the last digit.

3) 'inter quartum et quintum locum magnum esse spatium'

It was of course bad luck that this conclusion was held up to ridicule precisely the same year by the discovery of Ceres. Needless to say that all this was not favourable to the mutual esteem between scientists and philosophers (see quotation).

1.3 The discovery of Piazzi

On January 1, 1801, the Italian astronomer Giuseppe Piazzi discovered in the Taurus constellation a tiny little spot, and was able to follow its orbit until the 11th of February, when illness, bad weather, and the approaching Sun interrupted the observations. He named it *Ceres Ferdinandea* (Ferdinand is another King's name). The data of Piazzi's observations of Ceres were published in the September issue of the *Monatliche Corre-spondenz* (see [19]). We present these values in Table 1.1, where the latitudes β_i are taken southward. The great challenge was now to rediscover this lost body towards the

1801	Longitude	Latitude		Longitude	Latitude
Jan. 1	$53^0\ 23'\ 06.38''$	$3^0\ 06'\ 45.16''$	23	$53^0\ 44'\ 12.46''$	$1^0\ 38'\ 46.78''$
2	$53^0\ 19'\ 38.18''$	$3^0\ 02'\ 26.46''$	28	$54^0\ 15'\ 18.52''$	$1^0\ 21'\ 04.92''$
3	$53^0\ 16'\ 37.70''$	$2^0\ 58'\ 08.04''$	30	$54^0\ 30'\ 10.52''$	$1^0\ 14'\ 14.24''$
4	$53^0\ 14'\ 21.44''$	$2^0\ 53'\ 51.98''$	31	$54^0\ 38'\ 05.58''$	$1^0\ 10'\ 51.02''$
10	$53^0\ 07'\ 57.64''$	$2^0\ 28'\ 53.64''$	Feb. 1	$54^0\ 46'\ 27.14''$	$1^0\ 07'\ 34.18''$
13	$53^0\ 10'\ 05.60''$	$2^0\ 16'\ 46.08''$	2	$54^0\ 55'\ 01.52''$	$1^0\ 04'\ 18.10''$
14	$53^0\ 11'\ 54.20''$	$2^0\ 12'\ 54.02''$	5	$55^0\ 22'\ 45.20''$	$0^0\ 54'\ 34.54''$
19	$53^0\ 26'\ 01.98''$	$1^0\ 53'\ 37.82''$	8	$55^0\ 53'\ 04.52''$	$0^0\ 45'\ 08.28''$
21	$53^0\ 34'\ 22.68''$	$1^0\ 46'\ 13.06''$	11	$56^0\ 26'\ 28.20''$	$0^0\ 35'\ 55.02''$
22	$53^0\ 39'\ 11.58''$	$1^0\ 42'\ 28.80''$			

Table 1.1 The observations of Piazzi

end of the year, and many astronomers tried to extrapolate as good as possible its orbit (Burckhardt, Olbers, Piazzi).

But a certain "Dr. Gauss in Braunschweig" computed a totally different solution "nach einem eigenthümlichen Verfahren" and published it the 29th of September 1801. Not satisfied with that, with an enormous computational effort, he recomputed and readjusted the parameters continuously, and finally arrived in December 1801 at the values presented in Table 1.2 [4].

Sonnenferne	$326^0\ 53'\ 50''$
Ω	$81^0\ 1'\ 44''$
Neigung der Bahn	$10^0\ 36'\ 21''$
Logarithmus der halben grossen Axe	0.4414902
Excentricität	0.0819603
Epoche: 31 Dec. 1800 mittl. helioc. Länge	$77^0\ 54'\ 29''$

Table 1.2 The elements of Ceres (Gauss Dec. 1801)

---

4) The argument of the perihelion is given by $w = $ Sonnenferne $- 180^0 - \Omega$.

The 7th of December 1801, Freiherr von Zach re-discovered Ceres precisely at the position predicted by Gauss. This event "proved him to be the first of theoretical astronomers no less than the greatest of arithmeticians" [5].

2 The first computations of Gauss

Die von Kreis- und Parabel-Hypothesen unabhängige Bestimmung der Bahn eines Himmelskörpers aus einer kurzen Reihe von Beobachtungen beruht auf zwei Forderungen: I. Muss man Mittel haben, die Bahn zu finden, die drei gegebenen vollständigen Beobachtungen Genüge thut. II. Muss man die so gefundene Bahn so verbessern können, dass die Differenzen der Rechnung von dem ganzen Vorrath der Beobachtungen so gering als möglich werden.

(Gauss, *Summarische Übersicht*; see [8], p. 148)

Gewiss, jeder der die Rechnungen kennt, die die Bestimmung der Elemente eines Planeten und dann jeder daraus herzuleitende Ort erfordert, muss es bewundern, wie ein einzelner Mann in so kurzen Zeiträumen so vielfache mühsame Rechnungen zu vollenden vermögend war.

(von Zach, März 1805, see Gauss *Werke* 6, p. 262)

Le ciel est simple is the leitmotiv of the amateur observatory in St. Luc, Switzerland, and one might agree at least as long one does not try to understand the computations of Gauss. The great advantage of Gauss' ideas over his rivals was, that he assumed solely Kepler's laws for his planet and no other hypotheses. But Gauss never revealed details of his calculations. Urged by Olbers, he finally sent in August 1802 a manuscript *Summarische Übersicht* without any desire to see it published. This text was finally printed in 1809 by von Lindenau (see [8]) with all the excuses of the editor for the many 'imperfections'. An excellent English description of Gauss' calculations has appeared recently (see [18]).

All the difficulty stems from the great number of variables involved. Indeed, we have to work with

$$
\begin{array}{ccccc}
\text{Elements of orbit} & & \text{Heliocentric} & & \text{Geocentric} \\
w \text{ arg. of perihelion} & & \text{coordinates} & & \text{spherical} \\
\Omega \text{ long. of ascend. node} & (A) & & (B) & \text{coordinates} \\
i \text{ inclination of orbit} & \Longleftrightarrow & \begin{pmatrix} x \\ y \\ z \end{pmatrix} & \Longleftrightarrow & \begin{pmatrix} \rho \\ \lambda \\ \beta \end{pmatrix} \\
a \text{ semi-major axis} & & & & \\
e \text{ eccentricity} & & & & \\
l_0 \text{ mean heliocent. long.} & & & &
\end{array}
\qquad (2.1)
$$

The quantities measured are the angles λ and β (the distance ρ is unknown, of course) for several time values, the quantities to be computed are the elements of the orbit. So we need formulas for the connecting passages (A) and (B).

---

5) W.W.R. Ball, *History of Mathematics*, London 1901, p. 458; quoted from R.E. Moritz, *Memorabilia Mathematica*, p. 157.
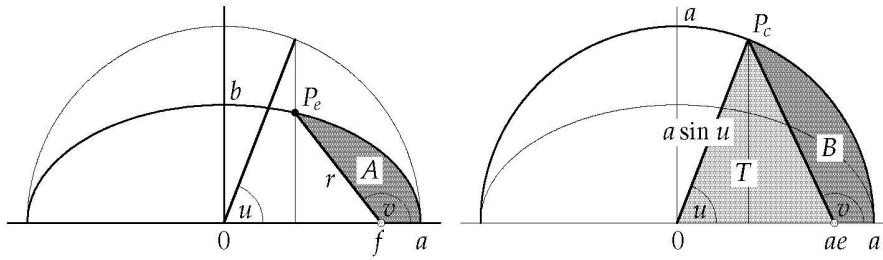
Fig. 2.1 Kepler orbit; P_e the planet, f focus (the sun), u the true anomaly, v the eccentric anomaly, a semi-major axis (perihelion), e eccentricity

Passage (A). For a given time t, we have first to find the position of the planet on the ellipse, i.e., to find the *eccentric anomaly* u (see Fig. 2.1, left). We first assume the point $t = 0$ at the perihelion. Then Kepler's second law ('same times, same areas') tells us that t is proportional to the area A. The period P of the orbit thereby corresponds to the total area $ab\pi$ of the ellipse, so we have

$$\frac{A}{ab\pi} = \frac{t}{P}.$$

We now stretch the ellipse to a circle (Fig. 2.1, right), so that $B = \frac{a}{b}A$, but also $B = \frac{a^2}{2}(u - e\sin u)$ (difference of the areas of a sector and the triangle T). The three equations lead to

$$nt = u - e\sin u \qquad \text{(Kepler's equation)} \tag{2.2}$$

where the constant $n = \frac{2\pi}{P}$ is called the *mean angular rate*. We finally transfer the origin of time to the correct place, i.e., t in (2.2) becomes $t - t_0$, where t_0 is the time of perihelion. Thus we have to add the epoch l_0 and to subtract the argument of the perihelion and the longitude of the ascending node. Then (2.2) becomes

$$l_0 - (w + \Omega) + nt = u - e\sin u. \tag{2.3}$$

To solve this transcendental equation for u we need to find the mean angular rate which is given by Kepler's *third* law. This law states that a^3 is proportional to P^2, i.e., that

$$n^2 a^3 \quad \text{is a known constant.} \tag{2.4}$$

Having now computed, for a given time t, the eccentric anomaly u, we express, using u, Ω, w, i, e, a and elementary spherical geometry, the position of the planet in the heliocentric coordinates (x, y, z) (see for example [4, pp. 182–186] or [3, pp. 84–90] for more details and explicit formulas).

Passage (B). For this, we have to know the solar geocentric coordinates (X, Y, Z) (for the same date and time) and we obtain the geocentric ecliptic coordinates of the planet by adding these and taking spherical coordinates

$$\begin{aligned}
\xi &= x + X = \rho\cos\beta\cos\lambda, \\
\nu &= y + Y = \rho\cos\beta\sin\lambda, \\
\zeta &= z + Z = \rho\sin\beta.
\end{aligned} \tag{2.5}$$

Gauss' procedure. At the time of the discovery of Ceres, it was well-known how to compute the six elements of the orbit of a planet from two sets of *heliocentric* coordinates x, y, z. This consists in solving 2×3 nonlinear equations in six unknowns. The great difficulty was that there were only *two geocentric* observed values λ_i, β_i per data point. After long formula manipulations of the above expressions, Gauss was able to reduce the computation of *one* set of heliocentric coordinates x, y, z to the knowledge of *two* sets of observed values λ_i, β_i via the solution of a complicated transcendental equation. Therefore, from *three* sets of observations, he obtained *two* sets of heliocentric coordinates and thus the desired elements. These developments, too long to be given here, are excellently presented in [18]; see also [11].

Thereby, it was advantageous to have the third point exactly in the middle of the two others. So Gauss started with the data Jan. 2, Jan. 22, and Feb. 11. The obtained values of the elements were then recomputed repeatedly by changing the dates, and by checking them for the remaining data. All these results of calculations, and Gauss' later results for the subsequent discoveries of *Pallas Olbersiana*, *Juno* and *Vesta*, are impressively documented in Gauss' *Werke*, vol. 6, pp. 199–402.

As a conclusion, we see that *these* computations were *not* performed with the method of least squares.

3 The method of least squares

> Der Verfasser gegenwärtiger Abhandlung, welcher im Jahre 1797 diese Aufgabe nach den Grundsätzen der Wahrscheinlichkeitsrechnung zuerst untersuchte, fand bald, dass die Ausmittelung der *wahrscheinlichsten* Werthe der unbekannten Grösse unmöglich sei, wenn nicht die Function, die die Wahrscheinlichkeit der Fehler darstellt, bekannt ist.
>
> (Gauss, Gött. gelehrte Anz. 33 (1821), pp. 321–327)

Things changed, however, after Ceres had been rediscovered in December 1801 and when more observations became available. Now the task was to improve the orbital elements to still higher accuracy with the help of all these new data. Here Gauss started to apply the *method of least squares*, again without ever revealing details to anybody. The only evidence is the last sentence of the *Summarische Übersicht*: "hat man schon Beobachtungen von 1 oder mehrern Jahren ..., so halte ich den Gebrauch der Differential-Änderung, wobei man eine beliebige Zahl von Beobachtungen zum Grunde legen kann, für das beste Mittel" – and the precision of the results (see also [17]).

Legendre contra Gauss. In 1805 appeared the work *Nouvelles méthodes pour la détermination des orbites des comètes* by A.-M. Legendre, containing in an appendix an extremely beautiful presentation of the *méthode des moindres quarrés*. The clarity of this work together with numerical examples made the least squares method immediately known in all scientific communities. Despite of Legendre's work, Gauss called, in his famous treatise [9] *Theoria motus corporum celestium*, published 1809, obstinately the least squares idea "my principle, which I have made use of since 1795". Legendre then protested in a long letter to Gauss, which is worth reading (Gauss *Werke* X/1, p. 380;

complete text in English translation in [15], pp. 242–243; "je n'ai jamais appellé *princip-ium nostrum* un principe qu'un autre avait publié avant moi"). Gauss never answered to Legendre, but mentioned to others the existence of a cryptic entry in his diary from June 17, 1798, simply saying "# Calculus probabilitatis contra La Place defensus" (see facsimile in Fig. 3.1). Legendre never forgot this, also because the young Jacobi (concerning

Fig. 3.1 Entry in Gauss' diary, June 17, 1798

elliptic functions, see [15], p. 246) and the young Bolyai (non-Euclidean geometry, see [2], p. 99) made similar experiences with Gauss.

Gauss' probabilistic justification of the least squares principle. Going much further than Legendre, Gauss gave an answer to the question: "Why least squares and not, for example, least fourth powers or least sixth powers?" To explain the idea, we treat a simple problem, i.e., the approximation of three 'observations' x_i, y_i ($i = 1, 2, 3$) by an 'orbit' which is a straight line

$$y = a + bx \tag{3.1}$$

(see Fig. 3.2). If now the three points don't lie on one line, there are three different lines,
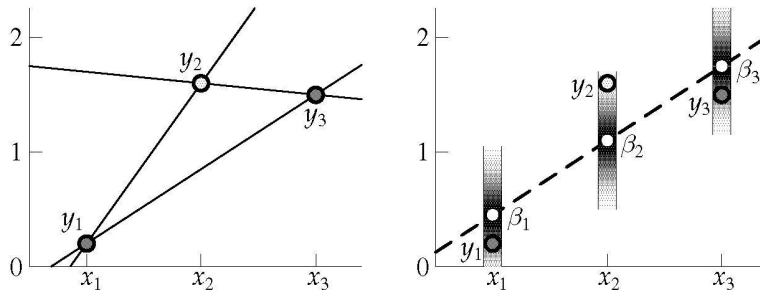

Fig. 3.2 Motivation for the least squares method

none of which is satisfactory (see the left picture). Now suppose that there *exist values*

$$\beta_i = a + bx_i \tag{3.2}$$

on a certain line, and that the measurements y_i are random samplings whose errors satisfy a certain probability law. The most common distribution is (see the right picture; the probability is symbolized by varying grey tones)[6]

$$P(0 \leq \beta_i - y_i \leq \Delta y) = \frac{e^{-\frac{(\beta_i - y_i)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} \, \Delta y.$$

6) Laplace arrived at this law from the binomial coefficients and a passage $n \to \infty$; for Gauss it was simply the law which reproduced the well-proved arithmetic mean.

Now, the probability of obtaining the *three* values y_1, y_2, y_3 (to a precision of Δy) is the product of the three above probabilities, i.e.,

$$\left(\frac{\Delta y}{\sigma\sqrt{2\pi}}\right)^3 \prod_{i=1}^{3} e^{-\frac{(\beta_i - y_i)^2}{2\sigma^2}} = \left(\frac{\Delta y}{\sigma\sqrt{2\pi}}\right)^3 e^{-\frac{\sum_{i=1}^{3}(\beta_i - y_i)^2}{2\sigma^2}}.$$

We have then *maximum likelihood* of our result, when this probability is maximal, i.e., when the exponent

$$\sum_{i=1}^{3}(\beta_i - y_i)^2 = \text{min !}$$

and with (3.2)

$$\sum_{i=1}^{3}(a + bx_i - y_i)^2 = \text{min !}$$

which is, precisely, the *principium nostrum*. Differentiating the last expression with respect to a and b we obtain

$$\begin{pmatrix} \Sigma 1 & \Sigma x_i \\ \Sigma x_i & \Sigma x_i^2 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \Sigma y_i \\ \Sigma x_i y_i \end{pmatrix} \quad \text{or} \quad A^T A \alpha = A^T y, \quad A = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{pmatrix}. \quad (3.3)$$

These are called the *normal equations*. Good luck, that the *principium* with best probabilistic justification also leads to the easiest possible problem to solve, a *linear* system of equations.

Further developments. Of the many important consequences which followed the least squares idea, we mention the following:

- *Gaussian elimination.* In order to prove the solvability of the normal equations, Gauss made in [9] the first clear description of the elimination algorithm for linear equations.

- *Gauss-Newton method.* In the same paper, Gauss also explained how *nonlinear* least squares problems are linearized in the neighbourhood of a first approximate solution, which is then iteratively refined.

- *Laplace's central limit theorem.* In 1809, Laplace published his central limit theorem, showing that *any* probability function, after taking arithmetic means, tends to the normal distribution for $n \to \infty$. Soon after, he extended this to justify the principle of least squares for arbitrary probability functions and $n \to \infty$. A great publication of all these results was *Théorie analytique des probabilités* from 1812.

- In 1823, Gauss publishes a second fundamental treatise on least squares, [10] *Theoria combinationis observationum erronibus minimis obnoxiae* in two parts, which contains a new justification of the least squares principle, independent of the probability function, which is today called the *Gauss-Markov theorem*.

- In 1828, Gauss publishes a *Supplementum*, which contains impressive calculations for the geodesic triangulations of the Netherlands and the country of Hannover.

- Also in 1828, Bessel discovers, originally for the discrete case, the relation between the least squares idea, the orthogonality relations, and the Euler-Fourier formulas for the trigonometric approximation. This discovery, extended by Gram (1883) to the continuous case, is the basis of the L^2 Hilbert space theory of Fourier series.

- In 1845, Jacobi publishes his method for solving the normal equations with the help of successive rotations in $\mathbb{R}^2$. These rotations lead in the 1950ies to Givens' method for triangularization and the first stable eigenvalue algorithm.

- In 1900 appears the classical paper of Karl Pearson [14], which combines the least squares method with the χ^2 distribution and led to the famous χ^2-test for the reliability of hypotheses.

- In 1958 appears Householder's reflection algorithm, which, by replacing Givens' rotations, leads to the QR decomposition, and, by Golub (1965), became the nowadays standard algorithm for least squares problems. All examples which follow, have been computed with this method, using a code written by E. Hairer for his course 'Analyse Numérique' (http://www.unige.ch/math/folks/hairer/polycop.html).

A complete modern treatment of numerics for the least squares method, which contains nearly 1000 bibliographical references, is the book of Björck [1]. Of valuable help for readers interested in Gauss' contributions is the bilingual edition of *Theoria combinationis observationum* with *Supplementum* and *Anzeigen*, and, most important, a carefully written *Afterword*, due to G.W. Stewart [16]. Many original texts translated into English can also be found in [12], Sect. 4.9 and 4.10.

The orbit of Ceres with the least squares method. In possession of a modern algorithm, we now want to compute the elements of the orbit of Ceres using the data of Piazzi with the least squares method. This we did as follows: for given orbital elements w, Ω, i, a, e, l_0, we designed a subroutine, computing with the aid of formulas (2.1) through (2.5), for the times t_i of Piazzi's observations, the geocentric longitudes λ_i and latitudes β_i. The necessary expressions for the solar heliocentric coordinates of the earth (X, Y, Z) were obtained from the server of *l'Institut de mécanique céleste et de calcul des éphémérides* (http://www.bdl.fr/ephemeride.html). These values compared to the *actual observations* $\widehat{\lambda}_i$, $\widehat{\beta}_i$ define a function

$$
\begin{aligned}
F : \mathbb{R}^6 \quad &\longrightarrow \quad \mathbb{R}^{38} \\
(w, \Omega, i, a, e, l_0) \quad &\longmapsto \quad \Big(\lambda_{t_1}(\ldots) - \widehat{\lambda}_1, \ldots, \lambda_{t_{19}}(\ldots) - \widehat{\lambda}_{19}, \\
&\qquad \beta_{t_1}(\ldots) - \widehat{\beta}_1, \ldots, \beta_{t_{19}}(\ldots) - \widehat{\beta}_{19} \Big),
\end{aligned}
\tag{3.4}
$$

and we have to find w, Ω, i, a, e, l_0 such that

$$
\|F(x)\|_2^2 = \sum_{i=1}^{19} \Big((\lambda_{t_i} - \widehat{\lambda}_i)^2 + (\beta_{t_i} - \widehat{\beta}_i)^2 \Big) = \min \, !
\tag{3.5}
$$

As initial values we chose values close to Gauss' values and after 5 Gauss-Newton iterations the least squares solution was precise to 7 digits (values displayed in Table 3.1).

Sonnenferne	$318^0\ 12'\ 27''$
Ω	$80^0\ 55'\ 9''$
Neigung der Bahn	$10^0\ 35'\ 38''$
Logarithmus der halben grossen Axe	0.4448506
Excentricität	0.0694885
Epoche: 31 Dec. 1800 mittl. helioc. Länge	$75^0\ 47'\ 31''$

Table 3.1 The elements of Ceres (least squares)

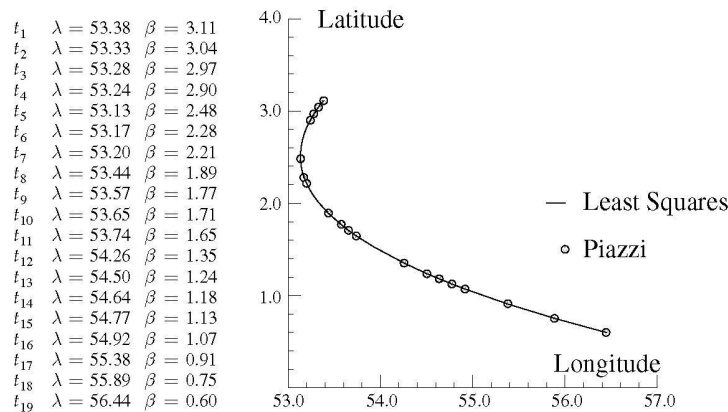|	
t_1	$\lambda = 53.38$ $\beta = 3.11$
t_2	$\lambda = 53.33$ $\beta = 3.04$
t_3	$\lambda = 53.28$ $\beta = 2.97$
t_4	$\lambda = 53.24$ $\beta = 2.90$
t_5	$\lambda = 53.13$ $\beta = 2.48$
t_6	$\lambda = 53.17$ $\beta = 2.28$
t_7	$\lambda = 53.20$ $\beta = 2.21$
t_8	$\lambda = 53.44$ $\beta = 1.89$
t_9	$\lambda = 53.57$ $\beta = 1.77$
t_{10}	$\lambda = 53.65$ $\beta = 1.71$
t_{11}	$\lambda = 53.74$ $\beta = 1.65$
t_{12}	$\lambda = 54.26$ $\beta = 1.35$
t_{13}	$\lambda = 54.50$ $\beta = 1.24$
t_{14}	$\lambda = 54.64$ $\beta = 1.18$
t_{15}	$\lambda = 54.77$ $\beta = 1.13$
t_{16}	$\lambda = 54.92$ $\beta = 1.07$
t_{17}	$\lambda = 55.38$ $\beta = 0.91$
t_{18}	$\lambda = 55.89$ $\beta = 0.75$
t_{19}	$\lambda = 56.44$ $\beta = 0.60$

Fig. 3.3 The computed and observed positions of Ceres

In Fig. 3.3 we compare Piazzi's observations and our computations. These latter computations fit better Piazzi's observations than Gauss' values, but Gauss' elements are closer to the true orbital elements. The reason is that some of Piazzi's measurements contain slight errors (this was already observed by Gauss), and that these errors influence enormously the solutions (this was also observed by Gauss). This phenomenon is today called *a badly conditioned problem*.

4 Some today's examples

The method of least squares is the automobile of modern statistical analysis;...

(The first sentence of Stigler [17])

Nearly everywhere, where data have to be analysed or models adjusted, one applies today the method of least squares, very often to problems of enormous dimensions. For particularly impressive examples and an advancement of the theory we refer the reader to a forthcoming book by Deuflhard [5]. Here, in this paper, we explain in some detail three nice examples from everyday life.

4.1 The position of a camera

Problem. We have a photograph (see Fig. 4.1), on which we distinguish a couple of points with measured local coordinates $(\hat{u}_k, \hat{v}_k)$. Of the same points, we determine the
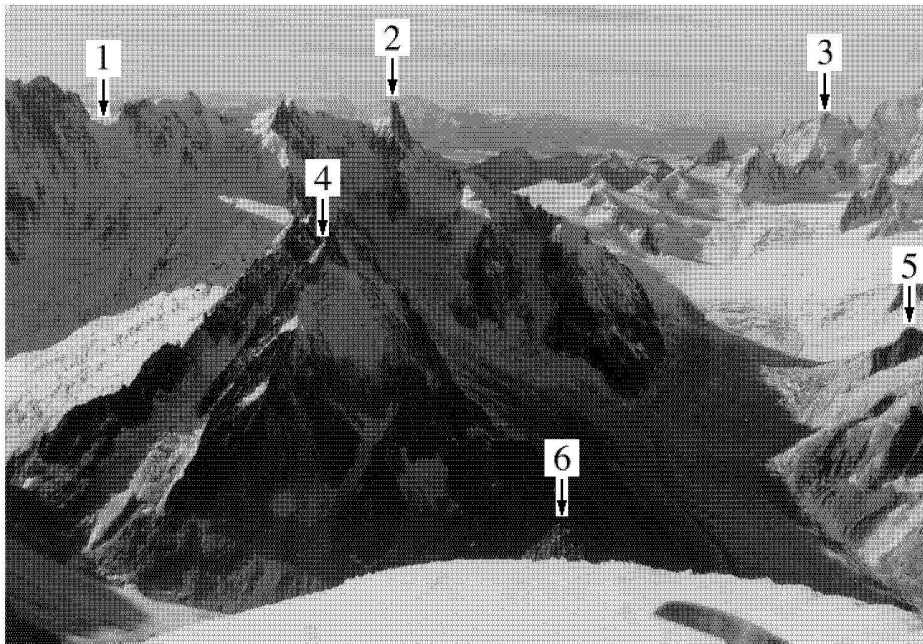
Fig. 4.1 A photograph from the Montblanc region (Photo: G. Wanner)

corresponding space coordinates (x_k, y_k, z_k) from a map, where the origin for x, y is placed arbitrarily and z are the altitudes. The task is to find out the position of the camera, its focus and its angles of inclination. A copy of the map, the Swiss national map 1:50 000 folio 5003, can be found under http://www.unige.ch/math/folks/ hairer/polycop.html. In Table 4.1 are given the values used in our calculations.

k	$\widehat{u}_k$	$\widehat{v}_k$	x_k	y_k	z_k
1. Col des Grandes Jorasses	−0.0480	0.0290	9855	5680	3825
2. Aiguille du Géant	−0.0100	0.0305	8170	5020	4013
3. Aig. Blanche de Peuterey	0.0490	0.0285	2885	730	4107
4. Aiguille du Tacul	−0.0190	0.0115	8900	7530	3444
5. Petit Rognon	0.0600	−0.0005	5700	7025	3008
6. Aiguille du Moine	0.0125	−0.0270	8980	11120	3412

Table 4.1 The data for the camera problem (in meters)

For the *solution* of our problem, we denote by $(\widetilde{x}, \widetilde{y}, \widetilde{z})$ the position in space of the camera's objective, and by $\vec{d} = (a, b, c)$ the perpendicular vector from the objective to the projection plane. Finally we allow the camera to be rotated around $\vec{d}$ by an angle θ. There are thus seven unknowns to determine. Very similar to the calculations of Gauss, but much easier, we have, once these 7 variables fixed, to find out the relations between
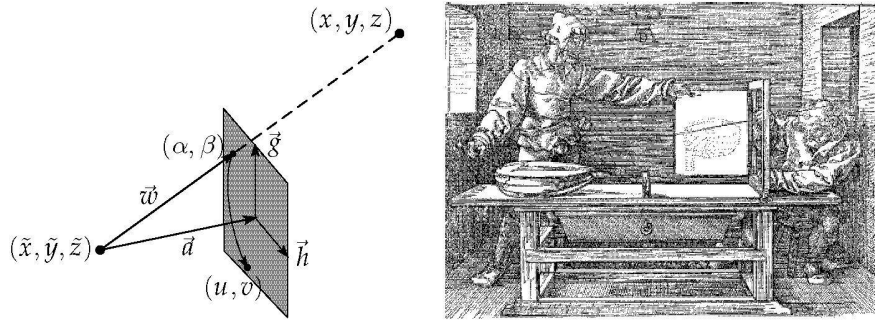
Fig. 4.2 Perspective representation: today (left) ... and 500 years ago (Dürer 1525, right)

the coordinates x, y, z in space and the corresponding projection points u, v on the photograph. For this, we fix two orthogonal vectors in the projection plane

$$\vec{h} = \frac{1}{\sqrt{a^2 + b^2}} \begin{pmatrix} b \\ -a \\ 0 \end{pmatrix}, \qquad \vec{g} = \frac{1}{\sqrt{(a^2 + b^2)(a^2 + b^2 + c^2)}} \begin{pmatrix} -ac \\ -bc \\ a^2 + b^2 \end{pmatrix}. \qquad (4.1)$$

Then, for a given point (x, y, z) (see Fig. 4.2) we compute a vector $\vec{w}$ by

$$\vec{w} = \lambda \cdot \begin{pmatrix} x - \tilde{x} \\ y - \tilde{y} \\ z - \tilde{z} \end{pmatrix} \qquad (4.2)$$

where the factor λ is determined by $\langle \vec{w} - \vec{d}, \vec{d} \rangle = 0$. Then $\alpha = \langle \vec{w}, \vec{h} \rangle$ and $\beta = \langle \vec{w}, \vec{g} \rangle$ are the coordinates of the projection point, which are finally rotated by θ:

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \qquad (4.3)$$

We have *then* the best solution, when these projected points (u_k, v_k), for the data (x_k, y_k, z_k), correspond in the best possible way to the measured data points $(\hat{u}_k, \hat{v}_k)$ of the photograph, i.e., according to 'principium nostrum', if

$$\sum_{k=1}^{6} \left((u_k - \hat{u}_k)^2 + (v_k - \hat{v}_k)^2 \right) = \min \, ! \qquad (4.4)$$

For this problem, the Gauss-Newton algorithm leads, with very rough initial values, after a couple of iterations, to the solution

$$\tilde{x} = 9679, \qquad \tilde{y} = 13139, \qquad \tilde{z} = 4131.$$

The photograph has thus been taken from the summit of the Aiguille Verte, whose altitude is known to be 4122 meters. The precision of these amateur calculations is not that of professional Swiss topographers.
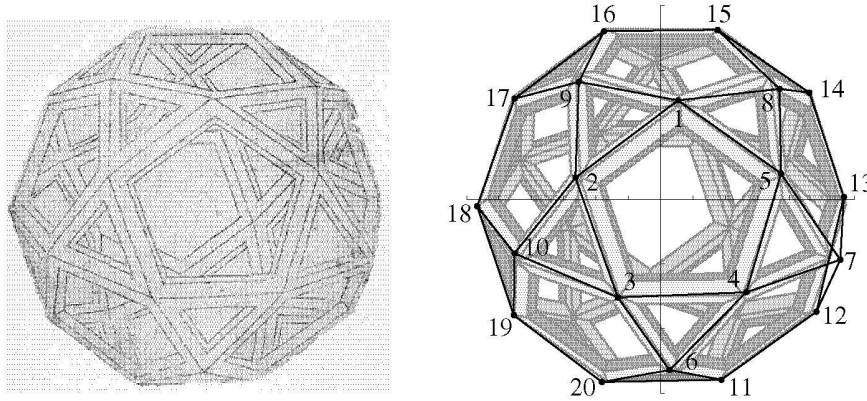
Fig. 4.3 Left: Drawing of Leonardo da Vinci (1510, Codex Atlanticus fol. 707r; Bibliotheca Ambrosiana, Milano), from [13], p. 100; right: Leonardo's vertices and, in grey, the 'corrected' drawing (Assyrus Abdullus & Gerhardus Wannerus, linguæ programmatoriæ Fortranus & Postscriptus, Calculatores SunBlade 100, Universitas Genavæ)

4.2 Leonardo's polyhedron

We now apply the same algorithm as above to an example from the history of art. The 'mountains' are now the exact vertices of a regular icosidodecahedron somewhere placed in space, and the 'photograph' is a drawing by Leonardo da Vinci (see Fig. 4.3, left) which was performed for the book *De divina proportione* by Luca Pacioli, Venice 1509. Pacioli says in his preface that it was done 'by the divine left hand of my friend Lionardo of Florence' (quoted from [6], p. 253). After having placed the 'camera' in the best possible position, we can project the exact vertices back to the picture and thus find out, nearly 500 years later with the aid of modern computing tools, the actual precision of Leonardo's drawing. The measured coordinates of the 20 visible vertices are given in Table 4.2.

k	u_k	v_k	k	u_k	v_k	k	u_k	v_k
1	5.409	30.691	8	36.865	34.219	15	17.609	52.536
2	−26.388	6.720	9	−25.283	36.394	16	−17.522	52.122
3	−13.259	−30.369	10	−45.244	−16.728	17	−45.244	31.161
4	26.517	−28.782	11	18.814	−55.828	18	−56.768	−2.147
5	37.265	8.054	12	48.271	−34.749	19	−45.433	−35.867
6	2.734	−52.888	13	56.767	0.764	20	−18.198	−56.563
7	55.650	−18.639	14	46.037	33.043			

Table 4.2 Measured vertices in Leonardo's drawing (in mm)

These points are re-drawn in the right picture of Fig. 4.3 in black, together with the 'corrected' polyhedron (in grey). We see that the drawing is very precise in the centre, but some vertices towards the periphery and in the background are less 'divine'.
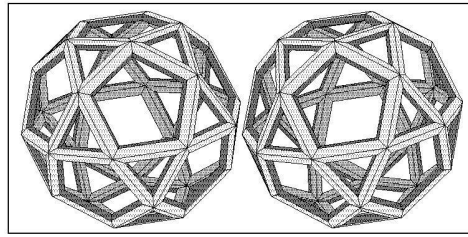
Fig. 4.4 Stereographic view of Leonardo's polyhedron (hold the picture rather close to the eyes and blend the
two drawings by staring through the paper)

Inspired by Leonardo's polyhedral skeleton, the authors have not resisted the pleasure
to produce a divine stereographic view of this beautiful object in Fig. 4.4.

4.3 Leonardo against Verrocchio

Dispirited is the pupil who does not surpass his master.

(Leonardo's maxim, see [13], p. 20)

In medieval paintings, holy persons were distinguished with a circular aureola behind
the head. During the Renaissance, progress of science transformed these aureolas into
ellipses. This gives us another occasion to submit the precision of one of Leonardo's
paintings under scientific investigation. We choose the painting *The Baptism of Christ*
from 1472, where the young pupil Leonardo added an additional angel to a painting of
his master Andrea del Verrocchio (see Fig. 4.5). We measure several points on the two
ellipses, and minimize for each of them the least squares problem

$$F = \sum_i \left(Ax_i^2 + 2Bx_iy_i + Cy_i^2 - Dx_i - Ey_i - 1 \right) = \min !$$

which (this time) is linear in the unknowns A, B, C, D, E. The comparison of the
minimal value of F for the two ellipses then showed that Leonardo *really* had already
'surpassed his master', although Verrocchio had an easier job, because a large part of
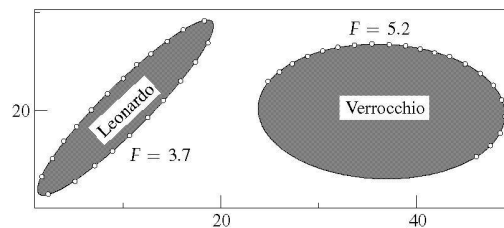his ellipse is not visible.

Fig. 4.5 Left: Leonardo's and Verrocchio's angels; right: best approximations of the ellipses

4.4 The hanging glacier above Grindelwald

In summer 1999 a hanging glacier high up in the mountains above Grindelwald (Switzerland) started to advance and threatened the region below by an enormous ice fall. In order to avoid a serious accident, a precise breaking off forecast was of great importance. Scientists from the ETH Zürich (M. Funk) therefore implanted a surveying stake on the ice masses (see Fig. 4.6, left) and observed carefully the advancing positions of the stake. The measured data are reproduced in Fig. 4.6 to the right. The time $t = 0$ corresponds to the 18th of July, 1999, at 7 a.m.
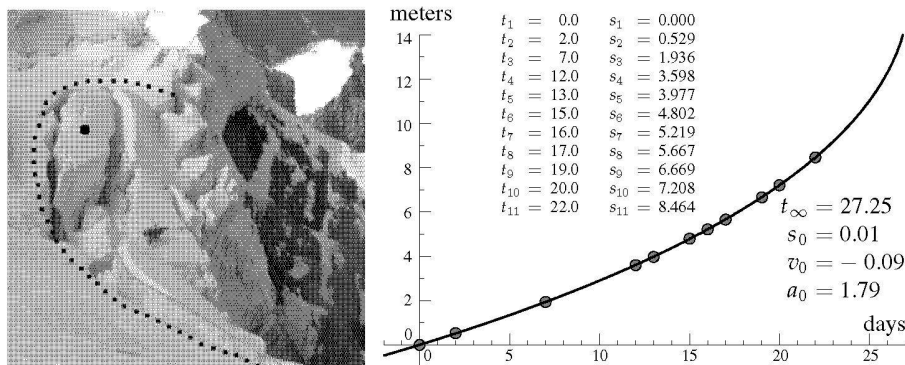

Fig. 4.6 Left: the Grindelwald hanging glacier with the surveying stake (Photo: M. Funk, from [7]); right: the data and the breaking off forecast

Earlier experiences with ice falls (in particular one at the *Weisshorn*) have shown that the increasing speed of such ice masses satisfies a formula

$$v(t) = v_0 + \frac{a_0}{(t_\infty - t)^n}$$

where $n \approx \frac{1}{2}$. By integrating this, we obtain for the positions

$$s(t) = s_0 + v_0 t + a_0 \left(\frac{(t_\infty - t)^{1-n} - t_\infty^{1-n}}{n - 1} \right). \tag{4.5}$$

The problem is now, to determine the unknown constants s_0, v_0, a_0 and t_∞ in such a way, that this function approaches the measured data points with minimal least squares error. The solutions obtained in this way are given in Fig. 4.6, and predicted the ice fall for $t_\infty = 27.25$, which corresponds to the 14th of August at 1 p.m. Actually, the glacier fell the 14th of August at 2 a.m. The forecast, 5 days before the event, was thus wrong by less than half a day. For more details, see [7].

Conclusion. After having seen, how the observations of a couple of stars have helped to develop modern science in such an extraordinary way, we *must* say, *really*, that stars influence our lifes, just not the way readers of horoscopes are believing.

References

[1] Björck, Å.: *Numerical methods for least squares problems*, 408pp., SIAM 1996.

[2] Bühler, W.K.: *GAUSS, eine biographische Studie*, Springer Verlag, 1987.

[3] Collins II, G.W.: *The Foundation of celestial mechanics*, Astronomy and Astrophysics Series, vol.16, Pachart Publishing House, Tuscon, 1989.

[4] Danjon, A.: *Astronomie générale*, Albert Blanchard, Paris, Seconde éd., 1986.

[5] Deuflhard, P.: *Newton methods for nonlinear problems. Affine invariance and adaptive algorithms*, to appear, Springer Verlag, 2002.

[6] Field, J.V.: Rediscovering the Archimedean polyhedra: Piero della Francesca, Luca Pacioli, Leonardo da Vinci, Albrecht Dürer, Daniele Barbaro, and Johannes Kepler, *Arch. History Exact Sc.*, 50 (1996), 241–289.

[7] Funk, M. & Minor, H.-E.: Eislawinen in den Alpen: Erfahrungen mit Schutzmassnahmen und Früherkennungsmethoden, *Wasserwirtschaft* vol. 91 (2001), 362–368.

[8] Gauss, C.F.: *Summarische Übersicht der zur Bestimmung der Bahnen der beiden neuen Hauptplaneten angewandten Methoden*, Monatliche Correspondenz, herauseg. Freiherr von Zach, Sept. 1809, Werke, vol. 6, 148–165.

[9] Gauss, C.F.: *Theoria motus corporum coelestium*, Perthes et Besser, Hamburgi (1809), Werke vol. 7, 1–288.

[10] Gauss, C.F.: *Theoria combinationis observationum erronibus minimis obnoxiae*, Pars Prior et Pars Post., Comm. Soc. Reg. Scient. Gott. 5, (1823), Werke, vol. 4, 1–26; 27–53; Supplementum Comm. Soc. Reg. Scient. Gott. 6, (1828), Werke, vol. 4, 55–93.

[11] Gauss, C.F.: *Elliptische Bahnbestimmung*, aus Gauss' Nachlass, Werke, vol. 11, 221–252.

[12] Goldstine, H.H.: *A history of numerical analysis from the 16th through the 19th century*, Springer Verlag, 1977.

[13] Letze, O. & Buchsteiner, T.: *Leonardo da Vinci, scientist inventor artist*, Exhibition Catalog 1999, Verlag Gerd Hatje, Ostfildern-Ruit, Germany.

[14] Pearson, K.: On a criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling, *Phil. Mag.* (5) 50 (1900), 157–175; corr. *Phil. Mag.* (6) 1 (1901), 670–671.

[15] Plackett, R.L.: Studies in the history of probability and statistics. XXIX, The discovery of the method of least squares, *Biometrika* 59 (1972), 239–251.

[16] Stewart, G.W.: Gauss, Theory of the combination of observations least subject to errors, bilingual edition of [10] with an *Afterword*, Classics in Appl. Math., SIAM 1995.

[17] Stigler, S.M.: Gauss and the invention of least squares, *The Annals of Stat.* 9 (1981), 465–474.

[18] Teets, D. & Whitehead, K.: The discovery of Ceres: How Gauss became famous, *Math. Magazine* 72 (1999), 83–93.

[19] Freiherr von Zach: *Fortgesetzte Nachrichten über den längst vermutheten neuen Haupt-Planeten unseres Sonnen-Systems*, in Gauss Werke, vol. 6, 199–204.

Assyr Abdulle
Mathematics Department (PACM)
Princeton University
Princeton, NJ 08544, USA
e-mail:
aabdulle@math.princeton.edu

Gerhard Wanner
Section de Mathématiques
Université de Genève
C.P. 240
CH-1211 Genève 24, Switzerland
e-mail:
Gerhard.Wanner@math.unige.ch