Chapter 1
Elliptic Functions

Scarcely had my work seen the light of day, scarcely could its title have become known to scientists abroad, when I learned with as much astonishment as satisfaction that two young geometers, MM. Jacobi of Königsberg and Abel of Christiania, had succeeded in their own individual work in considerably improving the theory of elliptic functions at its highest points.

Legendre

1.1 Legendre

On 28 June 1830 the Académie des sciences in Paris, the leading scientific institution of the day, announced that its Grand Prize in mathematics of 3,000 francs devoted to work which "presents the most important application of mathematical theories …or which contains a very remarkable analytical discovery" would be divided equally between Carl Gustav Jacob Jacobi in Königsberg and the family of the late Niels Henrik Abel of Christiania.[1] The Grand Prize was awarded regularly by the Académie, usually for meritorious answers to problems posed by the Académie itself. On this occasion they had decided instead to award the prize to two foreigners, Jacobi, then only 25, and Abel, who had died on April 6, 1829 at the age of 26, for work which had caused a sensation in the small but growing European mathematical community.

One reason for this decision lies behind the scenes with the elderly but still active and influential French mathematician Adrien Marie Legendre (Fig. 1.1). For 40 years he had published papers and books on what are called elliptic integrals. This work had not attracted the interest of others to the degree that he had hoped until unexpectedly in 1827 Abel and Jacobi had taken it up. However, they did not

[1] *Procès-verbaux des séances de l'Académie* 9 (1921), 466. The Commission consisted of Poisson, Poinsot, Legendre, with Lacroix as rapporteur.

U. Bottazzini and J. Gray, *Hidden Harmony—Geometric Fantasies*, Sources and Studies in the History of Mathematics and Physical Sciences, DOI 10.1007/978-1-4614-5725-1_2,

Fig. 1.1 Adrien-Marie Legendre (in Duren "Changing faces: the mistaken portrait of Legendre", *Notices of the AMS*, 56, 2009, 1440–1443)

so much extend Legendre's ideas as completely reformulate them. Each brought to the topic two remarkable new perceptions which enabled them to give it from the start a high degree of elegance, profundity, and power. These ideas were to invert the integrals, as we shall describe below, and to let them have complex endpoints, which is why their work was to be such a stimulus for later workers interested in the theory of complex functions. Of the two steps, Jacobi was often to remark later that the introduction of imaginaries on its own was enough to solve all the riddles of the early theory (see Dirichlet 1852, 10). Moreover, if at first Abel and Jacobi had seemed to be working in complementary ways an evident rivalry soon developed. Legendre was kept in touch mostly by Jacobi, who prudently wrote to him often about his work, for after publishing only a few short papers Jacobi had decided to publish his new theory of elliptic integrals as a book. Abel, however, published prolifically in a new German mathematical journal, although he too corresponded briefly with Legendre.

Legendre was delighted with the attention his favourite subject was receiving and generous enough to welcome it in its new and much altered form. He ensured that the Paris Académie asked him and Poisson to report on it, and Poisson presented their report on Jacobi's work on 21 December 1829.[2] This report, based largely on Jacobi's book, which had been published that April, helped ensure that the German was elected a corresponding member of the Académie on the 8th of February

[2] *Procès-verbaux des séances de l'Académie* 9 (1921), 373. Poisson's *Rapport* was published as his (1831).

1830.[3] By that time Poisson was one of the outstanding French mathematicians, if not the most authoritative after Laplace's death. A former, brilliant pupil at the École Polytechnique, and a professor there from 1802, Poisson, with the support of Lagrange and Laplace, had risen very quickly to the top of Paris mathematical milieu. In 1812 he was elected to the Institut de France, where he played an increasingly influential role.[4] Support like this undoubtedly helps explain why Abel and Jacobi's work was taken up when almost contemporary achievements of similar magnitude, such as the discovery of non-Euclidean geometry by Lobachevskii and Bolyai and the ideas of Galois, were not.

The widespread recognition of the importance of elliptic functions as a new but central domain in mathematics is worth documenting and explaining. It is one reason the subject of complex function theory was to grow so rapidly, for in the study of elliptic integrals, as reformulated by Abel and Jacobi, there was much to be done at every level from the deepest to the routine and superficial. Part of the explanation therefore lies in the traditional aspect of the new mathematics, its expression in terms of functions and their integrals. This "routine" mathematics was chiefly the creation of Euler and Lagrange. So far as questions in advanced calculus or analysis were concerned, it was a body of techniques highlighted by certain attractive results.[5]

The calculus of the eighteenth century was remarkably algebraic or formal. Functions had been defined quite generally by Euler on the input–output model but were at heart regarded either as closed expressions or as infinite series.[6] The meaning of x in, say,

$$(1-x)^{-1} = 1 + x + x^2 + \cdots$$

was deliberately vague. It is not true that Euler had no theory of convergence, or that he dealt with power series purely formally. He had, rather, several profound ideas, ranging from the now-standard elementary one to more sophisticated limit-processes to more formal ideas about series independent of their convergence, and he presented each of these views when he thought the context demanded it. The effect, nonetheless, was that he did not give convergence questions anything like the significance that Cauchy was to attach to them. The situation is more complicated with Lagrange. On the one hand it was a matter of definition for him that every function had a Taylor series, the coefficients of which were his definition of the derived functions of a function. On the other hand, he frequently gave estimates of the error involved in replacing such a series by a finite, polynomial expression such as its first n terms. As Grabiner has shown, Cauchy may well have drawn his

[3] *Procès-verbaux des séances de l'Académie* 9 (1921), 400.

[4] For Poisson's life and work, see Métivier et al. (1981).

[5] The work of Euler and Lagrange dominates eighteenth-century mathematics. On Euler, the reader may start with the online Euler Archive, which gives access to all of his works as well as many commentaries. Burckhardt's *Euleriana* updated to 1983 fills up pages 511–552 of Euler (1983a). As for Lagrange's life and work, see Loria (1913), Burzio (1963), and Borgato and Pepe (1990).

[6] For a recent discussion, see Ferraro (2008).

ideas about analysis in terms of the $\varepsilon-\delta-N$ methodology from yoking together these two apparently different ideas in Lagrange's work.[7] For various reasons and in various ways power series expansions were a legitimate and common device in the study of functions, albeit of a formal kind to modern eyes. They retained much of this character, for example, in the work of Jacobi.

Power series methods were commonly used in the study of differential equations when more informative solution methods failed. There was, by 1800, an impressive literature on ordinary differential equations. Linear equations had been given an (admittedly formal) theory by Lagrange, and some, like the hypergeometric equation, had been the object of more detailed study (as will be seen below). It had been well understood, for example, since Euler's work in the 1740s, that an nth order linear ordinary differential equation had a general solution depending on n arbitrary constants, or, equivalently, was a sum of n distinct solutions.[8] What was not at all understood was the nature of the singular points; this, as we shall see with Gauss, (see Sect. 1.5.2) had to wait for a firm grounding in the theory of a complex variable.

The contribution of Abel and Jacobi was to create a theory of functions at once traditional and novel; traditional in its algebraic aspect; novel in that it was, in ways that required to be understood, truly a theory of functions of a complex variable.

As for the topic of elliptic integrals, it had a long-established place in contemporary astronomy. Since Kepler's second law asserts that an elliptical orbit is parameterised by a satellite sweeping out equal areas in equal times, mathematicians were led straight away into questions involving the rectification of an ellipse and so to elliptic integrals. Newtonian theory then said that the orbit would be an ellipse only if the problem was a two-body one. For a 3- or n-body problem, the question was to compute the additional variation of that ellipse. The details of perturbation theory had called on some of the deepest ingenuity of eighteenth-century mathematicians such as Clairaut, Euler, Lagrange, and Laplace and acquired a new urgency in the early nineteenth century with the discovery of the asteroids. Moreover, since astronomical work is highly numerical, it lent some force to the whole thrust of Legendre's work, which was the simplification of the general elliptical integral, and the subsequent computation of the values of elliptical integrals as functions of the coefficients and their upper end points.

The computation of values of elliptic integrals proved to be far from the dreary task it might seem. The analogy between the trigonometric and elliptic integrals is helpful here as elsewhere. To compute tables of, say, the sine function, one would make repeated use of the addition formulae

$$\sin(u+v) = \sin u \cos v + \sin v \cos u,$$
$$\cos(u+v) = \cos u \cos v - \sin u \sin v,$$

[7] Grabiner (1981), see also Bottazzini (1992a).

[8] Publication is another matter and seems to have begun with Euler (1750).

with their important corollaries

$$\sin 2u = 2\sin u \cos u,$$
$$\cos 2u = \cos^2 u - \sin^2 u,$$

and their consequences

$$\sin(u/2) = ((1-\cos u)/2)^{1/2},$$
$$\cos(u/2) = ((1+\cos u)/2)^{1/2},$$

together with known values of cosine and sine such as $\cos 0 = 1$, $\sin 0 = 0$, $\cos \pi/2 = 0$, $\sin \pi/2 = 1$. Now, in terms of integrals,

$$u = \int_0^v \frac{dt}{\sqrt{1-t^2}}$$

may be taken to define the function $v = \sin u$, and the simplest and paradigmatic elliptic integral is

$$u = \int_0^v \frac{dt}{\sqrt{1-t^4}}. \tag{1.1}$$

It measures arc-length along the lemniscate $r^2 = \cos 2\theta$, which is a curve in the shape of a figure eight.[9] This integral can be regarded as defining either u as a function of v or v as a function of u. In either case, one is led to look for analogues of the behaviour of the trigonometric functions.

A remarkable analogy had been discovered by an Italian mathematician, Count Fagnano, in 1714 and republished by him in 1750, on the occasion of his submitting his life's work to the Berlin Académie (see Enneper 1876, Note III). His Fagnano (1750) was sent to Euler to read, with gratifying and dramatic effect. Fagnano had showed that a certain algebraic change of variable, which he gave explicitly, established that if

$$\int_0^v \frac{dt}{\sqrt{1-t^4}} = 2\int_0^w \frac{dt}{\sqrt{1-t^4}} \tag{1.2}$$

then v and w are algebraically related, so although arc-length along the lemniscate is a transcendental function of the parameter, values of the parameter for an arc and an arc of twice the length are algebraically related. Fagnano went on to show

[9] Differentiating the equation for the lemniscate gives $rdr = -\sin(2\theta)\,d\theta$. The element of length is given by $ds^2 = dr^2 + r^2 d\theta^2$, and eliminating θ gives this expression for arc-length: $ds^2 = \frac{dr^2}{1-r^4}$, whence the claim. The total arc length of the lemniscate is denoted 2ω.

that dividing the lemniscatic arc into 2.2^m, 3.2^m, or 5.2^m pieces was also possible algebraically. Euler was already interested in elliptic integrals and had studied of them in their setting of elastic beams.[10] On December 20, 1738 he had written to Daniel Bernoulli to say that he had come across "a singular property of the rectangular elastica"[11] corresponding to

$$\int_0^1 \frac{dx}{\sqrt{1-x^4}} \cdot \int_0^1 \frac{x^2 dx}{\sqrt{1-x^4}} = \frac{\pi}{4}.$$

Euler was immediately excited and generalised Fagnano's results to obtain a general addition theorem. He interpreted them as concerning solutions to the differential equation

$$\frac{dx}{\sqrt{1-x^4}} = \frac{dy}{\sqrt{1-y^4}}, \tag{1.3}$$

and by analogy with the trigonometric case

$$\frac{dx}{\sqrt{1-x^2}} = \frac{dy}{\sqrt{1-y^2}}$$

guessed correctly that the solution of this differential equation was

$$x^2 + y^2 = c^2 + 2xy(1+c^4)^{1/2} - c^2x^2y^2, \tag{1.4}$$

where c is an arbitrary constant. This fascinated him because it is an algebraic equation connecting solutions to a differential equation. It also means that

$$\int_0^x \frac{dt}{\sqrt{1-t^4}} + \int_0^y \frac{dt}{\sqrt{1-t^4}} = \int_0^c \frac{dt}{\sqrt{1-t^4}}, \tag{1.5}$$

so the result of adding together any two arcs of a lemniscate is an arc whose parameter, c, depends only algebraically on the parameter values of the original arcs. Equation (1.4) is called Euler's algebraic addition theorem for elliptic integrals. It makes explicit what Fagnano had also noticed, that doubling the lemniscatic arc

[10]The connection of elliptic integrals with elastic beams went back to Jakob Bernoulli who in his study of the elastica in the 1690s was led to consider the differential $dy = \frac{x^2 dx}{\sqrt{c^4 - x^4}}$ (c constant). After some attempts at expressing y in terms of exponentials he stated he had "weighty grounds" for believing that the integral could not be expressed by means of quadratures or rectifications of any conic section.

[11]According to Truesdell (1960, 174) Euler seemed "particularly proud" of this result, "and he comes back to it again and again, until finally it reveals itself to him as only a special case of the addition theorem for elliptic integrals".

yields a quartic equation, whereas doubling the circular arc yields only a quadratic equation.

This could have been the crucial step towards a theory of “lemniscatic integrals”

$$u = u(v) = \int_0^v \frac{dt}{\sqrt{1-t^4}} \tag{1.6}$$

and so it was to be, but not for Euler. The reason seems to have been, as Krazer argued (1912), that on this occasion Euler remained too close to the geometry of conics and lemniscates with which he began. Nor did anyone else pick up the problem for a generation. As Enneper pointed out (Enneper 1876, 542) “the writings of Fagnano seem to have been very little known, and the writings of the great Euler too little read”. The man who first studied the lemniscatic integral from a purely functional point of view was Legendre.

Like Euler, Legendre began, in his “Mémoire sur les intégrations par d’arcs d’ellipse” (1788a) with geometrical questions.[12] But in his “Mémoire sur les transcendantes elliptiques”, which he read to the Paris Académie in 1792, his attention was directed towards the new functions the integrals define, as is made clear by the title and still more by the sub-title: “Containing easy methods of comparing and valuing these transcendentals, which include elliptic arches, and which frequently occur in the application of the Integral Calculus”.[13] In this paper he showed for the first time how any integral of the form $\int \frac{Pdx}{R}$, where P is a rational function in x and R is the square root of a quartic (with real coefficients), can be simplified to one of the form

$$\int \frac{Qdt}{\sqrt{(1-t^2)(1-c^2t^2)}} \,. \tag{1.7}$$

The natural substitution $t = \sin\phi$ further reduces this to

$$\int \frac{Qd\phi}{\sqrt{(1-c^2\sin^2\phi)}} \,. \tag{1.8}$$

The variable ϕ Legendre called the amplitude of the elliptic integral, the parameter c the modulus, and quantity b defined as $\sqrt{1-c^2}$ the complementary modulus. Writing Δ or $\Delta(\phi)$ for the square root $\sqrt{(1-c^2\sin^2\phi)}$, Legendre showed how this reduction led to one of the three distinct kinds, a classification employed by

[12] As he did in Legendre (1788b).

[13] According to the English translator, Lacroix added a note to his *Traité* (1797–1798) to say that this interesting memoir had become very rare, and the translation was published at the request of several eminent mathematicians; see Leybourn’s *Mathematical Repository*, (2) 2 (1809, 1).

everyone since and which he was rightly to say in his *Traité* (1825–1832) 33 years later was at the basis of his method. In his terminology:
Elliptic integrals

$$\int_0^x \frac{d\phi}{\Delta} \tag{1.9}$$

are of the first kind;
elliptic integrals of the form

$$\int_0^x \Delta d\phi \tag{1.10}$$

are of the second kind; and elliptic integrals of the form

$$\int_0^x \frac{d\phi}{(1+n^2\sin^2\phi)\Delta}, \tag{1.11}$$

where n may be real or complex, are of the third kind.

All the integrals were regarded as functions of their upper end point x. He wrote F for a typical integral of the first kind and E (or G) for one of the second kind. The complete integrals $\int_0^{\pi/2} \frac{d\phi}{\Delta}$ and $\int_0^{\pi/2} \Delta d\phi$ he denoted F^1 and E^1, respectively, or $F^1(c)$ and $E^1(c)$ when he wanted to think of them as functions of the modulus c. He said it was indispensable that the modulus c and the amplitude ϕ were real, and that c be less that 1. Amongst other results, he showed that Euler's differential equation $\frac{dx}{R(x)} = \frac{dy}{R(y)}$, where R is the square root of a quartic, reduced to

$$\frac{d\phi}{\sqrt{1-c^2\sin^2\phi}} = \frac{d\psi}{\sqrt{1-c^2\sin^2\psi}}, \tag{1.12}$$

and that its integrals satisfy $F(\phi)+F(\psi)=F(\mu)$, where μ is an arbitrary constant. But when $\phi = 0, \psi = \mu$, so a comparison with Euler's result shows that

$$\cos\phi\cos\psi - \sin\phi\sin\psi\Delta(\mu) = \cos\mu. \tag{1.13}$$

This is his functional form of Euler's addition theorem (1.4) for elliptic integrals. The bulk of the paper was given over to a sketch of how tables of the values of $F(\phi)$ could be calculated for specified values of the modulus c.

That Legendre had thoroughly adopted the function-theoretic point of view is even clearer in his *Exercises de calcul intégral* (1811–1817) which begins with his study of "Des fonctions elliptiques". He meant $F(\phi)$, $E(\phi)$, and related functions, but to avoid confusion with their inverses, which we nowadays call elliptic functions, we shall call Legendre's functions elliptic integrals. The book is dominated by the spirit of Euler, and Legendre thought of himself as enriching mathematics as Euler had done with whole families of new, interesting functions. To this end he showed in Vol. II (1817) how they can be calculated numerically, then how the new functions can be used in geometrical and mechanical problems (see Sect. 4.6 below), and then he wrote extensively on Eulerian integrals. He devoted Vol. III (1816, sic) to the production of several sets of tables, which were also published independently, and computed tables for the Beta and Gamma functions which Euler had introduced (see Sect. 2.2 below) and shown, inter alia, to be closely connected to elliptic integrals.

The *Exercises* is a summary of his life's work on the subject, and it often builds on what he had written earlier. From his study of elliptic integrals (1792) he now deduced by differentiating under the integral sign that the complete elliptic integrals satisfy linear differential equations

$$(1-c^2)\frac{d^2F^1}{dc^2}+\frac{1-3c^2}{c}\frac{dF^1}{dc}-F^1=0 \tag{1.14}$$

(an equation first given by Euler in 1750)[14] and

$$(1-c^2)\frac{d^2E^1}{dc^2}+\frac{1-c^2}{c}\frac{dE^1}{dc}+E^1=0, \tag{1.15}$$

whence he obtained power series expansions for the complete integrals $F^1(c)$ and $E^1(c)$. He also established a strikingly attractive result connecting complete integrals of the first two kinds with complementary moduli (c and b):

$$\frac{\pi}{2}=F^1(c)E^1(b)+F^1(b)E^1(c)-F^1(b)F^1(c). \tag{1.16}$$

In this work he also showed (Vol. I, Sects. 17–22) how to calculate values of F by showing how to give accurate approximations when c is nearly 0 or 1, and how to reduce the general case to this one by a transformation. He had done this earlier, in his (1788a, b) and his (1792).

By analogy with the trigonometric case, he defined ϕ_n to be an amplitude such that $F(\phi_n)=nF(\phi)$ and sought to find $\sin(\phi_n)$ and $\cos(\phi_n)$ in terms of $\sin\phi$ and $\cos\phi$. He pointed out that this was easy when the modulus was 0 or 1 because then the elliptic integral can be evaluated explicitly. For example, when $c=0, F(\phi)=\phi$, which implies that $\phi_n=n\phi$. When c lies between 0 and 1 the addition formula gave him expressions for $\cos(\phi_2)$, $\sin(\phi_2)$ and $\tan(\phi_2)$. Identities between $\sin(\phi_{n+1})$, $\sin(\phi_{n-1})$ and $\sin(\phi_n)$ enabled him to obtain formulae for $\sin(\phi_3)$, $\sin(\phi_4)$, and so on. The formulae struck him as unexpectedly complicated. The trisection formula, for example, relating $\sin(\phi_3)=a$ to $\sin(\phi)=x$, is

[14] See Euler, *O.O.* (1) 20, 40.

$$\frac{(3-4(1+c^2)+6c^2x^4-c^4x^8)x}{1-6c^2x^4+4c^2(1+c^2)x^4-3c^4x^8}=a \tag{1.17}$$

It is of degree 9 in x, not of degree 3 as a näive analogy with trigonometry suggests. Quite generally, to divide ϕ into n parts leads to an equation of degree n^2, although the complexity of the formula was to invite later authors to fruitful reflection.

This approach to constructing a table of values therefore seemed blocked, but Legendre showed that $\tan\frac{1}{2}(\phi_{n+1}+\phi_{n-1})=\Delta\tan\phi_n$. Since Δ does not depend on n this formula can be iterated to give successive values of ϕ_n in terms of n. Legendre, however, preferred a second method, which altered the value of the modulus. The elliptic integral is trivial when $c=0$, for then it reduces to a linear function. As he had observed in his first paper, (1788a), it was already traditional in evaluating arc lengths along ellipses to make a substitution that transformed a given ellipse into another, more circular, one. Indeed, as Enneper later showed, Legendre's transformation is his re-working of Landen's transformation and is very close to some ideas of Lagrange, although neither are mentioned by name in Legendre's paper.[15]

In his paper Legendre considered the transformation $c'=\frac{2\sqrt{c}}{1+c}$. Iterating this transformation gives a series of values $c,c',c'',\ldots$ tending rapidly to the value 1. So iterating it backwards gives a series of values $c,c^o,c^{oo},\ldots$ tending towards the value 0. Explicitly,

$$c^o=\frac{1-\sqrt{1-c^2}}{1+\sqrt{1-c^2}},$$

or, in terms of the complementary modulus, $c^o,c^{oo},\ldots,c^{(n)},\ldots$ can be found from the sine tables by a quick use of the addition formulae.

Legendre applied this idea as follows. He defined ϕ^o by the equation

$$2\sin^2\phi=1+c^o\sin^2\phi^o-\Delta^o\cos\phi^o,$$

where $\Delta^o=\sqrt{1-(c^o)^2\sin^2\phi^o}$, and found that $F_c(\phi)=\frac{1+c^o}{2}F^o(\phi^o)$, where $F^o(\phi^o)=\int_0^{\phi^o}\frac{d\phi^o}{\Delta^o}$. He then deduced the useful little result that $\tan(\phi^o-\phi)=b\tan\phi$. So evaluating F at ϕ reduces to evaluating F^o at ϕ^o, which is easier to do because c^o is less than c, and by iterating he found that

$$F_c(\phi)=\frac{(1+c^o)(1+c^{oo})\ldots(1+c^{(n)})}{2^n}.F^{(n)}(\phi^{(n)}).$$

[15]See Enneper (1876, 353 and 358). Landen's transformation, introduced by him in his (1775), is equivalent to this transformation of the moduli: $k_1=\frac{1-k'}{1+k'}$.

When the value of $c^{(n)}$ was negligibly small, which Legendre observed it would be after a small number of steps, $\Delta = 1$ and $F(\phi) = \phi$, so he let $\Phi = \lim_{n\to\infty} \frac{\phi^{(n)}}{2^n}$ and found, on setting the product $(1+c^o)(1+c^{oo})\ldots = \alpha$ (it is a constant depending only on the choice of c, not the angle ϕ) that

$$F = (1+c^o)(1+c^{oo})\ldots = \alpha\Phi. \tag{1.18}$$

This meant that $F^1(c) = \frac{\pi}{2}\alpha$ and was easy to find from logarithm tables. When $\phi = \frac{\pi}{2}$ the limit $\Phi = \frac{\pi}{2}$, and the corresponding value of F is $\alpha\frac{\pi}{2}$.

Moreover, the value of c tended rather rapidly to zero, so that the convergence was quite rapid. Legendre illustrated this with an example. Starting from the value

$$c = \frac{\sqrt{2}}{2}\left(\frac{1+\sqrt{3}}{2}\right) = \sin 75^\circ,$$

which he noted was unfavourable to calculation because c was close to 1 (it is in fact 0.9659258262), Legendre showed that four iterations were enough to show that c^{oooo} vanished to seven decimal places, and $\log\alpha = 0.2460561$. He also showed that when c is small

$$c^o = \frac{1}{4}c^2 + \left(\frac{1.3}{2.4}\right)c^4 + \left(\frac{1.3.5}{2.4.6}\right)c^6 + \cdots \tag{1.19}$$

from which is followed that it was often enough to use just the first two terms. In the *Exercises* he gave a table of values of elliptic integrals to 14 decimal places.

Legendre also sought to show how useful his new functions would be in various parts of mathematics. In his (1792) he had mentioned the oscillations of a simple pendulum, which are given by the equation

$$dt = \frac{\sqrt{l}d\psi}{\sqrt{1-c^2\sin^2\psi}}, \tag{1.20}$$

where l is the length of the pendulum, $c^2 = \frac{h}{2l}$ where h is the height of the pendulum due to its speed at its lowest point (in units where the acceleration due to gravity =1) and the angle ψ is related to the angle of displacement from the vertical by the formula $\sin(\frac{\phi}{2}) = c\sin\psi$. So $t = F(\psi)$, which means that the familiar approximate equation for small arcs under-estimates the time needed to make a complete swing. Legendre was surprised to notice that he was the first to show that there were algebraic relations connecting the times of swings of a circular pendulum, just as there were for divisions of circular arcs.

In Vol. II of his *Exercises* he dwelt at length on three problems: the rotation of a solid about a fixed point; the motion either in plane or space of a body attracted to two fixed bodies; and the attraction due to an homogeneous ellipsoid. In the first volume of his *Traité* he added four more examples: motion under central forces, the surface area of oblique cones, the surface area of ellipsoids, and the problem of determining geodesics on an ellipsoid. The result was that Legendre's final book, surely intended by him as the definitive treatment of a topic that had engaged him all his life, was a systematic presentation of new functions of a real variable that embraced their definition, their fundamental properties, and tabulated their values, while displaying their utility in solving significant problems in applied mathematics.

At the time of the publication of the *Exercises* Legendre seems to have felt his work was not getting the attention it deserved. Once Abel and Jacobi had taken it as their point of departure, its fame was assured. In the first supplement to his *Traité*, dated 12 August 1828, Legendre wrote that[16]

> Until then geometers had taken almost no part in this kind of research, but scarcely had my work seen the light of day, scarcely could its title have become known to scientists abroad, when I learned with as much astonishment as satisfaction that two young geometers, MM. Jacobi of Königsberg and Abel of Christiania had succeeded in their own individual work in considerably improving the theory of elliptic functions at its highest points.

And Dirichlet, in his memorial address on Jacobi (see Dirichlet 1852, 9), said that

> It is Legendre's eternal glory to have discovered the kernel of an important branch of analysis and by the work of half a lifetime to have erected on these foundations an independent theory .... Only with the continued determination that enabled the great mathematician to return again and again to the subject was he able to overcome difficulties that, with the means he had at his command, must have seemed scarcely subduable.

Legendre's is firmly a real theory of functions of a real variable, and the obvious analogy Euler had pointed out between $\int \frac{dt}{\sqrt{1-t^2}}$ and $\int \frac{dt}{\sqrt{1-t^4}}$ never fruitfully suggested to Legendre that he should invert his new-found functions. Inversion in this context is to take an integral, say $y(x) = \int_0^x f(t)dt$, and to regard it not as a function, y, of its upper end point, x, but inversely, as defining a function $x = x(y)$. The familiar motivation for doing this is the trigonometric functions, as Jacobi spelled out in a paper of 1832, after the first rush of discovery was over.[17] Inverting the integral $y(x) = \int_0^x \frac{dt}{\sqrt{1-t^2}}$ gives the function $x = x(y) = \sin y$. This is a much more tractable function than $y(x) = \arcsin(x)$; it is periodic, whereas the arcsin function is infinitely many-valued (and for that reason not strictly a function at all in the modern sense of the term).

As Krazer pointed out, Legendre regarded the inversion problem as essentially solved by his tables, which in a sense it is—for real functions of a real variable (see Krazer 1909, 55). In view of the great successes of Abel and Jacobi, it is interesting to speculate on why Legendre never did in 40 years what they did almost at once, and there is surely much justice in Jacobi's opinion that[18]

[16] Quoted in Ore (1957, 210).

[17] See Jacobi (1832b, 5). Partial English translation in Birkhoff (1973, 207–212).

[18] Quoted in Koenigsberger (1904b, 54) and Krazer (1909, 55n).

> The fear that the mathematician most chiefly concerned with the determination of numerical values had of the imaginary, was the reason that Legendre was prevented from taking the most important step in modern analysis, the introduction of doubly periodic functions.

We would add that in many of the applied problems that Legendre discussed the elliptic integral arises naturally, and the problem and its solutions make sense at that level. Moreover, in Legendre's time the distinction between a function and a multi-function was not sharply made and indeed, would one throw the logarithm out of mathematics in favour of its inverse the exponential function on the grounds that it is not a proper function?

The crucial breakthrough in the study of elliptic integrals proved to be the simultaneous recognition of inversion and complexification as governing ideas. As we shall see, the integral $\int_0^v \frac{dt}{\sqrt{1-t^4}} = u(v)$ is more tractable when regarded as defining not $u = u(v)$ but, inversely, $v = v(u)$, and it is seen even more clearly when v is allowed to be complex. Without this second step, the first one loses much of its force. Although Legendre knew that his tables of values for elliptic integrals solved the inversion problem, he did not appreciate the fact. Abel, Jacobi, and Gauss almost immediately took up the question as one involving complex u and v, and for them inversion burst with the full force of a revelation.

Another stimulus to the work of Abel was the cryptic hint that Gauss had dropped in his *Disquisitiones arithmeticae* (1801, Sect. 335) about the principles of his theory of cyclotomy: "Not only can they be applied to the theory of circular functions, but also to many other transcendental functions, e.g. those which depend on the integral $\int \frac{dt}{\sqrt{1-t^4}}$". But although Gauss promised a treatment of this and related topics, he never provided it, and it was the young Abel who, with difficulty, made it yield its secrets.

1.2 Abel

Niels Henrik Abel was born on the 5th of August 1802 on the island of Finnöy, off the southern coast of Norway, where his father was a pastor and his mother the daughter of a local shipowner, and educated at home until he was 13, when he was sent to school in Christiania (now Oslo). In his final year there his extraordinary mathematical talents were discovered when a new mathematics teacher, Bernt Michael Holmboe, himself only 22, came to the school.[19] Soon Holmboe was giving Abel, not yet 17, private tuition in the works of Lagrange; not surprisingly he wrote in Abel's school report that the youth was "an excellent mathematical genius". Abel

[19] Abel's short life has been well told in many places, for example Ore (1957) and most recently Stubhaug (2000). Stubhaug's fascinating account is much fuller on Abel's life and times, Ore's remains more reliable on the mathematics. The richest account of his mathematics is Houzel (2004).

Fig. 1.2 Niels Henrik Abel (Portrait by Gorbitz, in Stubhaug, *Niels Henrik Abel and his times*, Springer ©1996, p. 437)

graduated from school and went to Christiania university, where he was encouraged by Hansteen, an expert in terrestrial magnetism, to think of a career in mathematics (Fig. 1.2). When in 1823 Abel thought, erroneously, to have shown how to solve the quintic by radicals, Hansteen sent him to Copenhagen to talk to Degen, who was better qualified than anyone in Norway. Degen could not find the flaw, but his advice, to test the method on examples, soon revealed that there was one. Writing to Holmboe about the trip, Abel commented for the first time on record about elliptic integrals and on Degen's reaction to some of Abel's ideas on a work he showed him[20]:

> You remember the little paper which treated the inverse functions of the elliptic transcendentals, where I proved something impossible; I asked him to read it from one end to the other; but he could not discover any false conclusion nor understand where the fault might be; God knows how I will get out of it.

Krazer (1909, 56), following many other commentators, observed that the "mistake" may only have been a surprising novelty, for example that the division equation is of order n^2 and not order n. Be that as it may, Abel's letter reveals that he had already taken one of the two crucial steps in the transformation of Legendre's work and inverted the integrals.

In 1825 Abel was funded by the university to continue his studies abroad, and he left with three friends on September 7th. They first went to Berlin, where Abel was fortunate enough to meet Privy Councillor Leopold Crelle. Forty-five years old, Crelle was an influential engineer who had built the railway from Berlin to nearby

[20]Quoted in Ore (1957, 65).

Potsdam, but he was also an exponent of the German philosophy of neo-humanism. Thus he laid great emphasis on the study of mathematics for its own sake, believing that worthwhile applications of new mathematics would arise naturally, so that the subject need not be shackled, as had previously been the case in many parts of Germany, to a narrowly utilitarian view. His success as an engineer, allied to his skills as an organiser, made him a very powerful supporter of pure mathematics. Even more valuably for Abel he was in the process of establishing a new journal for mathematics, the first of its kind in Germany. Crelle hoped thereby to raise what he saw as the poor standards of German mathematics, and he invited Abel to contribute to it. This journal, the *Journal für die reine und angewandte Mathematik*, rapidly became the leading journal devoted to mathematics in Germany, and after Crelle's death ever more securely the journal of mathematics as Berlin saw it. Despite its title it concentrated on pure mathematics, becoming what wits called the *Journal für die reine unangewandte Mathematik*, the Journal for pure unapplied mathematics.

In Abel, Crelle gained a prolific contributor who brought fame to his journal. Through Crelle, Abel gained access to an up-to-date mathematical library. He also gained a lifelong admirer, and a translator of the papers which he began to produce at a great rate—no less than seven in the first volume of Crelle's *Journal* alone. His interest in the solvability of equations by radicals deepened. Before leaving Norway he had succeeded in showing, independently of Ruffini (1799), that the general polynomial equation of degree 5 is not solvable by radicals. Now he began to consider those of higher degree which nonetheless are solvable by radicals and to investigate when this occurred. He also became aware of Cauchy's critiques of divergent series in analysis and wrote his own paper on the convergence of the binomial series. The divergence of certain Fourier series struck him most forcefully.

He decided against visiting Gauss in his lair in Göttingen, perhaps on the advice of Crelle. Crelle, like everyone in Berlin, thought highly of Gauss, but regretted that everything the great man wrote was "an abomination—gruel—so obscure that it is almost impossible to understand anything" and Hansteen added that Gauss "is like the fox who covers his tracks in the snow with his tail".[21] Abel wrote with evident surprise to Hansteen on 5 December 1825 (in Abel 1902, 11–12) that

> The degree to which young mathematicians here in Berlin and, I hear, all over Germany almost worship Gauss is extraordinary. For them he is the epitome of mathematical perfection, but if he is indeed a great genius it is also certain that he writes badly.

Instead Abel went first on a holiday and then to the centre of the mathematical world, Paris, where he arrived on July 10, 1826. But he did not find the French mathematicians as approachable as the Prussians had been, indeed his best contact was with the young Lejeune Dirichlet who called on Abel under the mistaken impression that he was a fellow German. He did meet Cauchy but found him mad and impossible to deal with, as he wrote to Holmboe (24 October) adding that "He is

[21] The quotes from Crelle and Hansteen are given in Bjerknes (1885, 92).

extremely Catholic and bigoted, which is strange for a mathematician". On the other hand he was the only person actively engaged in pure mathematics, everyone else being engaged in magnetism and other kinds of physical theories, while Legendre, then 74, was "as old as stone".[22]

In December Abel took up the lemniscate, whose arc length is given by the simplest of elliptic integrals: $s = \int_0^x \frac{dt}{\sqrt{1-t^4}}$. He wrote to Crelle that he had found that the lemniscatic arc can be divided by ruler and compass into n equal arcs for exactly the same values of n as the circular arc, as Gauss had hinted in the *Disquisitiones arithmeticae*. To Holmboe he said that he now saw as clear as day how Gauss had come to discover his results.[23]

In January 1827 he was back in Berlin, which he preferred to Paris. It was in this year that his first major paper on elliptic integrals appeared in Crelle's *Journal*, and it will be worthwhile considering it in some detail, for it presents the first account of the theory of elliptic functions. The paper, his *Recherches sur les fonctions elliptiques* (Abel, 1828a), amply displays Abel's lucid style of exposition, fostered under the firm guidance of Crelle. It also indicates the benefits that flowed from taking the second step and allowing the variables to have complex values. In it Abel developed the theory of elliptic functions in close analogy with the theory of trigonometric functions. He introduced three functions (replacing, as it were, sine and cosine) and derived addition laws for them and investigated their division formulae. There he presented a thorough analysis of when the corresponding equations are solvable by radicals and proved Gauss's claims about the lemniscate. He concluded with finding expressions for the new functions as infinite series and infinite products.

Abel began his paper by summarising what Legendre had shown in his *Exercises*. Abel took the general elliptic integral of the first kind in a slightly different form:

$$\alpha = \int_0^x \frac{dt}{\sqrt{(1-c^2t^2)(1+e^2t^2)}}. \tag{1.21}$$

The inverse function $x = \phi(\alpha)$ therefore satisfies

$$\frac{d\phi}{d\alpha} = \sqrt{(1-c^2x^2)(1+e^2x^2)}.$$

Abel introduced the useful abbreviations

$$f(\alpha) = \sqrt{1-c^2\phi^2(\alpha)}, \quad F(\alpha) = \sqrt{1+e^2\phi^2(\alpha)}. \tag{1.22}$$

[22] All quotations from Abel (1902, 45–46); see also Ore (1957, 147) and Stubhaug (2000, 405).

[23] See Abel (1881, 2, 261), and Ore (1957, 154).

He then remarked that these functions have some interesting properties, for example, each of the equations $\phi(\alpha)=0$, $f(\alpha)=0$, $F(\alpha)=0$ has infinitely many roots which can be determined, and for any integer m, $\phi(m\alpha)$ can easily be found in terms of $\phi(\alpha)$, etc. But the converse problem is harder: to express $\phi(\alpha)$, $f(\alpha)$, or $F(\alpha)$ in terms of $\phi(m\alpha)$ requires solving an equation of degree m^2, the study of which is the principal point of the memoir. The analogy with the trigonometric functions is evident.

He defined

$$\frac{\omega}{2}=\int_0^{1/c}\frac{dt}{\sqrt{(1-c^2t^2)(1+e^2t^2)}} \tag{1.23}$$

and observed that ϕ is positive in the range $0<\alpha<\frac{\omega}{2}$, and that $\phi(0)=0$ and $\phi\left(\frac{\omega}{2}\right)=\frac{1}{c}$. Moreover, ϕ is an odd function, $\phi(-\alpha)=-\phi(\alpha)$. If α is replaced by $i\beta$, so $ix=\phi(i\beta)$ then $\beta=\int_0^x\frac{dt}{\sqrt{(1+c^2t^2)(1-e^2t^2)}}$ is real and positive between $x=0$ and $x=1/e$.[24] So inversely, defining

$$\frac{\tilde{\omega}}{2}=\int_0^{1/e}\frac{dt}{\sqrt{(1+c^2t^2)(1-e^2t^2)}}, \tag{1.24}$$

x is positive in the range $0<\beta<\frac{\tilde{\omega}}{2}$. It is clear that $\phi(i\alpha)=i\phi(\alpha)$, and $F(i\alpha)=f(\alpha)$, and also that $\phi\left(\frac{i\varpi}{2}\right)=\frac{i}{e}$. Now that many values of the functions ϕ, f, F are known along the real and imaginary axes, it remained to find their general values. Abel showed that

$$\phi(\alpha+\beta)=\frac{\phi(\alpha)f(\beta)F(\beta)+\phi(\beta)f(\alpha)F(\alpha)}{R},$$

$$f(\alpha+\beta)=\frac{f(\alpha)f(\beta)-c^2\phi(\alpha)\phi(\beta)F(\alpha)F(\beta)}{R},$$

$$F(\alpha+\beta)=\frac{F(\alpha)F(\beta)+e^2\phi(\alpha)\phi(\beta)f(\alpha)f(\beta)}{R},$$

[24] For the sake of convenience here, and everywhere in this book, we write i instead of $\sqrt{-1}$ even though $\sqrt{-1}$ was commonly used by mathematicians until the late 1840s. Abel himself wrote "where i to abbreviate, represents the imaginary quantity $\sqrt{-1}$". In this he followed Gauss's usage in the *Disquisitiones arithmeticae* of 1801. The first to use i to denote the imaginary unit was Euler in his (1777d), presented to the St. Petersburg Academy on May 5, 1777 but published only in 1794 as Supplement IV to Vol. 1, Chap. V of his *Institutiones calculi integralis*. There (p. 184) he stated: "In the following I will designate the formula $\sqrt{-1}$ by the letter i such that it will be $ii=-1$ and $-1/i=i$ as well." It is likely that Gauss was prompted by this to make a systematic use of i.

where $R = 1 + c^2e^2\phi^2(\alpha)\phi^2(\beta)$, observing that these results followed from those in Legendre's *Exercises*, but also that they could be proved directly, as follows. Letting r denote the right-hand side of the first equation, he calculated $\frac{dr}{d\alpha}$ and observed that this expression was symmetric in α and β so $\frac{dr}{d\alpha} = \frac{dr}{d\beta}$ whence r is a function of $\alpha + \beta$ which can be found by setting $\beta = 0$, when $\phi(\alpha)$ is obtained, so $r(\alpha+\beta) = \phi(\alpha+\beta)$.

As Abel then said, a slew of other formulae then follow from these addition formulae, formulae for $\phi(\alpha+\beta)+\phi(\alpha-\beta)$, for $f(\alpha+\beta) + f(\alpha-\beta)$, for $\phi(\alpha \pm \frac{\omega}{2})$, and so forth. From

$$\phi(\alpha+\tfrac{\omega}{2}) = \phi(-\alpha+\tfrac{-\omega}{2}) \text{ and } \phi(\alpha+\tfrac{i\tilde{\omega}}{2}) = \phi(-\alpha+\tfrac{i\tilde{\omega}}{2}),$$

Abel deduced (in Sect. 4) that, as he put it,

$$\phi\left(\frac{\omega}{2}+\frac{i\tilde{\omega}}{2}\right) = f\left(\frac{\omega}{2}+\frac{i\tilde{\omega}}{2}\right) = F\left(\frac{\omega}{2}+\frac{i\tilde{\omega}}{2}\right) = \frac{1}{0}. \tag{1.25}$$

In due course this became the recognition that these functions have poles at $\frac{\omega}{2} + \frac{i\tilde{\omega}}{2}$. The formulae at once imply the functions are periodic:

$$\phi(\alpha+\omega) = -\phi(\alpha),\ \ \phi(\alpha+2\omega) = \phi(\alpha), \tag{1.26}$$

$$\phi(\alpha+i\tilde{\omega}) = -\phi(\alpha),\ \ \phi(\alpha+2i\tilde{\omega}) = \phi(\alpha), \tag{1.27}$$

with similar formulae for f and F, and Abel observed (Sect. 5)

> The formulae that have been established make clear that one will have all the values of the functions $\phi(\alpha)$, $f(\alpha)$, $F(\alpha)$ for all real or imaginary values of the variable if one knows them for the real values of this quantity lying between $\frac{\omega}{2}$ and $\frac{-\omega}{2}$ and for the imaginary values of the form $i\beta$, where β lies between $\frac{-\tilde{\omega}}{2}$ and $\frac{\tilde{\omega}}{2}$.

One says that the functions are doubly periodic, because they have two distinct periods, 2ω and $2i\tilde{\omega}$. To find the zeros of $\phi(\alpha+i\beta)$, therefore, it is enough to look in the restricted range $\frac{-\omega}{2} \le \alpha < \frac{\omega}{2}$, $\frac{-\tilde{\omega}}{2} \le \beta < \frac{\tilde{\omega}}{2}$. By the addition formula, the equation $\phi(\alpha+i\beta) = 0$ reduces to two equations on separating out real and imaginary parts:

$$\phi(\alpha)f(i\beta)F(i\beta) = 0, \quad \phi(i\beta)f(\alpha)F(\alpha) = 0.$$

These equations in turn yield $\alpha = m\omega$, $\beta = n\tilde{\omega}$ (or, if you pursue $f(i\beta)F(i\beta) = 0$, $\alpha = \left(m+\frac{1}{2}\right)\omega$, $\beta = \left(n+\frac{1}{2}\right)\tilde{\omega}$, values for which ϕ is infinite). So the zeros of $\phi(x)$ are $x = m\omega + ni\tilde{\omega}$. A similar calculation enabled Abel to solve the equation $\phi(x) = \frac{1}{0}$ and find $x = \left(m+\frac{1}{2}\right)\omega + \left(n+\frac{1}{2}\right)i\tilde{\omega}$, and to solve the equation $\phi(x) = \phi(a)$ in terms of a and find $x = (-1)^{m+n}a + m\omega + ni\tilde{\omega}$. Although Abel did not use the phrase, this amounts to saying that the values of an elliptic function are known everywhere once they are known on a period parallelogram.

Nothing so far can be called difficult, indeed it was all made to look rather easy. Nor was it much more work for Abel to express $\phi(m\alpha)$ as a rational function of

$\phi(\alpha)$, $f(\alpha)$, and $F(\alpha)$, the analogy with the trigonometric functions was still a good guide. Abel set $\phi(n\alpha) = \frac{P_n}{Q_n}$ and found recurrence relations for P_n and Q_n, and dealt similarly with $f(n\alpha)$, and $F(n\alpha)$. It was with the converse question—express, say, $\phi(\alpha/m)$ as a function of $\phi(\alpha)$, $f(\alpha)$, and $F(\alpha)$—that the research became hard, and it is in this connection that Abel did some of his finest work on the theory of equations, which was to have considerable implications for the future development of Galois theory.

To see why the solutions are so numerous leads us directly to the importance of introducing complex variables. Recall that to find $\sin(u/3)$ given $\sin u$ requires solving a cubic equation, for, from the equation $\sin u = 3\sin(u/3) - 4\sin^3(u/3)$ we deduce that $\sin(u/3)$ satisfies the cubic equation $4x^3 - 3x + \sin u = 0$. We also know that the angles $(u/3), (u+2\pi)/3$ and $(u+4\pi)/3$ are such that their sines satisfy this equation. However, in the elliptic case, the division equation that gives $\phi(u/3)$ in terms of $\phi(u)$ is of degree 9. The key realisation is that in the elliptic case some of the roots are complex, whereas as in the trigonometric case they are all real. Once the elliptic function is treated as a function of a complex variable, the values of u that satisfy the equation for $\phi(u/3)$ will all clearly be of the form $u_0 + (m\omega + \tilde{m}i\tilde{\omega})/3$, and so there are 9 of them. This makes it clear why the number of solutions is unexpectedly large.[25]

Abel next picked up the hint dropped by Gauss about the lemniscatic functions and investigated the equations that arise to see when, for example, they are solvable by radicals. When $m = 2$ it was easy for Abel to show (Sect. 13) that

$$\phi\left(\frac{\alpha}{2}\right) = \frac{1}{c}\sqrt{\frac{1-f(\alpha)}{1+F(\alpha)}}$$

with similar formulae for $f\left(\frac{\alpha}{2}\right)$ and $F\left(\frac{a}{2}\right)$. It follows that if the values $\phi(\alpha)$, $f(\alpha)$, and $F(\alpha)$ are constructible, so are the values for $\phi\left(\frac{\alpha}{2}\right)$ and $\phi\left(\frac{\alpha}{2^n}\right)$, and similarly the f's and F's. Repeated bisection followed by iteration.

Abel then turned to the study of the equations

$$\varphi(\alpha) = \frac{P_p}{Q_p},$$

and the corresponding equations for the functions f and F—the determination of the functions $\varphi\left(\frac{\alpha}{p}\right)$, $f\left(\frac{\alpha}{p}\right)$, and $F\left(\frac{\alpha}{p}\right)$ as a functions of $\varphi(\alpha)$, $f(\alpha)$, and $F(\alpha)$ when $p = 2n+1$ is an odd prime, as he put it in Sect. 12. These equations are all of degree p^2, and Abel showed that they are all solvable algebraically (that is, by radicals). The structure of his argument reveals how deeply he had penetrated the theory of

[25]The actual number is n^2 if n is odd and $2n^2$ if n is even. The solutions are $\phi\left(\frac{u}{n} + 2\frac{m\omega+\mu\varpi i}{n}\right)$ when n is odd and $\phi\left((-1)^{m+\mu}\frac{u}{n} + 2\frac{m\omega+\mu\varpi i}{n}\right)$ when n is even, $0 \le m, \mu < n$. See Houzel (1978, 24).

the solution of equations; only Gauss had gone as far, and he had for once published his ideas, in the *Disquisitiones arithmeticae*. There Gauss had made a deep analysis of the cyclotomic equation, which has the property that all of its complex roots are known once a primitive root is found. In the present setting also, once one root is known so are all the others—if $\phi\left(\frac{\alpha}{p}\right)$ is a root of the equation $\phi(\alpha) = c$, then so are the $p^2 - 1$ quantities $\phi\left(\alpha + \frac{m\omega+\tilde{m}\tilde{\omega}}{p}\right)$. So in principle a great deal is known about the polynomial equation $P_p = \phi(\alpha)Q_p$.

After a lengthy argument, which we omit (Sects. 14–16), Abel concluded that $\phi(\alpha/p)$ was an algebraic function of $\phi(\alpha)$ and added that the same was true for $f\left(\frac{\alpha}{p}\right)$ as a function of $f(\alpha)$ and for $F\left(\frac{\alpha}{p}\right)$ as a function of $F(\alpha)$. Moreover the values $\phi(\alpha/p)$ could be written in terms of radicals involving the values of $\phi(\alpha)$ and the values of φ at the points $\frac{2\omega}{p}$ and $\frac{2\varpi}{p}$, so the equation was even solvable by radicals once those numbers have been adjoined. They are found by solving the equation $P_p(\alpha) = 0$, whose roots are $x = \phi\left(\frac{m\omega+\tilde{m}i\tilde{\omega}}{p}\right), \quad -n \le m, \tilde{m} \le n$. But $m = \tilde{m} = 0$ gives the uninteresting solution $x = 0$, so one has $P_{2n+1} = xR$, and $R = 0$ is an equation of degree $p^2 - 1 = (p+1)(p-1)$. Abel found it helpful to study this equation after making the substitution $x^2 = r$, when it is of degree $(2n+2)n$ in r, and he showed that its solution reduced to two equations, one of degree $2n+2$ and the other of degree n. Indeed by considering their coefficients, he showed that the roots of $R = 0$ could be found by solving one equation of degree $2n+2$ and then $2n+2$ equations of degree n. Moreover, these equations of degree n were solvable algebraically, while the equation of degree $2n+2$ was not generally so, although it could be in special cases. In this context he was happy to state as a particular case of his findings (Sect. 22):

> One can divide the entire circumference of the lemniscate into m equal parts *by ruler and compass alone* if m is of the form 2^n or a prime of the form $2^n + 1$, or if m is a product of numbers of these two kinds. This theorem, as one sees, is exactly the same as the theorem of Gauss for the circle. (Italics Abel's.)

Abel concluded the first part of his memoir with expressions for $\phi(\alpha)$, $f(\alpha)$ and $F(\alpha)$ as infinite series and infinite products. Heuristically he derived them from expressions for $\phi\left(\frac{\alpha}{2n+1}\right)$ by letting n become infinite, when, as he put it, $\phi\left(\frac{\alpha}{2n+1}\right)$ disappears and one obtains for $\phi(\alpha)$ an algebraic expression with infinitely many terms. This kind of argument vividly recalls Euler's treatment of the trigonometric functions in his *Introductio in analysin infinitorum*, I, (1748a). However, Abel attempted a more rigorous limiting argument, but in the absence of any concept of uniform convergence it could not be said to be successful. Houzel (1978, 38) remarks at this point that Abel's treatment of the convergence questions is delicate. It is certainly very thorough, but we do not think it is convincing. The upshot is a rather intimidating expression for $\phi(\alpha)$, $f(\alpha)$, and $F(\alpha)$ that leave no doubt of Abel's immense technical facility with the traditional style of analysing functions.

Abel's ideas about complex variables were entirely formal, but that said he presented a fully developed theory of elliptic functions as complex functions, analogous to the trigonometric functions. We shall now see that Jacobi's way in to the subject was significantly different, and closer to Legendre's.

1.3 Jacobi

Jacobi was born into a banker's family in Potsdam near Berlin on 10 December 1804 and displayed his prodigious gifts for mathematics and languages from an early age (Fig 1.3). He entered the local gymnasium before he was 12 and within 6 months rose to the top class, where he had to remain for 4 years because he was not allowed to enter the university until he was 16. It was at school that he first read Euler. When he went to the University of Berlin he had to choose between mathematics and philology, then enjoying its golden age in Germany. He chose mathematics, but not the mostly elementary lectures then on offer, rather the heady diet of Euler, Lagrange and Laplace. By 1825 he had submitted his doctoral thesis and was given permission to begin his Habilitation, the qualification that gave anyone who possessed it the right to teach in any German university (although not necessarily to be paid). He also became a Christian at this time, a high point of the assimilationist trend among Germans of Jewish descent, which enabled him to become a Privatdozent or instructor at Berlin University. He was not yet 21. It is worth pointing out that in 1801, before the University of Berlin was founded, the city was little more than a garrison town, its middle class largely Jewish but subject to many legal restrictions and extra taxes. Foreign Jews had to enter the city by the Rosenthaler gate, which was otherwise used for cattle.[26]

In May 1826 Jacobi transferred to the University at Königsberg, where he joined a small but strong science faculty consisting of Franz Neumann and Heinrich Dove in physics and the astronomer Friedrich Bessel. He became an associate professor there at the end of 1827. Inspired by his colleagues, his first publications are in applied mathematics, and he remained interested in the subject all his life. Ironically, success came easily to him with almost every thing he touched except the topic that made his name internationally and which interests us most: elliptic functions. His friend in later life, Dirichlet, tells this story in his memorial address of Jacobi[27]:

> One of his friends who noticed him in a bad mood one day, received this answer when he asked why he was out of sorts: You see me on the point of returning this book (Legendre's *Exercises*) to the library, with which I've been exceedingly unlucky. On other occasions when I have studied an important work this has stimulated me to some thoughts of my own and there has always been something in it for me. This time I have come away quite empty handed and have not been inspired in the least.

[26] Information taken from Elon (1999, 20). See also Elon (2002).

[27] Dirichlet (1852, 7).

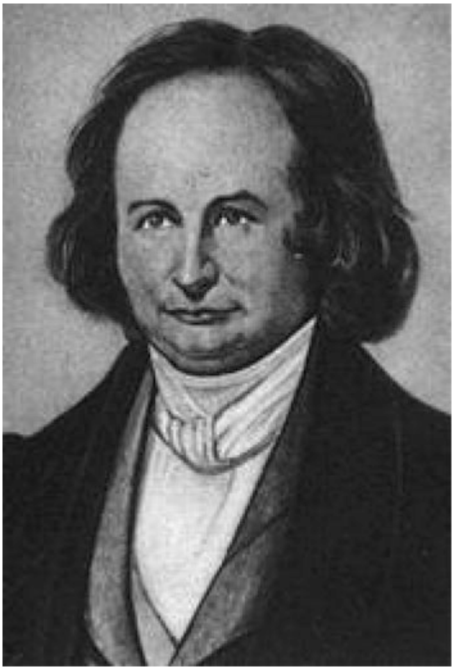

Fig. 1.3 Carl Gustav Jacob Jacobi (From his *Gesammelte Werke*)

Later, but richer, was Dirichlet's comment,

By June 1827 Jacobi had indeed come to some new ideas of his own. Specifically, he had found new ways to transform one elliptic integral into another by rational changes of the variable. He was the first to discover the existence of transformations of every degree. Not having seen Legendre's *Traité* he did not know that Legendre had found a transformation of order 3. He sent a report of his discoveries to Schumacher for publication in his *Astronomische Nachrichten*, where they appeared in the September issue.

In his two notes Jacobi considered the transformation of elliptic functions, which he approached via the equation

$$\int_0^{\phi} \frac{dt}{\sqrt{1-c^2\sin^2 t}} = m\int_0^{\psi} \frac{dt}{\sqrt{1-k^2\sin^2 t}}\,. \tag{1.28}$$

Given an elliptic integral in terms of ϕ, it is required to transform it into another involving ψ. Each such transformation determines the values of the new modulus, k, and the number m. Jacobi observed that one can find such transformations by writing $\sin\phi$ as a rational function of $\sin\psi$ thus: $\sin\phi = \frac{U}{V}$, where U is a polynomial in odd powers of $\sin\psi$ up to the mth, and V is a polynomial in even powers of $\sin\psi$ up to the $(m-1)$th. Such a transformation is said to be of order m. However, at this stage Jacobi only knew how to find the polynomials U and V when $n = 3$ or 5; the rest of his claim was strictly speaking only a conjecture. He also pointed out that if one takes two or more transformations one can obtain solutions to the multiplication problem of order a composite number n relating two or more elliptic integrals:

$$\int_0^{\phi} \frac{dt}{\sqrt{1-c^2\sin^2 t}} = n\int_0^{\psi} \frac{dt}{\sqrt{1-c^2\sin^2 t}}\,.$$

In his second paper he showed how his new transformations, like Legendre's, could assist in calculating tables of values for the integrals.

These papers are rather small by comparison with Abel's and very much in Legendre's spirit. However, in the December edition of Schumacher's journal he broke new ground by proving his general transformation formula and inverting the elliptic integral. Here he argued that the substitution $y = \frac{U}{V}$ will transform

$$\frac{dy}{\sqrt{(1-a_0y)(1-a_1y)(1-a_2y)(1-a_3y)}} = \frac{dy}{\sqrt{f_4(y)}}$$

into

$$\frac{dy}{M\sqrt{(1-b_0y)(1-b_1y)(1-b_2y)(1-b_3y)}} = \frac{dy}{\sqrt{g_4(y)}},$$

where, if

$$T^2 = V^4\frac{f_4(\frac{U}{V})}{g_4(x)}\,,\text{ then } \frac{T}{M} = V\frac{dU}{dx} - U\frac{dV}{dx}$$

and M is to be a constant.

So to transform $\frac{dy}{\sqrt{(1-y^2)(1-\lambda^2y^2)}}$ into $\frac{dx}{M\sqrt{(1-x^2)(1-k^2x^2)}}$ he denoted the integral $\int_0^{\phi} \frac{dt}{\sqrt{1-k^2\sin^2 t}}$ by $F(\phi)$ and defined

$$K = \int_0^{\pi/2} \frac{dt}{\sqrt{(1-k^2\sin^2 t)}}.$$

Observing that if $F(\phi)+F(\psi) = F(\sigma)$, which is the expression given by Legendre (see (1.12) above) and writing $F(\phi) = \xi$, Jacobi said

$$F(\phi) = \int_0^{x} \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}}$$

and $x =$sinam ξ (Fig. 1.4).

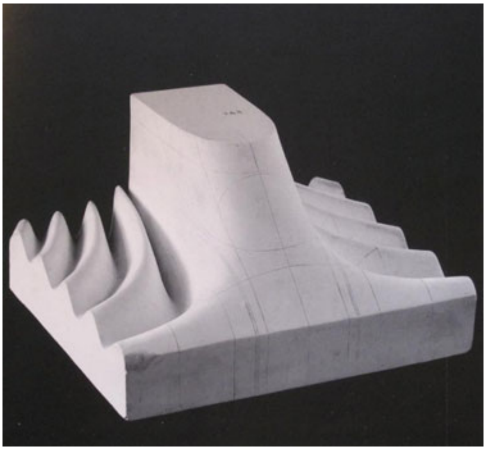

Fig. 1.4 The real part of Jacobi's amplitude function (Fischer, *Mathematische Modelle – Mathematical Models*, Vieweg ©1986, number 132)

This is, of course, the crucial step of inverting the elliptic integral.[28] Formulae like the one for σ in terms of ϕ and ψ can be reinterpreted so as to apply to $x = \mathrm{sinam}\xi$, and pursuing this insight led Jacobi to his formulae first for V in terms of x then U in terms of x and finally for M. The coefficients of U and V and the expression for M are full of terms like $\mathrm{sinam}\frac{2K}{2n+1}$ and $\mathrm{sinam}\frac{2nK}{2n+1}$, so it is quite evident that without inversion Jacobi could never have proved his general claim. When, then, did he discover it?

In this connection the famous correspondence between Legendre and Jacobi is most informative. It was begun when Jacobi wrote to Legendre on 5th August 1827 to inform him of the discoveries which Schumacher was shortly to publish. The letter is rather odd in tone: Jacobi was trying to be modest—which never came easily to him, he was a notoriously irascible person—yet he knew he had something to say. He described his work on elliptic integrals, mentioned in passing that Gauss had discovered a transformation of order 7, and ended with some remarks about cubic and biquadratic reciprocity in number theory (see Sect. 4.7.2).

Legendre replied on 30 November 1827, claiming that Jacobi's letter had been delayed and only just reached him. He welcomed the transformation of order 5 although he recognised the one of order 3 as his own. However, he could not find polynomials U and V for the 7th order transformation, and rather doubted that they could exist. Jacobi had not published a proof, and he urged him to do so. But the remarks about Gauss could not pass unnoticed: without proof, Gauss had claimed the law of quadratic reciprocity as his when he, Legendre, had published it in 1785, and Gauss had then gone on to try and poach the method of least squares. So Jacobi should not worry that Gauss claimed to have discovered Jacobi's ideas in 1808, such excessive impudence was unbelievable in a man of Gauss's abilities.

[28] Jacobi wrote the new function as a composite: the sine of the amplitude of ξ. As the function became more widely used its name contracted to *sinam* and eventually to *sn*.

Historians observe that Gauss only claimed, correctly, to be the first to prove the law of quadratic reciprocity, attributing its discovery to Legendre, and priority of discovery for the method of least squares is indeed due to Gauss, although Legendre published first. It must have been difficult to endure the lofty utterances coming from Göttingen, and Legendre had had more than his share. Legendre did more than write privately to Jacobi, he communicated his results to the Académie in the warmest terms on 5 November 1827, and his report was published in the *Globe* on the 29th.

Jacobi was delighted with the much older man's response and took the opportunity of his reply on 12 January 1828 to acquaint Legendre with the content of Abel's *Recherches*, the first part of which was now in print. He put Abel's work into his own notation, but otherwise summarised it much as we have above, dwelling on the striking division of the lemniscate. He then pointed out that the new edition of Schumacher's journal carried a proof of the existence of the general transformation and started to discuss modular equations (which we will define below, see p. 47). In Legendre's reply, 9 February 1828, Legendre explained that he had already seen Abel's paper, although he was happy to have had it analysed in a language closer to his own. In what has become a famous passage he went on[29]:

> It is a great satisfaction to me to see two young mathematicians, like you and him, cultivate successfully a branch of analysis which has for so long been my favourite object of study and which has not attracted the attention in my country that it deserves.

In his reply to this letter (12 April 1828) Jacobi retraced his route to his original discovery. This makes it clear that he had had the idea of inverting the integrals before he wrote his second letter to Schumacher, in August 1827. This seemingly establishes his independence from Abel, by just over a month. But the case is not convincing, although Krazer accepted it (see Krazer 1909, 59). Jacobi first admitted that his published proofs were different from his discovery method and then made his first study of the transformations hinge on the substitution $\operatorname{sinam}\frac{2K}{3}$. Other views are possible. Bjerknes in his biography of Abel (Bjerknes 1885, 83–85), was the first to raise the idea that Abel's work might have provided the necessary inspiration for Jacobi, which he later covered up. This was not accepted by Sophus Lie, who was working on the second edition of Abel's collected works at the time, and was the cause of some friction between the two (see Stubhaug 2002, 284). But Ore argues convincingly (1957, 180–190) that Jacobi could well have seen Abel's *Recherches* at a time when he was stuck for a general proof of his first important insight. His application was in for a professorship at Königsberg but not yet decided, Legendre was putting him under pressure to provide a proof, and suddenly another man opens up the whole subject. How natural to use his ideas, and how smoothly, we add, does Jacobi flow from describing Abel's work to solving his own problem. Jacobi is usually generous to Abel in letters but avoids mentioning his name in print. Ore is properly cautious and we incline to agree with him that the idea of inversion "in

[29] See Legendre and Jacobi (1875, 224).

all probability was revealed to him through reading Abel's *Recherches*" (Ore, 1957, 184). What Jacobi did with the idea in any case exonerates him from what small weakness not acknowledging published work may comprise.

Now that both men were in print, it soon became evident that an exciting development in mathematics was taking place. Abel, who must have thought he had this subject to himself, was at first quite alarmed. We have this account of his reaction in a letter from Hansteen to Schumacher that Schumacher quoted to Gauss[30]:

> Abel sends herewith an article about elliptic transcendents, which he asked to have printed as soon as possible, since Jacobi is on his heels. The other day, when I handed him the last number of the *Astronomische Nachrichten*, he became quite pale and was compelled to run to the confectioner's shop and take a schnapps of bitter to counteract his alteration. For several years he has been in possession of a general method which he communicates in this paper, and which includes more than Jacobi's theorems.

As Ore pointed out, Abel dealt with the sudden appearance of Jacobi by switching his attention to Schumacher's *Astronomische Nachrichten*, and contributing to it an article on the transformation of elliptic functions but one written from a vastly more novel standpoint. "My knockout of Jacobi" he called it in a letter to Holmboe on 29 July, 1828,[31] also calling his (1828b) the "Death-ification of Jacobi" (Stubhaug 2000, 454). Only one participant was not impressed: Gauss replied to Schumacher (23 May 1828) who told Crelle who in turn told Abel that "he (Abel) has come about one third of the way that I have gone in my researches, with the same aim and even with the same choice of notation."[32] We shall consider in Sect. 1.5 the extent to which this unsympathetic verdict is correct, but both men were soon to surpass him and Gauss still did nothing to encourage them.

Evidently, 1828 and 1829 were to be hectic years in the development of elliptic functions. The contemporary response was affected by the context in which all the new work appeared, which had been set by Legendre's theory of elliptic integrals as real functions of their upper end point, x, together with a study of the complete integral K as a function of the modulus, k. The use of transformations which relate elliptic integrals to different moduli appeared in this theory as a means to computing tables of values of the integrals. Questions of multiplication or division of elliptic integrals (given x_0 find x such that $\int_0^x = m\int_0^{x_0}$ for specific rational m) had also been studied by various writers, chiefly because of their geometric interpretation and the curious fact that their solutions often displayed a surprising algebraic, rather than transcendental, dependence of x on x_0.[33] Finally, elliptic integrals arose naturally in astronomy because planets and asteroids traverse ellipses, so there was a working

---

[30] Quoted in Ore (1957, 189).

[31] See Abel (1902, 68). Engl. trl. in Ore (1957, 190).

[32] Gauss to Bessel, 30 March 1828, in Gauss (1880, nr. 63).

[33] The work of Landen and Lagrange is prominent in this connection. There was also a long tradition of work on the rectification of ellipse. See Enneper (1876, Sect. 44) and Houzel (1978).

theory of approximations to elliptic integrals there, which indeed was the occasion of Gauss's only publication on the subject.[34]

Abel's *Recherches* presented a reformulation of much of this material in terms of the complex valued inverse function $x = x(y)$ of the complex variable y. The virtue of this theory was that this inverse function is easier to understand. In particular, the multiplication and division theory of elliptic functions, illuminated by the observation of double periodicity, led to a rich family of polynomial equations many of which were unexpectedly solvable by radicals. Even the more complicated and more general theory of transformations of elliptic functions was illuminated in this way, not only by the discovery of infinitely many unexpected transformations, but by the hint of conceptual order that came with their discovery. The implications of this for applied mathematics were less surely evident. Abel had not discussed them, and Legendre's examples often required the integral, not its inverse function.

In all of this the question of transformations is central. We have seen examples in Jacobi's work. Abel, in the later parts of his *Recherches* took up the theme, in a way that provoked Enneper (1876, 292) to the comment that Abel here showed "a depth of insight, and a richness of surprising ideas which alone would suffice to keep the memory of his wonderful genius alive in the history of mathematics". Independently of Jacobi, Abel introduced (Sect. 41) an infinite number of transformations:

$$\int_0^y \frac{dt}{\sqrt{(1-c_1^2t^2)(1+e_1^2t^2)}} = \pm a\int_0^x \frac{dt}{\sqrt{(1-c^2t^2)(1+e^2t^2)}}, \tag{1.29}$$

where the new moduli, c_1 and e_1, and the number a depended in an explicit way on x and the quantity

$$\alpha = \frac{(m+\tilde{m})\omega + (m-\tilde{m})i\tilde{\omega}}{2n+1}$$

at least one of m or $\tilde{m}$ being prime to $2n+1$. He also gave the formulae for c_1, e_1, and x; they involve his elliptic function ϕ. So Abel used the double periodicity of the elliptic functions to obtain his transformations, and found, as he put it (Sect. 45) "a certain number of transformations corresponding to each value of $2n+1$". Indeed he claimed (Sect. 49) that his transformations combined with Legendre's led to all possible transformations.

A special case of great interest observed by Abel concerned transformations leaving the modulus unaltered and the integral multiplied by a complex number, a:

$$\int_0^y \frac{dt}{\sqrt{(1-t^2)(1+\mu^2t^2)}} = a\int_0^x \frac{dt}{\sqrt{(1-t^2)(1+\mu^2t^2)}}. \tag{1.30}$$

[34] See Gauss (1818), discussed in Geppert (1927).

On such occasions the transformation is called complex multiplication by a.[35] Abel claimed (Sect. 50) that when a was rational or of the form $m \pm i\sqrt{n}$, m and n rational, then the equation above had algebraic solutions, but in the second case the modulus μ could only have certain special values that satisfied an equation with infinitely many real or imaginary roots. These became known as the singular moduli and Abel conjectured that they were the roots of polynomials which were solvable by radicals, a conjecture first proved in Kronecker (1857, see Houzel 2007). Since it will appear (see the Appendix below) that for the elliptic functions x and y that occur in (1.30) to be algebraically related it is necessary and sufficient that the lattice of periods be mapped onto itself, Abel's problem is precisely captured by the theory of complex multiplication, and the condition on a has been explained. The fact that the lattice is special underlies Abel's claim about the corresponding modulus.[36]

Abel's presentation was a significant step in the direction mapped out by Euler and Lagrange; algebraic solutions to differential equations (1.30) were now shown to exist for some irrational a's but only for a precisely delineated class. Abel gave explicit solutions in simple cases: $a = \sqrt{-3}$, when $e = \sqrt{3}+2$, and $a = \sqrt{-5}$ and e satisfies a cubic equation; in each case y is a rational function of x. Finally, when $a = i\sqrt{2n+1}$, he found explicitly that e was given by an infinite series (Sect. 52):

$$e = \frac{4\pi}{\omega}\left(\frac{h}{h^2+1} + \frac{h^3}{h^6+1} + \cdots\right), \tag{1.31}$$

where $h = \exp\left(\frac{\pi}{2\sqrt{2n+1}}\right)$. The proofs of his claims about these differential equations were given in his papers (1828b and Addition) in Schumacher's journal, which is the "knockout of Jacobi" we referred to earlier. Abel ended his *Recherches* with a note comparing his results with Jacobi's, noting that their results were in agreement but that his came with proofs on occasions when Jacobi's did not, and that there were times when his proofs were simpler than those Jacobi had provided.

The Jacobi–Legendre correspondence shows that Jacobi at once began to use Abel's ideas in his own study of transformations, admitting to Legendre (12 April 1828) that he could not do without Abel's analysis. Here appear for the first time several of the striking infinite series that Jacobi was so adept at handling, of which the most remarkable in his opinion was

$$\sqrt{\frac{2K}{\pi}} = 1 + 2q + 2q^4 + 2q^9 + 2q^{16} + \cdots, \tag{1.32}$$

where K is a complete integral and $q = e^{-\pi K'/K}$. Notice that the coefficients give the number of ways the exponents n^2 can be written as a square: as n^2 and as $(-n)^2$—this is trivial, but Jacobi was to draw profound consequences. Jacobi also found this

[35] For a short modern explanation, see the Appendix at the end of this chapter.

[36] We shall not pursue the ways the theory of complex multiplication was to enrich number theory. See the references in Goldstein et al. (2007) or, for a modern mathematical account (Silverman, 1986).

striking identity concerning eighth powers:

$$\Big((1-q)(1-q^3)(1-q^5)\dots\Big)^8 + 16q\Big((1+q^2)(1+q^4)(1+q^6)\dots\Big)^8 = \Big((1+q)(1+q^3)(1+q^5)\dots\Big)^8. \quad (1.33)$$

We mention these identities here because they will entertain us below, when we shall explain how Jacobi derived them (see p. 51). They are deduced ultimately from expressions for sinam u as a quotient of infinite products, which at that time had only been proved by Abel, and Jacobi's admission that he relied on Abel's work at this point provoked Legendre to reply (see Legendre and Jacobi 1875, 415) that it

> testifies to your candour, a quality that accords so well with real talent [...] but having done justice to the beautiful work of M. Abel I place it far below your own discoveries and would like the glory of them, that is to say of their proofs, to belong entirely to you.

Jacobi indeed found independent proofs of the crucial theorems soon afterwards and wrote to Legendre to tell him so on 16 June 1828.

When Jacobi wrote to Legendre on 9 September 1828 he noted that Legendre had not commented on the remarkable series Jacobi had found and noted that the series for $\sqrt{\frac{2K}{\pi}}$ was the key to results about the number of ways a number can be written as a sum of squares. He pointed out that

$$\begin{aligned}\left(\frac{2K}{\pi}\right)^2 &= 1+\frac{8q}{1-q}+\frac{16q^2}{1+q^2}+\frac{24q^3}{1-q^3}+\cdots\\ &= 1+\frac{8q}{(1-q)^2}+\frac{8q^2}{(1+q^2)^2}+\frac{8q^3}{(1-q^3)^2}+\cdots\\ &= 1+8\sum\phi(p)\left(q^p+3q^{2p}+3q^{4p}+3q^{8p}+\cdots\right),\end{aligned}$$

where the sum is taken over all odd numbers p and $\phi(p)$ is the sum of the factors of p. Since this last series contains every power of q, by comparison with fourth power of the series for $\sqrt{\frac{2K}{\pi}}$, one finds that every number is the sum of four squares. A little more work was to allow Jacobi in the *Fundamenta nova* to deduce that the number of ways a number is the sum of four squares is 8 times the sum of its divisors that do not divide 4.

Jacobi's response to the "knockout" is also to be found in his letter to Legendre of 9 September 1828. "It is" he said "as far above my praise as it is above my own work" valuing it particularly for supplying the rigorous proofs his own work lacked. But he pointed out in a brief note in Crelle's *Journal* (Jacobi, 1828b) that he knew there were $p+1$ transformations of order a prime p and gave them explicitly as the result of substituting q^p or any of the p distinct values of $q^{1/p}$ into

$$\sqrt{k} = \frac{2\left(q^{1/4} + 2q^{9/4} + 2q^{25/4} + \cdots\right)}{1 + 2q + 2q^4 + 2q^9 + \cdots}. \tag{1.34}$$

Jacobi was also generous in his praise of Abel to Crelle, who was trying desperately to get a job anywhere in Europe for the young Norwegian, whose health was now being weakened by tuberculosis. Also in September 1828 Legendre, Poisson, Lacroix and Maurice wrote from the Institut de France to the Swedish King Karl Johan urging him to find a place for Abel in Stockholm. On October 25th Legendre finally wrote directly to Abel,[37] and Abel in his reply (25 November 1828) said that "Jacobi will certainly perfect to an undreamed-of degree the theory of elliptic functions but even mathematics in general. No-one can esteem him more highly than I do".[38] This letter is interesting for other reasons, because in it Abel alludes to a general theory whereby it can be determined whether an arbitrary polynomial equation is solvable by radicals or not. We have already seen how closely linked in Abel's mind were the theories of elliptic functions and solvability by radicals, a fact which was to be decisive for the future of both subjects.

The events of 1829 can be briefly described. Abel was by then back in Norway, where a temporary improvement in his health was followed by its collapse towards the end of February. While measures proceeded to get him a suitable academic position, his tuberculosis deepened its hold on him, and on the 6th of April, 1829, Abel died.[39] He was only 26.

The manuscript Abel had been working on for the last few months, his *Précis*, was published posthumously in the fourth volume of Crelle's *Journal* later that year. In it, Abel defined the following functions:

$$\Delta(x,c) = \pm\sqrt{(1-x^2)(1-c^2x^2)} \tag{1.35}$$

$$\tilde{\omega}(x,c) = \int \frac{dx}{\Delta(x,c)} \tag{1.36}$$

$$\tilde{\omega}_0(x,c) = \int \frac{x^2dx}{\Delta(x,c)} \tag{1.37}$$

$$\Pi(x,c,a) = \int \frac{dx}{\left(1-\frac{x^2}{a^2}\right)\Delta(x,c)}. \tag{1.38}$$

The last three are, respectively, elliptic integrals of the first, second and third kinds. The fundamental problem he tackled in the first part of the paper is to find when a sum of such functions (depending on different parameters and evaluated for different values of the variable) reduces to the sum of an algebraic function and the logarithms

[37] See Abel (1902, 77–79).

[38] Abel (1902, 90). Engl. trl. in Ore (1957, 213).

[39] Gauss wrote to Schumacher on 19 May to say that "Abel's death, ... is a very great loss to science". Quoted in Dunnington (2004, 255).

of other algebraic functions of the variables. This is a significant generalisation of the addition formulae of Fagnano and Euler and foreshadows his later paper with arbitrary algebraic integrands (discussed below, see Sect. 4.4). His solution led on to a complete solution of the transformation problem for elliptic functions, which depends on the equation:

$$\frac{dy}{\Delta(y,c')} = \varepsilon \frac{dx}{\Delta(x,c)}, \tag{1.39}$$

where ε is a constant. From that Abel deduced the general transformation of an elliptic function of the first kind and showed that the corresponding moduli were related by an algebraic equation.

In order to be a little more precise, we repeat some of Abel's definitions. Abel introduced the complete elliptic integrals

$$\frac{\tilde{\omega}}{2} = \int_0^1 \frac{dx}{\Delta(x,c)}, \quad \frac{\omega}{2} = \int_0^1 \frac{dx}{\Delta(x,b)}, \tag{1.40}$$

where $b = \sqrt{1-c^2}$ and noted that if $\lambda(\theta)$ is defined by the equation $\theta = \int_0^{\lambda(\theta)} \frac{dx}{\Delta(x,c)}$, then $\lambda(\theta + 2\tilde{\omega}) = \lambda(\theta)$, and $\lambda(\theta + \omega i) = \lambda(\theta)$.

He defined $q = \exp\left(-\frac{\omega}{\tilde{\omega}}\pi\right)$ and set $q_1 = q^{\mu} = \exp\left(-\mu\frac{\omega}{\tilde{\omega}} + \mu' i\right)\pi$, where μ and μ' are rational numbers. Abel then showed that when c is real and less than 1 (but c' may be real or imaginary) and y is an arbitrary function of x, (1.39) can be satisfied if and only if either ε is real and an integer multiple of $\tilde{\omega}'/\tilde{\omega}$, or ε is purely imaginary and a half-integer multiple of $i\omega'/\tilde{\omega}$. In these cases, the modulus c is given by quotient of two infinite products involving q_1, which Abel gave explicitly. Abel also noted that the modular equation connecting c and c', while not generally solvable by radicals, has the property that all roots are known rationally once any two are.

In the second part of the paper, Abel replaced the above functions with their inverses expressed in terms of the function $\lambda(\theta)$, so

$$\tilde{\omega}_0(x,c) = \int \lambda^2(\theta)d\theta\,, \quad \Pi(x,c,a) = \int \frac{d\theta}{\left(1 - \frac{\lambda^2(\theta)}{a^2}\right)}. \tag{1.41}$$

The moduli are now required to be real and less than 1. "In this form", Abel remarked[40], "the elliptic functions have many interesting properties and are much easier to study. Above all, it is the function $\lambda(\theta)$ that merits particular attention". He showed that it can be represented as the quotient of two everywhere convergent power series,

[40]See Abel (1829b, 521) where he referred to his (1826a).

$$\phi(\theta) = \theta + a\theta^3 + a'\theta^5 + \cdots \tag{1.42}$$

$$f(\theta) = 1 + b'\theta^4 + b''\theta^6 + \cdots, \tag{1.43}$$

and that the functions ϕ and f have similar properties to λ.

Abel also broached the subject of arbitrary integrands. In a paper submitted to the French Académie in 1826 he dealt magnificently with the general case, but at the time of his death the paper was sitting as good as forgotten in a pile of papers in the care of Cauchy. He had, however, also published a short note (Abel, 1828c) on a special but significant case in Crelle's *Journal* in 1828, in which he discussed integrals of the form $\int \frac{rdx}{\sqrt{R}}$, where $r(x)$ is an arbitrary rational function and $R(x)$ is an arbitrary polynomial. The theme of that paper was a generalisation of the addition theorem of the *Précis*, and we shall look at it in Sect. 4.4, where we shall see that this paper innocently brought about the rescue of Abel's declining reputation.

Also in 1829 Jacobi published his first definitive account of the theory of elliptic functions, his book *Fundamenta nova theoriae functionum ellipticarum* (New foundations of a theory of elliptic functions). Writing in Latin was a dying custom, kept alive largely in Germany perhaps out of a desire to reach an international audience, and perhaps because a mastery of Latin was a requirement of all German gymnasia. One recalls that Jacobi was particularly good at languages; in 1832 his appointment as a full professor followed a four-hour disputation in Latin.

Jacobi began his treatise with an account of the first problem that had drawn his attention to the subject, the question of the transformation of elliptic integrals. He wrote:

> The problem we shall propose is generally this: to find a rational function y of x such that
>
> $$\frac{dy}{\sqrt{A' + B'y + C'y^2 + D'y^3 + E'y^4}} = \frac{dx}{\sqrt{A + Bx + Cx^2 + Dx^3 + Ex^4}}.$$

He also raised the multiplication problem, which is to find solutions to

$$\frac{dy}{\sqrt{A' + B'y + C'y^2 + D'y^3 + E'y^4}} = \frac{Mdx}{\sqrt{A + Bx + Cx^2 + Dx^3 + Ex^4}}, \tag{1.44}$$

where the quantity M, which may depend on the parameters $A, B, \ldots E$, does not depend on x.

Jacobi's idea was to write $y = \frac{U}{V}$, where U and V are polynomials in x of degrees p and m, respectively, where $p \geq m$. This reduces the integrand in y to another of the same form, possibly multiplied by a constant, provided the constants can be suitably chosen. When this is the case, the transformation is said to be of order p.

Jacobi showed by counting constants that this could be done for any value of p, in other words that there are more coefficients in the polynomials U and V than there are equations for them to satisfy—a typically nineteenth-century approach. Indeed, there will always be three arbitrary constants in any such transformation.

So, as he showed, one can transform $\frac{dy}{\sqrt{(1-y^2)(1-\lambda^2x^2)}}$ into $\frac{Mdx}{\sqrt{(1-x^2)(1-k^2x^2)}}$ with this transformation of order 3:

$$U = x(a + a'x^2) \quad , \quad V = 1 + b'x^2\ . \tag{1.45}$$

The moduli λ and k are then related by this equation, where $u^4 = k$ and $v^4 = \lambda$:

$$u^4 - v^4 + 2uv(1 - u^2v^2) = 0 \tag{1.46}$$

and the multiplier is $M = \frac{v}{v+2u^3}$. The transformation of order 5 produces this relation between u and v:

$$u^6 - v^6 + 5u^2v^2(u^2 - v^2) + 4uv(1 - u^4v^4) = 0\ , \tag{1.47}$$

where $M = \frac{v(1-uv^3)}{(v-u^5)}$. Equations like these between two moduli he called modular equations (Sect. 24); they later caught the attention of Galois, because the cases of low degree can be remarkably simplified.[41]

As we mentioned earlier, Jacobi's solution to these problems proceed by inverting the elliptic integral,

$$u = \int_0^{\phi} \frac{dt}{\sqrt{1-k^2\sin^2 t}} \quad = \int_0^{x} \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}}$$

to obtain elliptic functions. Following Legendre he called ϕ an angle, the amplitude of the function u, and wrote ϕ = am u, x = sinam u. He called x an elliptic function. He defined the complete elliptic integrals

$$K = \int_0^{\pi/2} \frac{dt}{\sqrt{1-k^2\sin^2(t)}}\ , \quad K' = \int_0^{\pi/2} \frac{dt}{\sqrt{1-k'^2\sin^2(t)}}, \tag{1.48}$$

where k and k' are complementary moduli ($k^2 + k'^2 = 1$). To obtain values for an elliptic function with a purely imaginary argument he changed to a new variable defined by the transformation $\sin\phi = i\tan\psi$, $\cos\phi = \sec\psi$. He then used an addition theorem to define elliptic functions for arbitrary complex variables and finally deduced that the new functions were doubly periodic:

Elliptic functions enjoy two periods, one real and the other imaginary, at least when the modulus k is real. Both are imaginary when the modulus itself is imaginary (we could call this the principle of the two periods).

The first half of the *Fundamenta nova* dealt at length with the theory of transformations of elliptic functions. Jacobi showed that transformations of every order do exist and investigated the number of each order. He studied the effect of a

[41] See Gray (2000a, 116–118).

transformation on the moduli and how the composition of two transformations can be made to yield a solution to the multiplication problem. He showed, as Legendre had earlier, that the complete elliptic integrals satisfy the differential equation

$$k(1-k^2)\frac{d^2Q}{dk^2}+(1-3k^2)\frac{dQ}{dk}-kQ=0 \tag{1.49}$$

and applied this differential equation to the study of the multiplier M.

In the second half of the book he showed how to write the elliptic functions as quotients of one infinite product by another, and how the complete elliptic integrals depended on the modulus. Building on Euler's (1768b) he found such striking formulae as (Sect. 36, nr 7):

$$k=4\sqrt{q}\left(\frac{(1+q^2)(1+q^4)(1+q^6)\dots}{(1+q)(1+q^3)(1+q^5)\dots}\right)^4, \tag{1.50}$$

where $q=e^{-\pi K'/K}$. The same expression appears in Abel's posthumous *Précis* (Sect. 7).

In Sect. 52 there appeared for the first time in Jacobi's work a quantity that was to play a more important, foundational role in later versions of the theory: the theta function, defined on p. 198 by the formula

$$\theta(u)=\theta(0)\exp\int\limits_0^u Z(t)dt, \tag{1.51}$$

where Z was defined in terms of the complete and incomplete elliptic integrals of the first and second kinds by the formula $Z(u)=\frac{F^1E(\phi)-E^1F(\phi)}{F^1\Delta(\phi)}$.

Jacobi derived a great many equations involving the theta function and the elliptic functions. He investigated it as a function of a complex variable, established addition formulae for it and showed that it was periodic:

$$\theta(u+2K)=\theta(u) \tag{1.52}$$

but (see Sect. 57) not doubly periodic, and instead that

$$\theta(u+2iK')=-e^{\frac{\pi(K'-iu)}{K}}\theta(u). \tag{1.53}$$

Many of the power series he introduced, and the book is full of them, were in terms of the variable q. He found series for $\mathrm{sinam}(2Kx/\pi)$ and $\theta(2Kx/\pi).\theta(0)$, and came close to writing the function sinam as a quotient of two theta functions. Although the book may have shown to him, as he wrote to Dirichlet, that introducing complex quantities "alone solved all the riddles", his formal approach was to raise as many problems as it solved.[42] It was enormously rich in power series and identities

[42]See Dirichlet (1852, 10).

between new functions. What it is not, and could not have been, is anything like a study of functions in the sense of a theory of functions of a complex variable. It is clear in historical grounds why that is. How it prompted such a theory, and was itself changed by it, forms an important part of this book.

1.4 Elliptic Integrals and Elliptic Functions

To what extent are the first theories of elliptic functions theories of functions of a complex variable? The answer is made harder to obtain by the confused state of the early developments of complex function theory, then almost exclusively in the hands of Cauchy. As Chap. 3 describes, the situation was doubly confused: mathematically confused, in that Cauchy's own ideas were in a turbulent state of development, and confused as regards their dissemination.

There is a straightforward sense in which Abel's and Jacobi's ideas do belong in any history of complex function theory. The objects they discussed were genuinely functions and indeed functions of a complex variable. But if one goes beyond this superficial observation and seeks to insist on points of detail, the ground rapidly becomes slippery. What, for example, was a complex variable to either Abel or Jacobi? What, indeed, was a complex number? Neither ever addressed the point head on. Jacobi's answer would surely have been formal: a complex variable is an expression of the form $x+iy$, to which the usual rules of arithmetic apply and $i^2=-1$. Such formal mathematics was his stock in trade. Abel's answer would have been along the same lines, for if he did not quite match Jacobi's skill with formulae, his interest in mathematics lay heavily on the side of algebra; Sylow, who edited Abel's works, described him as primarily an algebraist. One does not need a geometrical interpretation of complex numbers to pursue the theory of equations.

But how can one justify in a formal way all the manipulations with integrals that they carried out? What does it mean to deduce from $\int_0^u \frac{dt}{\sqrt{1-t^4}} = v$ by means of the substitution $s=it$ that $\int_0^{iu} \frac{-ids}{\sqrt{1-s^4}} = v$, and so $u(v)=iu(iv)$? What does it mean to extend the addition theorem defined for real variables, and for purely imaginary ones if one accepts the validity of the substitution just described, to functions of a complex variable? The answer we propose is that for Abel and Jacobi the validity of the formal operations conferred a common-sense meaning on the formulae that resulted. So the fact that the integral is not well defined once the variables have become complex but is an infinitely many-valued expression is somewhat beside the point. It is the same with $\log x$; it does not have a unique meaning once x can be complex, but $\log x$ does not arise outside of any context. And given a context it will be clear what it means. Common-sense and a little mathematical sophistication enables one to interpret $\int_1^x \frac{dt}{t}$, and to know that $\frac{d\log x}{dx}=\frac{1}{x}$. In the same way, simple formal algebra enables you to work with (many-valued) elliptic integrals and complex valued elliptic functions.

It is interesting that in the first flood of invention, the question did not present itself. At no stage in this work did Abel or Jacobi feel compelled to raise it, much less address it. Cauchy's theory of complex integrals is never mentioned. Nor in their entire correspondence did Legendre see fit to hint that, even at some later stage, they might like to benefit from consulting Cauchy's work. Such a robust, common-sense view has a lot to commend it, but there is a price to pay: the nature of a complex integral and its dependence on the path of integration cannot be raised. Perhaps surprisingly therefore we conclude that in the first phase of the creation of a theory of elliptic functions there was no theory of complex integrals, contrary to the claim in Dieudonné (1974, 1, 44). We shall return to this point in Chap. 4, when we shall argue that the absence even in Cauchy's theory of a good grasp of many-valued integrands drove people to seek other foundations altogether for elliptic functions, and to abandon, at least for a while, the starting point of elliptic integrals.

In the same way, there is no hint in this theory of a classification of infinite points into poles of various orders, let alone a distinction between a pole of finite order and an essential singularity. The fact that these functions have points where they become infinite is described, but the recognition that poles are part-and-parcel of a complex function still lay in the future.[43] As we shall see, this recognition, nowadays encapsulated in Liouville's theorem, was first obtained when mathematicians sought to define a theory of doubly periodic functions without recourse to elliptic integrals.

The division of elliptic integrals into three kinds, and Abel's later extension of Euler's theorem to any algebraic integral, does not of itself imply a recognition that integrands either have no poles, or simple poles, or poles of higher orders. We shall argue in Chap. 4, after this part of Abel's work has been described in detail, that this view is false if it implies that Abel had such an insight, but true if taken to mean that this way of classifying integrands can be defended in such terms. Accordingly, we see the threefold classification of elliptic integrals as another challenge the theory of elliptic functions presented to any one seeking to base it on a theory of complex functions.

In short, Abel's and Jacobi's theories of elliptic functions were purely formal in so far as they were a theory of complex integrals and complex functions of a complex variable, and so presented an important stimulus to anyone seeking to develop it on the basis of an autonomous theory of functions of a complex variable.

On the other hand, the novelty of the theory cannot be denied. Other than polynomials and the trigonometric and exponential functions, there were very few other functions of a complex variable known at the time. The new theory commended itself to its readers by presenting a natural generalisation of the trigonometric functions in a complex setting. Handling these functions, deducing consequences of their addition laws, and so forth, is intrinsically attractive to a

[43]The term "pole" was first introduced in Neumann (1865a, 38).

mathematician. For example, the fact that the division of elliptic functions leads to polynomial equations with the striking property of being solvable by radicals, known to Gauss, Abel and Jacobi, was a profound stimulus to Galois.

Another early application was to number theory. An old assertion of Fermat's is that every positive integer is the sum of at most four squares, but it was not proved until Lagrange in 1770 consummated an argument begun by Euler in 1749 (see Weil 1984, 178–179). Jacobi found that the coefficients r_n in one of his power series were essentially the number of ways n could be expressed as a sum of four squares. Because he also knew about the behaviour of this function, he could obtain a lovely identity for r_n:

$$r_n = 8\sum_{d|n, 4\nmid d} d.$$

In words, r_n is eight times the sum of the divisors of n that are not multiples of 4. He announced this result in his (1828a) and showed how to prove on the last page of his *Fundamenta nova*. Weil (1984, 186) made the interesting observation that Euler had pointed out in a letter to Goldbach in 1750 that a proof using power series would be the most natural way of proving Fermat's theorem; Jacobi certainly did not know of that letter. Because knowing in how many ways a positive integer is a sum of four squares is deeper than simply knowing that it can be done, number theorists were quick to follow Jacobi's lead (we shall describe some of those achievements in more detail in Sect. 4.7).

Jacobi obtained his results[44] by working with the function later labelled

$$\theta_3(0,q) = 1 + \sum_{n=1}^{\infty} 2q^{n^2}.$$

The coefficient of q^j is the number of ways the integer $j = n^2$ can be expressed as a sum of one square (i.e., a single square), where we count n^2 and $(-n)^2$ as distinct. It follows that the coefficient of j in the power series expansion of

$$\theta_3(0,q)^m = 1 + \sum_j a_j q^j$$

represents the number of ways j can be written as a sum of m squares, where we count as distinct ways that differ in the order or sign of the summands. For example, $5 = 1^2 + 2^2$, from which we get a total of 8 representations by replacing 1 with -1, or 2 with -2 or switching the order of 1 and 2 and so on.

This series did not enable Jacobi to write down the number of ways any given number j can be expressed as a sum of squares: a formula for the coefficient of q^j is needed for that. But Jacobi had come at the power series via his theory of elliptic functions, and he knew that they were periodic. He therefore followed the

[44]See the very helpful exposition by Eric Conrad: Jacobi's Four Square theorem, at http://www.math.ohio-state.edu/econrad/Jacobi/sumofsq/sumofsq.html.

clear example set by Fourier in 1822 and expressed his periodic functions as Fourier series. He obtained these Fourier series:

$$\frac{2K}{\pi}\sin\operatorname{am}\left(\frac{2Kx}{\pi},k\right) = \frac{4}{k}\sum_{j=0}^{\infty}\frac{q^{(2j+1)/2}}{(1-q^{2j+1})}\sin(2j+1)x \ .$$

$$\frac{2K}{\pi}\cos\operatorname{am}\left(\frac{2Kx}{\pi},k\right) = \frac{4}{k}\sum_{j=0}^{\infty}\frac{q^{(2j+1)/2}}{(1+q^{2j+1})}\cos(2j+1)x \ .$$

$$\frac{2K}{\pi}\Delta\operatorname{am}\left(\frac{2Kx}{\pi},k\right) = 1+4\sum_{j=0}^{\infty}\frac{q^{j}}{(1+q^{2j})}\cos 2jx.$$

Knowledge about elliptic functions therefore turns into knowledge about these power series. More precisely, the series for Δam, with $z=0$ is

$$\frac{2K}{\pi} = 1+4\sum\frac{q^j}{1+q^{2j}};$$

and the series for its square is

$$\left(\frac{2K}{\pi}\right)^2 = 1+8\sum\frac{q^j}{(1+(-q)^j)^2};$$

and, of course,

$$\theta_3(0,q) = \sqrt{\left(\frac{2K}{\pi}\right)}.$$

Jacobi expanded the series for $\frac{2K}{\pi}$ as a power series and obtained

$$1+4\sum_{j=1}^{\infty}(d_1(j)-d_3(j))q^j,$$

where $d_1(j)$ is the number of divisors of j that are congruent to 1 mod 4 and $d_3(j)$ is the number of divisors of j that are congruent to 3 mod 4. This gave him this theorem: The number of ways a positive integer can be a sum of two squares is equal to four times the difference of the numbers of divisors congruent to 1 and 3 modulo 4.

When Jacobi expanded the series for $\left(\frac{2K}{\pi}\right)^2$ he obtained

$$1+8\sum_{k|j,4\nmid k|}^{\infty}\left(\sum k\right)q^j \ ,$$

where the sum is taken over all divisors of j that are not divisible by 4. This gave him his four squares theorem: The number of ways an integer can be written as the sum of four squares is equal to eight times the sum of all its divisors which are not divisible by 4.

One other application of the new functions is worth mentioning. It rapidly became important, and remained so, and it has recently been studied again by mathematicians (see Griffiths and Harris 1978). This is the interpretation Jacobi gave to a striking theorem of Poncelet about polygons, called the Poncelet closure theorem. The theorem has a complicated history but the crucial step had been taken by Poncelet in his *Traité des propriétés projectives des figures* (1822).[45] Take any two conics C and D with C inside D, say, and suppose for simplicity that both of them are circles, centres C_0 and D_0, respectively. Pick a point P_0 on D and draw a tangent from it to C, which meets D again at P_1, say. Repeat this construction starting with P_1 to obtain P_2, P_2 to obtain P_3, and so on.

In this way a sequence of points $\{P_n\}$ is obtained on D, and Poncelet's theorem asserts that if, for some n, $P = P_0$, the sequence closes and therefore there is an n-gon inscribed in D and circumscribing C, then there is always such an interscribed n-gon whatever position is taken for P_0. Poncelet's proof is a tour de force of geometry. Jacobi, basing his approach on Steiner's investigation of the special cases $n = 5$ and $n = 6$, investigated the relationship between ϕ_n and ϕ_{n-1} where ϕ_n is the angle $P_n\,C_0$ makes with C_0D_0. It is easy enough to find this relationship:

$$\tan\left(\frac{\phi_{n+2}+\phi_n}{2}\right) = \frac{R-a}{R+a}\tan\phi_{n+1},$$

where R is the radius of D and $a = |C_0D_0|$. "In this form of the equations" said Jacobi (1828c, 285) "it springs to the eyes at once that they coincide with those for the multiplication of elliptic functions" and he showed that

$$\phi_n = \mathrm{am}(u+nt),$$

where in his notation $\int_0^{\phi} \frac{dt}{\sqrt{1-k^2\sin^2 t}} = u$, so $\phi = \mathrm{am}u$, and k is an explicit function of R and a. So the n-gon closes when $\mathrm{am}(u+nt) = \mathrm{am}u$, which it does for suitable t and n independently of u (nt has merely to be a period). So unexpected and simple was Jacobi's proof that many later writers referred to it as Jacobi's geometric proof of the multiplication formula for elliptic functions, thus completely reversing the intention of his original argument.

1.4.1 Immediate Responses

The dramatic originality of the work of Abel and Jacobi, its richness, subtlety, and depth, did not make it easy to contribute to their work, and indeed the first responses

[45] Given a fascinating discussion in Bos et al. (1987).

are little more than reworkings of aspects of the original papers. But they do show how widely the news spread.

The Italian geometer and astronomer Giovanni Plana expressed his view in his (1829), which was devoted entirely to giving a quicker derivation of some of the formulae in Jacobi's work and made no mention of Abel's. His fundamental criticism still strikes one as fair: "one does not easily understand by what chain of ideas he could have been led naturally to the singular form that he gives a certain rational function of a single variable that serves as his point of departure for his demonstration .... But once the discovery is made, one can be curious to know if there is effectively another, elementary and direct, way". (Plana, 1829, 333).

The same opinion was held by James Ivory, in a memoir to the Royal Society of London (Ivory, 1829) although he was much more laudatory in his comments on the work of Abel and Jacobi as a whole. Even so (Ivory, 1829, 351), " ...it seldom happens that an inventor arrives by the shortest road at the results which he has created, or explains them in the simplest manner. The demonstrations of M. Jacobi require long and complicated calculations; and it can hardly be said that the train of deductions leads naturally to the truths which are proved". His own contributions was to simplify and unify the Jacobian theory of transformations by pursuing the analogy with the trigonometric functions more closely.

The first book on the subject came out in 1841: Pieter Frans Verhulst's *Traité élémentaire des fonctions elliptiques, [etc.]*, published in Brussels in 1841. This was entirely devoted to the requisite calculations but is further evidenced that the topic of elliptic functions was widely considered to be important.

1.5 Gauss

Gauss's ideas about complex functions emerged in the course of his work on a variety of topics, many of them to do with elliptic integrals and elliptic functions and then with the theory of numbers, so it seems best to follow him in that respect, and postpone a discussion of his ideas about complex numbers and functions until the end of this section.

Carl Friedrich Gauss has, rightly, been the subject of several biographies of which Dunnington (2004) and Wussing (2011) are the most complete and Bühler (1981) the best introduction to Gauss as a mathematician (Fig 1.5). Gauss was born in 1777 in Brunswick to a working class family. His intellectual precocity brought him to the attention of the Duke of Brunswick at the age of 14, who paid him a stipend to attend the Collegium Carolinum in Brunswick. From there he went to the University at Göttingen, still on the Duke's stipend, and by now he was making the first of his many original discoveries. The first to be published was his proof that the regular 17-sided polygon can be constructed by ruler and compass alone, which concealed a profound insight into the structure of polynomial equations from the mathematician he showed it to, the by-then elderly Georg Kaestner. In 1801 Gauss

Fig. 1.5 Carl Friedrich Gauss (Portrait by Schwartz, in Bühler, *Gauss: a biographical study*, Springer ©1981, frontispiece)

published his major book on number theory, his *Disquisitiones arithmeticae*,[46] some of whose implications we shall trace below, but he became famous in the astronomical community that year too.

The asteroid Ceres had been found by the Italian astronomer Piazzi on 1 January, 1801 but it had disappeared behind the sun after 41 nights of observations. The news of a possible new planet generated much excitement, and by September 1801 Piazzi had published all his data. Professional astronomers now raced to find it, but in 1 January, 1802 two German astronomers, Olbers and von Zach, independently confirmed that Gauss's predictions had proved sufficiently accurate to locate Ceres once again (Dunnington 2004, 49–55). It has since become clear that Gauss had invented the least square law in statistics in order to make the best sense of Piazzi's data.[47]

It was astronomy that was to provide Gauss with a career. German telescopes were the state of the art, the heavy computational load was congenial to Gauss because he was also one of the few major mathematicians to have extra-ordinary powers of mental arithmetic, and the value the German intellectual community placed on astronomy helped Gauss to feel he was repaying his debt to the Duke of Brunswick. He was active for many years in the discovery of more and more asteroids and enjoyed his contacts with the astronomical community; men like Bessel and Olbers counted among his friends. It led in turn to the arduous work

[46] See Goldstein et al. (2007).

[47] See Dunnington (2004, 498) and for a vigorous defence of Legendre's rights in the matter Stigler (1986).

Gauss spent on the survey of estates of Hannover in the 1820s, from which he distilled his reformulation of differential geometry and the crucial discovery of the intrinsic nature of (Gaussian) curvature. In the 1830s and 1840s his interest in physics led him to work on geomagnetism, where he organised the first international survey of the Earth's magnetic field and its variation and confirmed that the Earth has only two magnetic poles (some people had suggested four). He and his assistant Wilhelm Weber also wired up Göttingen to equip it with a functioning telegraph, but it worked without amplifiers and so was confined to uneconomic distances. Out of this work came Gauss's influential treatment of potential theory.

And all this time he worked on mathematics: four proofs of the fundamental theorem of algebra; six proofs of quadratic reciprocity; he was the first to be convinced that the geometry of space might be non-Euclidean (ahead of Bolyai and Lobachevskii). It is to his achievements in the fields of complex function theory and elliptic functions that we now turn.

1.5.1 Gauss on Elliptic Integrals and Elliptic Functions

We have already quoted Gauss's remark that Abel had come about a third of the way in his theory of elliptic functions. We shall now consider in detail what Gauss had accomplished in 35 years of intermittent work on that theme. The richness of material forces us to be selective; we shall concentrate on how he organised and reorganised his way through the subject, looking particularly at the implications for a theory of functions of a complex variable. Paradoxically, it was Gauss's marked reluctance to publish that has left the historian with such a wealth of material. With no polished version of his ideas in print, Gauss had no incentive to destroy his working notes. Instead of a book that surely would have hidden the routes to his discoveries we have a profusion of drafts and formulae. All these are collected in the Gauss *Werke*, where, moreover, they have been exceptionally well analysed by Schlesinger and others. Indeed the literature on this topic, from Klein and Schlesinger through Geppert and down to Cox in our own day forms one of the high points in the history of modern mathematics.[48]

Gauss first considered the lemniscatic integral

[48]This material includes Gauss's diary (for an English translation, see Gray 1984b, reprinted with corrections in Dunnington 2004) which enables some discoveries to be dated precisely, some of Gauss's pocket notebooks, often dated on the first page, and other jottings including marginalia. From this it is possible to build up a reasonably detailed chronology of events, as was done by Schlesinger (Gauss, *Werke*, 10.2), and we shall follow his chronology except at one point, to be mentioned below. Accordingly, we have suppressed all the analysis of how the discoveries are dated. The interested reader should consult Cox (1984) and the essay by Schlesinger.

$$z = \int_0^x \frac{dx}{\sqrt{(1-x^4)}} \tag{1.54}$$

in September 1796, when he expanded the integrand as a power series and integrated it term by term, thus obtaining z as a power series in x. This he then inverted by the formal methods of reversion of series, obtaining x as a power series in z. In January 1797 he read Euler's posthumously published paper (1775d) on elliptic integrals, and Stirling's book *Methodus differentialis* of 1730. From Euler's paper he learned this remarkable result: if

$$A = \int_0^1 \frac{dx}{\sqrt{(1-x^4)}} \quad \text{and } B = \int_0^1 \frac{x^2 dx}{\sqrt{1-x^4}}$$

then $AB = \pi/4$. Gauss now began "to examine thoroughly the lemniscate" (Diary entry nr. 51). He called the function z of x defined above by several names; settling on sl for *sinus lemniscaticus* and cl for the cosine, so $cl(x) = sl\left(\frac{\varpi}{2} - x\right)$. We may therefore write one of his findings this way[49]:

$$sl^2(x) + cl^2(x) + cl^2(x)sl^2(x) = 1.$$

This is Euler's addition formula for elliptic integrals in the new setting. Other formulae followed, for such things as $sl\,(2x)$ and $sl\,(x_1 + x_2)$. The one for $sl\,(3x)$ is interesting and caught Gauss's attention. Writing s for $sl(x)$, it is:

$$sl\,(3x) = \frac{s\left(3 - 6s^4 - s^8\right)}{1 + 6s^4 - 3s^8}. \tag{1.55}$$

So division of the lemniscatic sine by 3 leads to an equation of degree 9, whereas division of the ordinary sine leads only to a cubic equation. On March 19 he noted in his diary (entry nr. 60): "Why dividing the lemniscate into n parts leads to an equation of degree n^2". The reason is that all but n of the roots are complex.[50] He thereupon regarded sl and cl as functions of a complex variable. The change of variable from x to ix in the lemniscatic integral establishes that $sl\,(ix) = isl\,(x)$, from which formulae for $sl\,(x_1 + ix_2)$ were written down. Similarly, $cl(ix) = 1/cl(x)$.

Gauss had already seen that the addition formula implies that, as a real function, sl was a periodic function with period 2ϖ, where $\int_0^1 \frac{dx}{\sqrt{1-x^4}} = \frac{\varpi}{2}$. It followed that the complex function sl had two distinct periods, 2ϖ and $2i\varpi$, and so, when m and n are integers,

[49] Here we follow the draft Elegantiores integralis $\int \frac{dx}{\sqrt{1-x^4}}$ proprietatis, which dates from before 1801, see Gauss *Werke* 3, 404–412.

[50] Two days later he showed that the lemniscate is divisible by ruler and compass into five parts, thus finding an unexpected parallel between the divisibility theories of the lemniscate and the circle.

$$sl\,(x+(m+in)\,2\varpi) = sl\,(x)\,,$$

the first occurrence of the Gaussian integers, $m+in$ (see *Werke* 3, 411–412). It was therefore easy to write down all the points where sl and cl were zero or infinite, and so for Gauss to write sl and cl as quotients of two infinite series. Gauss wrote $sl(x) = M(x)/N(x)$ and looked at the functions M and N in their own right as functions of a complex variable. The formulae giving $sl\,(2x)$ in terms of $sl\,(x)$ yielded formulae for such things as $N\,(2x)$, and special values of $sl\,(x)$ could be found from the simplest ones by using the addition formulae. So Gauss did one of his provocative little sums. He calculated $N\,(\varpi)$ to five decimal places and its log to 4 and noted that this seemed to be $\pi/2$. He wrote "Log hyp this number $= 1.5708 = \frac{1}{2}\pi$ of the circle?" (*Werke*, 10.1, 158). In his diary (29 March 1797, nr. 63) he checked this coincidence to six decimal places and commented that this "is most remarkable and a proof of this property promises the most serious increase in analysis".

This observation excited Gauss because he connected it to Euler's remarkable result $AB = \pi/4$. Here $A = \varpi/2 = \int\limits_0^1 \frac{dx}{\sqrt{(1-x^4)}}$ is the value of the complete lemniscatic integral, and so B must be $\pi/2\varpi$. This yields another connection between ϖ and π.

1.5.1.1 The Arithmetic–Geometric Mean

In July 1798 an improved representation of the numerator and denominator of the lemniscatic functions led Gauss back to the calculation of ϖ, and he commented in his Diary (July, nr. 92): "We have found out the most elegant things exceeding all expectations and that by methods which open up to us a whole new field ahead". His entry to this new field was, however, still a year away. He found it on returning to an old love of his, the arithmetic–geometric mean. This is defined for two real numbers a and b as follows: set $a_0 = a$ and $b_0 = b$, and recursively

$$a_{n+1} = \frac{1}{2}\,(a_n + b_n)\;, \quad \text{and} \quad b_{n+1} = \sqrt{a_n b_n}. \tag{1.56}$$

Then it is easily seen that the two sequences (a_n) and (b_n) converge to the same value, called the arithmetic–geometric mean of a and b, abbreviated to the agm and written $M\,(a,b)$ here.[51] Since $M\,(\lambda a, \lambda b) = \lambda M\,(a,b)$, $M\,(a,b) = aM\left(1, b/a\right)$, and it is enough to study $M\,(1,x)$. On 30 May 1799 Gauss wrote in his Diary (nr. 98): "We have found that the arithmetic–geometric mean between 1 and 2 is π/ϖ to 11 places, which thing being proved a new field in analysis will certainly be opened up". A letter from Pfaff to Gauss (*Werke* 10.1, 232) makes it clear that even in November 1799 the proof was eluding Gauss.

[51] Convergence only fails if the initial values chosen include 0.

One of the Gauss's earliest results about the agm was a power series for $M(1,1+x)$. It seems not to be known how Gauss first came to this series, but in 1800 he gave this method (*Werke* 3, 365). He found, on writing $x = 2t + t^2$, that

$$\begin{aligned} M(1,1+x) &= M\left(1,1+2t+t^2\right) \\ &= M\left(1+t,1+t+t^2/2\right) = (1+t)M\left(1,1+\frac{t^2}{2(1+t)}\right) \end{aligned}$$

so

$$M\left(1,1+2t+t^2\right) = (1+t)M\left(1,\frac{t^2}{2(1+t)}\right). \tag{1.57}$$

So, on setting $M(1,1+x) = 1 + ax + bx^2 + \cdots$ and expanding both sides and collecting like terms, he obtained the coefficients of the power series expansion of $M(1,1+x)$.[52] The series begins

$$M(1,1+x) = 1 + \frac{x}{2} - \frac{x^2}{16} + \frac{x^3}{32} - \frac{21x^4}{1024} + \cdots.$$

Gauss also considered two functions related to the agm and which also have attractive power series expansions. The first is $M(1+x,1-x)$, and the second its reciprocal.[53] This second function turned out to be most important in all the future developments. It has this expansion:

$$K(x) = \frac{1}{M(1+x,1-x)} = 1 + \frac{x^2}{4} + \frac{9x^4}{64} + \frac{25x^6}{256} + \cdots. \tag{1.58}$$

The coefficient of x^{2n} is $\left(\frac{1.3.5....(2n-1)}{2.4.6....2n}\right)^2$. He showed formally that the function $K(x)$ satisfies the linear ordinary differential equation

$$\left(x^3 - x\right)K'' - K'\left(3x^2 - 1\right) + xK = 0 \tag{1.59}$$

and noted that another independent integral of this differential equation is $M^{-1}(1,x)$.

Gauss connected this result with an earlier calculation of his (*Werke* 10.1, 267). As soon as he had considered the functions *sl* and *cl* as functions of a complex variable, he had written down the formal Fourier series expansion of $\frac{1}{\sqrt{(1+\sin^2 V)}}$:

[52]Gauss did not check the convergence of this series, but the agm converges very fast and its use is the key to the rapid evaluation of elliptic integrals.

[53]Schlesinger (1912, 63) suggested that it was the appearance of the reciprocal of $M(1,\sqrt{2})$ that led Gauss to consider not the agm in general but its reciprocal.

$$\frac{1}{\sqrt{(1+\sin^2 V)}} = a + b\cos 2V + c\cos 4V + d\cos 6V + \cdots \tag{1.60}$$

and observed that

$$a = 1 - \frac{1^2}{2^2} + \frac{1^2.3^2}{2^2.4^2} - \cdots = \frac{2}{\pi}\int_0^1 \frac{dx}{\sqrt{1-x^4}}\,, \tag{1.61}$$

which is, of course ϖ/π. Now he expanded $1\Big/\sqrt{(1+y\sin^2 V)}$ similarly and wrote down that the reciprocal of $M\left(1,\sqrt{1+y}\right)$ is the V-free part (as he called it, *Werke* 10.1, 183) of the expansion of $\frac{1}{\sqrt{1+y\sin^2 V}}$. This gave him that

$$M\left(1,\sqrt{1+y}\right).\int_0^1 \frac{dV}{\sqrt{1+y\sin^2 V}} = \pi/2\,.$$

He now set $y = -k^2$, and this became

$$M\left(1,\sqrt{1-k^2}\right).\int_0^1 \frac{dV}{\sqrt{1-k^2\sin^2 V}} = \pi/2\,. \tag{1.62}$$

So Gauss had shown that the complete elliptic integral, regarded as a function of the modulus k, was given by the reciprocal of the agm at the complementary modulus, $k' = \sqrt{1-k^2}$.

The significance of the clue given by the value of π/ϖ is now revealed: when $k = i$ the above expression for M reduces to

$$M\left(1,\sqrt{2}\right).\int_0^{\pi/2} \frac{dV}{\sqrt{1+\sin^2 V}} = \pi/2.$$

By means of the known value of $\varpi/2$ for the complete lemniscatic integral

$$\int_0^{\pi/2} \frac{dV}{\sqrt{1+\sin^2 V}} = \int_0^1 \frac{dx}{\sqrt{1-x^4}}\,,$$

this reduces to $M\left(1,\sqrt{2}\right) = \pi/\varpi$.

Thus Gauss had the conceptual explanation he sought for the numerical coincidence that had intrigued him for so long, and it was obtained by generalising the original question about complete lemniscatic integrals to the setting of complete elliptic integrals. The enigmatic value of $N(\varpi)$ was simultaneously explained from the power series expansion of N, for the value of $\int_0^\pi \frac{d\phi}{\sqrt{1-k^2\cos^2\phi}}$, a complete elliptic integral of the first kind, is $\pi N(k)$.

During the first few months of 1800 Gauss pursued this insight until he could confide in his Diary on May 6 (nr. 105) that he had led the theory of elliptic integrals "to the summit of universality". Even that estimation had to be promptly set aside. On May 22 the theory was "greatly increased and unified", becoming "most beautifully bound together and increased infinitely" (nr. 106). These entries, and two more written on or before June 3, make it clear that Gauss, who had already studied the lemniscatic functions and obtained a crucial insight into the value of the general complete elliptic integral, now proceeded to invert the general elliptic integral. In short, he began to study the general elliptic function with (real) modulus k. The formula connecting the complete integral and the agm immediately gave him an expression for a period in terms of the modulus.

Some aspects of the generalisation are straightforward, although never trivial.[54] A substantial result was obtained by Gauss in May 1800 (Diary entry nr. 108): the new elliptic functions can be expanded as quotients of power series. So Gauss now, as other authors were to later, pursued two approaches. One proceeded from the inversion of an elliptic integral depending on a parameter k. The other developed elliptic functions directly as quotients of entire functions, without reference to an integral. The natural question then is: do these two approaches describe the same objects?

On the first approach, an elliptic integral depends on its modulus, k. The corresponding inverse elliptic function has periods that are functions of this modulus, and these periods are multiples of the complete integrals. Gauss's second approach proceeded in a way that was to be paradigmatic in the theory of complex functions: he sought to represent his new functions as quotients of (to use modern language) holomorphic[55] functions. That is, he represented a function f as a quotient $f(z) = \frac{g(z)}{h(z)}$ where g and h have no poles in the complex plane and the zeros of g correspond to the zeros of f while the zeros of h correspond to the poles of f. This approach would have been more influential if Gauss had published it in 1800 when he discovered it, but he left it to others to discover, as, happily, they did, and it was only published in 1868 in the third volume of his *Werke*.

In the lemniscatic case, Gauss had introduced functions P and Q as follows. He introduced a new variable, $s = \sin(\varphi\pi/\varpi)$. The function $sl(\varphi)$ has period 2ϖ, so it can be written as a function of s. Gauss found that $sl(\varphi) = P(\varphi)/Q(\varphi)$, where P and Q had zeros where the quotient function sl has, respectively, zeros and infinities, and the relationships between $P(\varphi)$ and $P(\varphi + \text{period})$ and $Q(\varphi)$ and $Q(\varphi + \text{period})$ ensure that the quotient is a doubly periodic function.

[54]For example, the addition theorem, written down in November 1799, see (*Werke* 10.1, 196).

[55]This term was introduced by Briot and Bouquet in (1875, 14).

More precisely, Gauss showed (*Werke* 3, 416) that P can be written as this infinite product, in which $s = \sin(\frac{\pi}{\varpi}\varphi)$:

$$P(\phi) = (\varpi/\pi)s\prod_{k=1}^{\infty}\left(1+\frac{s^2}{\sinh^2 k\pi}\right)$$

and Q as this one:

$$Q(\phi) = \prod_{k=0}^{\infty}\left(1-\frac{s^2}{\cosh^2(2k+1)\pi/2}\right).$$

Gauss also expressed P and Q as Fourier series (*Werke* 3, 418), where

$$P(\phi\varpi) = \\ 2^{3/4}\sqrt{\frac{\pi}{\varpi}}\left(e^{-\pi/4}\sin(\phi\pi) - e^{-9\pi/4}\sin(3\phi\pi) + e^{-25\pi/4}\sin(5\phi\pi) - \cdots\right) \quad (1.63)$$

and

$$Q(\phi\varpi) = 2^{-1/4}\sqrt{\frac{\pi}{\varpi}}\left(1+2e^{-\pi}\cos(2\phi\pi)+2e^{-4\pi}\cos(4\phi\pi)+\cdots\right).$$

This makes it very clear that Gauss had already come to the theory of theta functions that was later to animate Jacobi (and indeed Gauss had four such functions—we have suppressed his R and S).

To obtain the functions that do for a general elliptic function what P and Q do for the lemniscatic function, Gauss merely had to observe how the quotient P/Q has the appropriate periods in the lemniscatic case and make appropriate modifications.[56] Since he knew how the periods depend on the modulus, his journey from integral to quotient was concluded.

The difficult way round is to see that every function defined as a quotient arises as the inverse of a suitable elliptic integral. Gauss had shown (*Werke* 10.1, 69) that the periods of a general elliptic function obtained from an elliptic integral with modulus k (i.e. of the form $\int \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}}$) are

$$\varpi = \frac{\pi}{M\left(1,\sqrt{1-k^2}\right)}, \quad \text{and } \varpi' = \frac{\pi}{kM\left(1,\sqrt{1+\frac{1}{k^2}}\right)}. \quad (1.64)$$

[56] The word "merely" here is ironic; the step is hardly an easy one.

If one follows Gauss and writes $k = \tan v$, these expressions for ϖ and ϖ' become[57] $\pi \cos v / M(1, \cos v)$ and $\pi \cos v / M(1, \sin v)$. So the problem is to show, given functions P and Q, say, whose quotient P/Q has periods $m\varpi + m'i\varpi'$, that ϖ and ϖ' can be expressed as these functions of k: $\varpi = \pi \cos v / M(1, \cos v)$ and $\varpi' = \pi \cos v / M(1, \sin v)$.

By a simple scaling argument, it is enough to find k such that the ratio ϖ/ϖ' is correct. This ratio is $M(1, \cos v)/M(1, \sin v)$, and so Gauss defined the function $z = \exp\left(-\frac{\pi}{2} \cdot \frac{M(1, \cos v)}{M(1, \sin v)}\right)$. The problem is now to show that this function takes every possible value between 0 and 1 as v goes between 0 and $\pi/2$. This is reasonably clear from the power series expansion, which shows that as t tends to 0, $M(1,t)$ tends to infinity.[58]

There was a further generalisation to be made. All the elliptic functions that have arisen so far have had one real and one purely imaginary period. This came about because the corresponding elliptic integrals had a real modulus. This restriction arose historically because real quantities were regarded as completely general, and because elliptic integrals with real moduli make geometrical and physical sense. Nonetheless, once elliptic functions have been shown to be best considered as complex functions of a complex variable, it is possible to abandon this restriction on the periods. Once this is done, the ratio of the periods can no longer be purely imaginary, and so the agm must itself become a complex-valued function of a complex variable. This raises a very difficult problem indeed.

The nub of the difficulty is that the square root of a complex number is not uniquely defined, nor is there a workable convention for selecting a square root akin to the rule "choose the positive root of a positive number". Instead, at each stage in the computation of an agm of two complex numbers, a choice must be made.[59] This freedom of choice affects the outcome, but it does not wreck the whole story. It turns out that at each stage there is always a correct choice for the square root which ensures that the sequences defining the agm converge to a common limit. Indeed, if the numbers at the nth stage are a_n and b_n (and therefore $a_{n+1} = (a_n + b_n)/2$), then b_{n+1} is the correct root of $a_n b_n$ if $|a_{n+1} - b_{n+1}| \leq |a_{n+1} + b_{n+1}|$ and, if equality holds, $\mathrm{Im}(b_{n+1}/a_{n+1}) > 0$. Moreover, if a finite number of incorrect choices are made, the series still converge to a common limit, although generally one different from the limit obtained when all the choices were correct.

These convergence claims are not difficult to verify. Gauss seems to have yet again found the way forward by a computation, not of the most general case but of a case in which one of the initial values was negative. Much more striking is his observation, made in the diary on 3 June 1800 (nr. 109), that connection

[57] See (*Werke* 10.1, 194 and 197), and (Cox, 1984, 325).

[58] Schlesinger suggests that Gauss was aware of the much more precise limit $\lim_{x\to\infty} \frac{1}{x} M(1,x) \log 4x = \pi/2$, see (*Werke* 10.1, 268).

[59] How this can be done is well described in Cox (1984).

between the infinitely many means has been completely cleared up. It is followed two days later by the remark (nr. 110): "We have now immediately applied our theory to elliptic transcendents", confirming the interpretation that the way Gauss took to generalising his theory of elliptic functions was to make the agm into a complex function. Unfortunately, nothing survives from 1800 to indicate what Gauss's complete solution was. The best indication we have is from a note from as late as 1825, which says (*Werke*, 10.1, 219) "The agm changes if one chooses a negative value (for the square root): yet all results take the following form: $\frac{1}{(\mu)} = \frac{1}{\mu} + \frac{4ik}{\lambda}$". This asserts a fairly simple connection between any three values λ, μ and (μ), of the agm of two numbers, which is indeed valid under certain assumptions one may suppose Gauss made tacitly; one requires that both $(\mu)/\lambda$ and μ/λ have positive real part. If this requirement is dropped, then one obtains instead:

$$\frac{1}{(\mu)} = \frac{\delta}{\mu}\frac{\gamma}{\lambda}, \tag{1.65}$$

where δ and γ are integers satisfying $\delta \equiv 1, \gamma \equiv 0 \pmod 4$.

It would be interesting to know when Gauss came to this insight into the agm as a complex function. It is intimately tangled up with other questions, such as his developing knowledge of the study of certain related functions as complex functions. The mathematical situation is that there is no way known of proving the claim about the agm which does not lead through some of the deepest parts of the theory of these other functions as complex functions.

Schlesinger (1912, 95) argued that Gauss seems not to have calculated agms of complex numbers directly, but to have worked with functions p, q and r defined as follows (see Gauss *Werke* 10.1, 218)

$$p(t) = 1 + 2t + 2t^4 + 2t^9 + \cdots, \tag{1.66}$$

$$q(t) = 1 - 2t + 2t^4 - 2t^9 + \cdots, \tag{1.67}$$

$$r(t) = 2t^{1/4} + 2t^{9/4} + 2t^{25/4} + \cdots \tag{1.68}$$

and connected to the agm of a and b this way: $p^2(t) = a/M(a,b)$, $q^2(t) = b/M(a,b)$, $r^2(t) = a^2 - b^2$. At one stage in 1810 Gauss (*Werke* 10.1, 224) studied how the functions p, q, and r transform under the transformation

$$t \mapsto \frac{\alpha t - \beta i}{i\gamma t + \delta} \tag{1.69}$$

for the six cases where the matrix is congruent to the identity matrix mod 2. These transformations show (following Gauss, *Werke* 3, 478, dated 1827) that the ratios $p^2 : q^2 : r^2$ are unaltered by transformations where γ and δ are as above. Schlesinger interpreted this as the mathematical meaning of the Diary entry nr. 109 of June 3.

Elsewhere (in 1805, see *Werke* 3, 477 and 10.2, 102–103), Gauss drew the diagram in Fig. 1.6, which was misunderstood by Schering when editing the third

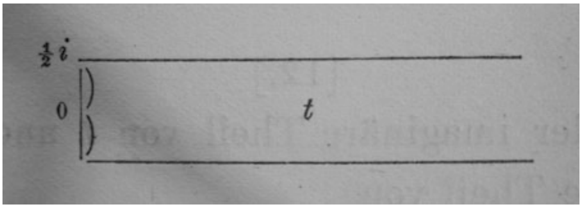

Fig. 1.6 Gauss's diagram as reproduced in *Werke* 3, 477

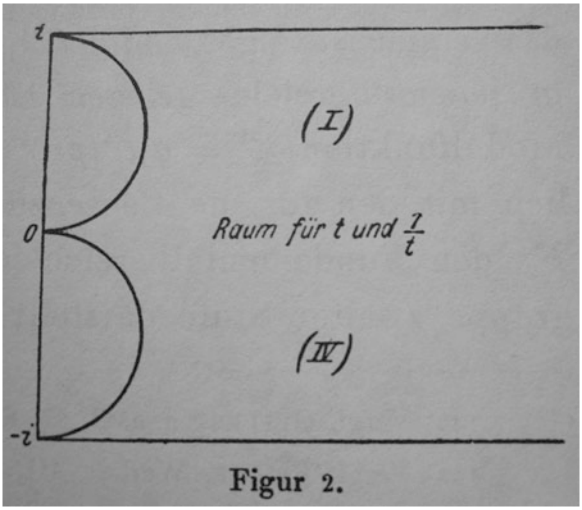

Fig. 1.7 Gauss's diagram as corrected in *Werke* 8, 105

volume of Gauss's *Werke*, p. 478, but corrected by the editors of Vol. 8 (see *Werke* 8, 105 and Fig. 1.7). Its significance is that any function of a complex variable t which is invariant under the substitutions $t \mapsto t+i$ and $t \mapsto 1/t$ takes all its values on precisely the domain indicated.

In another place (*Werke* 8, 101) Gauss observed that, if $q(t)/p(t) = A$ is regarded as an equation for t, then the solution is $t = \mu/\lambda$ where, if $A^2 = n/m, \mu = M(m,n)$ and $\lambda = M(m+n, m-n)$. To connect this with what has gone before, suppose without loss of generality that $m = 1$ and replace $n = A^2$ by $x = \cos v$ so $\mu/\lambda = M(1,x)/M(1,(1-x^2)) = M(1,\cos v)/M(1,\sin v)$, which we shall temporarily write $g(x)$. Write the other quotient $q(t)/p(t)$ as $f(t)$. Then the claim is that $f^2(t) = x$ if $t = g(x)$, in other words, f^2 and g are inverse functions. Moreover, Gauss went on, "one obtains thus only one value t; all others will be contained in the formula $\frac{\alpha t - 2\beta i}{\delta - 2\gamma t i}$, where α, β, γ, and δ are all integers that satisfy the equation $\alpha\delta - \beta\gamma = 1$". Putting this together with the previous results, we see that Gauss had a complete solution to the inversion of elliptic integrals even when they depended on a complex parameter k.

The only question is when. Schlesinger and Cox argue that the Diary entry nr. 109 of June 3 1800 enables one to date the discovery to 1800. The absence of accompanying material, and the presence of appropriate material of a later date is, on this view, just an accident of transmission. An alternative view would be that the surviving material indicates accurately when these difficult insights were made, and that the Diary entry is ambiguous. There are occasions when Gauss made claims in his Diary of results whose truth he had become certain but whose proof eluded him; the striking claims made in entry nr. 92 and their delayed resolution have already been discussed, and nr. 118, May 1801 is another case in point (see below Sect. 4.7.1). It is surely possible that the diary entries mean only that Gauss had begun to investigate the agm for complex values, and had been pleased to suspect a connection between them. The detailed verification had to wait for a much later date.

Since Gauss never published an account of his discoveries, dating them precisely is interesting but not vital, and we shall not insist on it further. In this book we have three questions to ask about Gauss's work:

- How does it compare with the later discoveries of Abel and Jacobi?
- To what extent was it an example of complex function theory?
- And to what extent did it influence his ideas about complex functions in general?

The theory Gauss developed of what might be called elliptic functions with a real modulus had, as we saw, two aspects: a direct approach starting with the inversion of the corresponding elliptic integral and an indirect approach starting from some suitable power series. The agreement between these two theories was assured by establishing certain properties of the agm, thought of as a function of the real modulus. The fruits of these theories were such things as the rich collection of identities between various functions, and results about the division of elliptic functions. When it came to the general theory for an arbitrary complex modulus, Gauss, however, abandoned the twofold structure of the real theory. There is no study of elliptic integrals with complex modulus starting from the Euler addition theorem and proceeding via the introduction of complex variables and the inversion of the integral to the double periodicity of the corresponding elliptic functions and their division equations. Nor, as we have seen, was there any extensive numerical investigation of the agm of pairs of complex numbers. Instead, the weight of the investigation was placed on generalising to the complex case the power series side of the real theory. So our first conclusion is that complex integrals were not considered an acceptable starting point for Gauss.

This raises the question of the extent to which he continued to regard elliptic integrals with a real modulus but complex end points and paths of integration as well understood. We shall discuss this point further below, when we look at his developing ideas about complex function theory.

The theory Gauss did develop was based on power series in complex variables. Gauss was certain that a viable theory of a complex variable could be based on the representation of complex numbers as points in the plane. In his *Disquisitiones arithmeticae* of 1801 he had illustrated his discoveries in number theory with that representation, for example in his theory of the ruler and compass construction constructibility of the regular 17-gon. So it is entirely natural that some of his deepest results about the transformation of the complex functions he was studying should be couched in three overlapping modes. He stressed the number-theoretic side of his transformations, he presented them explicitly in terms of their effect on the power series, and he accompanied them with the geometrical diagram for the function q/p discussed above. Our second conclusion is therefore that Gauss went beyond a formal or purely algebraic theory of a complex variable and embraced its geometric representation. The implications this was to have for his general study of complex functions will be further discussed below.

How far then did his theory surpass that of Abel's in 1828? Gauss himself wrote to Bessel that Abel "has followed the same path that I embarked upon in 1798, so that the great coincidence of the results is not to be wondered at ..." (Gauss, 1880, 477). Indeed, in some cases even the notation was the same; for example, both denoted the value of the complete lemniscatic integral by ω. Schlesinger suggested (1912, 184) that Gauss's mostly unpublished theory looked like this by 1828:

> The first third was the general theory of functions arising from the [hypergeometric series], the second the theory of the agm and the modular function, and finally the third, which Abel published before Gauss, was the theory of elliptic functions in the strict sense.

Schlesinger might have added that Gauss would have appreciated Abel's consideration of the extent to which the division equations are solvable algebraically, and the investigation of the divisibility of the lemniscate by ruler and compass would have struck him most forcefully, inspired as it was by the hint he had dropped in his *Disquisitiones arithmeticae*. He would have savoured the expansion of the elliptic functions in infinite series and infinite products but would have noted that Abel did not seem to have appreciated their full significance. Gauss would also have seen, however, that what Abel presented was only an account of elliptic functions with a real modulus.

Had Gauss chosen to comment a year later on Jacobi's *Fundamenta nova* instead, he would have seen a general theory of the transformations of elliptic functions surpassing anything he had written down. He would also have found something more like his theory of theta functions, but again only in the context of elliptic functions with real modulus. He might also have become aware of Abel's account of the same power series in Abel's (1828a). In that paper Abel also allowed the modulus to become purely imaginary, but in all respects his theory of a complex variable was, like Jacobi's, entirely formal.

It is the truly complex nature of Gauss's theory that stands out. It is not just of greater scope, it was deeper mathematically. As we discuss further below, Gauss was clear about the importance of the complex domain. Here, let us only recall the geometrical account of how the ratios of the functions $p^2 : q^2 : r^2$ depended on the complex variable t described above.

1.5.2 The Hypergeometric Series and Equation

As Schlesinger noted, it is therefore rather paradoxical that Gauss chose not to follow his own advice when he published the most fully developed of his ideas about complex functions. Or rather, when he published part of his ideas; the resolution of this inconsistency may well, as Schlesinger suggested, lie in Gauss's habitual reluctance to publish. Gauss's published paper of 1812 deals with the hypergeometric series[60]

$$F(\alpha,\beta,\gamma,x) = 1 + \left(\frac{\alpha.\beta}{1.\gamma}\right)x + \left(\frac{\alpha(\alpha+1)\beta(\beta+1)}{1.2.\gamma(\gamma+1)}\right)x^2 + \cdots \tag{1.70}$$

It is, as Gauss showed, capable of representing many known functions, such as the trigonometric functions. Cases of it arise as the coefficients in the Fourier series expansion of $(a^2+b^2-2ab\cos\theta)^{-n}$, so it is of practical use in astronomy.[61] There are intriguing, simple relationships between what Gauss called contiguous F's, two F's where one value of α, β, or γ differ by 1. Finally, the value of $F(\alpha,\beta,\gamma,1)$ depends on an interesting way on Gauss's factorial function. This is a function which satisfies $\Pi(n) = n!$ for integer n and $\Pi(z+1) = (z+1)\,\Pi(z)$ and is in fact related to Euler's Gamma function $\Gamma(z)$ by the simple formula $\Pi(z) = \Gamma(z+1)$.[62] Using his newly developed theory of this function, Gauss (*Werke* 3, 150) gave an immediate proof of formulae that, he said, Euler had worked very hard to obtain, such as the formula that had intrigued Gauss for so long, $A.B = \pi/4$. Gauss still did not reveal any of his theory of the agm and elliptic functions. Moreover, throughout this, the published part of the paper, the only comment on the nature of the variable x is that it must be a complex variable of modulus less than 1 for the series to converge (nor may γ be a negative integer).

Matters are different with the part of this paper that Gauss did not publish.[63] The hypergeometric series satisfies this differential equation,[64] called the hypergeometric equation:

$$\left(x-x^2\right)\frac{d^2F}{dx^2} + (\gamma-(\alpha+\beta+1)x)\frac{dF}{dx} - \alpha\beta F = 0. \tag{1.71}$$

But conversely, which is what interested Gauss, the hypergeometric equation defines a function of a complex variable x everywhere except at the points $x=0$, $x=1$, and $x=\infty$. This function can be represented inside the unit circle by the hypergeometric series, but unlike the series it makes sense outside it. Like the

---

[60]For its prehistory see Schlesinger's essay and for the later history of this important equation (Gray, 2000a).

[61]Gauss used it to deal with the net gravitational effect of a planet in an elliptical orbit in Gauss (1818). See the German translation and commentary by Geppert in 1927.

[62]Legendre introduced the symbol Γ for the Gamma function in his (1811–1817, 1, 277). Euler had come across the Gamma function in his (1729).

[63]This is Gauss (1812b), which continues (Gauss, 1812a) and is numbered in consecutive sections.

[64] The observation that the hypergeometric series satisfies the hypergeometric equation had been made by Euler in his *Institutionum calculi integralis*, vol. 2, Part 1, Chaps. 8–11 and later in the paper (Euler, 1778) published posthumously.

series, the equation has a prehistory. A special case of it, with $\alpha = \beta = 1/2$ and $\gamma = 1$, is Legendre's differential equation (1.14) after the substitution $c^2 = x$, and another, $\alpha = -\beta = 1/2$ and $\gamma = 1$, is Legendre's differential equation (1.15) after the substitution $c^2 = x$. Again, like the series, there are simple connections between the hypergeometric equation and its transform under maps such as $x \mapsto 1/x$ and $x \mapsto 1-x$.

The hypergeometric equation illuminates an understanding of Gauss's ideas about functions of a complex variable in two ways. Gauss made it clear that all solutions to the equation can be expressed as a sum of two of them, but the series expressions he had for the solutions were restricted to their domain of convergence. He put forward many solutions to it in the form $f(y)P(y)$, where $P(y)$ is a suitable hypergeometric series, the complex variable is x, $1-x$, or $\frac{1}{x}$ and the corresponding $f(y)$ is a power of x, $1-x$, or $\frac{1}{x}$. This exhibits the solutions to the differential equation as an analytic function valid in a neighbourhood of $x = 0$, $x = 1$, or $x = \infty$ multiplied by a term that captures the behaviour of the solution in that neighbourhood.

Although Gauss gave no special word to those three points in this context, it is clear that he appreciated their role both as setting bounds on the radius of convergence of a hypergeometric series, and in terms of their effect on the nature of the solutions to the hypergeometric equation. This is most apparent in his discussion of the following paradoxical result (his term, Sect. 55). By first setting $x = 4y - y^2$, and $\gamma = \alpha + \beta + 1/2$, when F becomes equal to $F(2\alpha, 2\beta, \alpha + \beta + 1/2, y)$, and then changing y to $1-y$, Gauss found that his earlier results implied that

$$F\left(2\alpha, 2\beta, \alpha + \beta + \frac{1}{2}, y\right) = F\left(2\alpha, 2\beta, \alpha + \beta + \frac{1}{2}, 1-y\right) \tag{1.72}$$

"which equation is certainly false" (*Werke* 3, 226). To resolve the paradox, he said, one has to distinguish between F as a function and F as an infinite series. The former is defined for all finite values of the variable except 0 and 1, while the latter is only defined when the variable is less than 1 in absolute value. However, when defined the series takes a unique value for each value of the variable, whereas the function does not. The false equation is in fact meaningless because as a result about series they have distinct domains of convergence, and as a result about functions the functions are many-valued. One could no more deduce a false result here, he remarked, than one could validly infer from $\arcsin(1/2) = 30^\circ$ and $\arcsin(1/2) = 150^\circ$ that $30^\circ = 150^\circ$.

What Gauss did not say was that he had another powerful reason for studying the hypergeometric series, one that derived from his deep study of the complex theory of elliptic functions. Indeed, as he wrote in 1800, the functions $y_1 = \frac{1}{M(1+x, 1-x)}$ (which we denoted by K above, see (1.58)) and $y_2 = \frac{1}{M(1,x)}$ satisfy the differential equation:

$$\left(x^3 - x\right)\frac{d^2y}{dx^2} + \left(3x^2 - 1\right)\frac{dy}{dx} + xy = 0. \tag{1.73}$$

This is not the hypergeometric equation, but on making the transformation $x^2 = z$ it becomes the equation

$$z(1-z)\frac{d^2y}{dz^2} + (1-2z)\frac{dy}{dz} - \frac{y}{4} = 0, \tag{1.74}$$

which is the hypergeometric equation with $\alpha = \beta = 1/2, \gamma = 1$ (in fact, Legendre's equation (1.14)).[65] We can be certain Gauss knew this. Indeed, Schlesinger points out (1917, 91) that Gauss's later derivation of the hypergeometric equation for the hypergeometric series follows exactly his derivation (*Werke* 10.1, 177–180) of the power series for $\frac{1}{M(1+x,1-x)}$ described above. Schlesinger's plausible conjecture is that Gauss's recognition of the great generality of the hypergeometric equation caused him to change the direction of his research.

What finally does this comparison of the work Gauss with that of Abel and Jacobi establish? Most simply, that there was a way. Gauss and Abel were proceeding in the same direction, as far as the theory of functions was concerned. Secondly, that Abel's theory, as also Jacobi's, could be criticised for lacking a good way of writing the new functions (such as theta series were to provide), a rigorous theory of convergence, and an explanation of double periodicity. Thirdly, that the theory of the modular function $k = k(K/K')$ and the connection with differential equations (such as Legendre's) would be a fruitful area to explore, whether or not it was also connected to the theory of transformations and modular equations. Those three points, so to speak, come from Gauss's side. Our comparison also establishes that the theory of transformations is novel, and better developed by Abel than Gauss (if not in his *Recherches* then in his *Précis*) and that it should or could perhaps be emancipated from the theory of the functions themselves (a task that was to grow in importance until finally it was solved, essentially by Dedekind, in 1877 (see Sect. 8.2.2). These observations were all to be made by nineteenth-century mathematicians, and in later chapters we shall see when and to what effect.

1.5.3 Gauss on Complex Numbers and Complex Functions

In the course of all this highly innovative and focused work Gauss naturally considered what he should or could say about complex numbers and complex functions in general. For much of the eighteenth century mathematicians had thought that there was a hierarchy of different kinds of imaginary quantities, something that Gauss (1799, 14) called "veritable shadow of shadows" (*vera umbrae umbra*). It was only gradually accepted that "generally all imaginary quantities, no matter how complicated they might be, are always reducible to the form $M + N\sqrt{-1}$" where M and N are real numbers, as Euler expressed it in the concluding lines

[65] See Gauss's unpublished notes of 1809, *Werke* 10.1, 343 for a mention of this special case.

of his paper (1749) on the fundamental theorem of algebra. No one before 1800 considered that complex variables might enter into equations for loci, although it was generally understood that curves might sustain Bezout's theorem if one allowed that their points on intersection might be complex. This device, however, was not appreciated as a fact of geometrical significance inviting further exploration, but as an accountant's trick to balance the books.

In these respects Gauss differed little from his predecessors, but from his earliest days as a research mathematician he was more willing to promote the geometrical point of view. This is readily apparent in his remarkable discovery that the regular 17-sided polygon is constructible by ruler and compass alone. Remarkable not because it adds to the limited number of figures that are constructible in this way but because of the way it foreshadows elements of Galois theory and because it resolutely presents the non-trivial complex roots of the equation $z^{17} - 1 = 0$ as points in the plane. When the discovery was described anew in his *Disquisitiones arithmeticae* the roots are written out in their real and imaginary parts—each to ten places of decimals.

In the *Disquisitiones arithmeticae* Gauss spoke, as Euler had done, of imaginary quantities, and (see Sect. 337) following Euler used "i for the imaginary quantity $\sqrt{-1}$". He still spoke of imaginary quantities when he gave his first proof of the fundamental theorem of algebra (see below Sect. 1.5.4) in 1799, and he used that term in his work on the hypergeometric series. But in 1831 there was an interesting shift of emphasis. In the second part of his study of biquadratic residues (1832), he argued that number theory is revealed in its "entire simplicity and natural beauty" (Sect. 30) when the field of arithmetic is extended to the imaginary numbers. He explained that this meant admitting numbers of the form $a+bi$. "Such numbers", he said, "will be called complex integers". More precisely, he went on in the next section, the domain of complex numbers $a+bi$ contains the real numbers, for which $b=0$ and the imaginary numbers, for which b is not zero. Then, in Sect. 32, he set out the arithmetical rules for dealing with complex numbers. We read this as a step away from the idea that i is to be understood or explained as some kind of a square root, and towards the idea that it is some kind of formal expression to be understood more algebraically.

In a report on his own work for 1831 Gauss (1831, 177) commented that "The demonstrability of the intuitive meaning of $\sqrt{-1}$ can now be completely established and more is not required for this quantity to be used in the domain of arithmetic". Then in 1850 Gauss expressed the opinion that "imaginary quantities are not so much naturalised in mathematics as tolerated, and we are far from putting them on the same level as the real numbers. But there is no longer any ground for such an insult once the metaphysics of such quantities is put in its true light and it is shown that these have their own representational meaning, just as good as the negatives".[66] Sadly, as Stäckel (1917, 66–67) went on to note, what we do not have is Gauss's true light on the issue, one that would take us from the issue of the representation

[66] See Gauss (*Werke* 10.1, 404–405).

of complex numbers to their metaphysical nature. We can perhaps feel that what Gauss was contemplating was enlarging, if not indeed abandoning, the prevailing metaphysics of mathematics in this regard.

Instead, we can return to Gauss's ideas about the value of going complex, which he expressed in letters to his close friend Bessel in the 1810s. This also sheds light on Gauss's ideas about the meaning of a complex integral. In this connection, Schlesinger (1912, 152–153) noted that Gauss's discussion of the hypergeometric series, although inspired by Euler's work on it, did not follow Euler into a theory of its representation as an integral. He connected this with Gauss's earlier refusal to invert the general elliptic integral with complex modulus, and put it down to an awareness that a complex integral may well define a many-valued function of its upper end-point.

Gauss's often-cited letter to Bessel of 18 December 1811 sets out what was then a radical position.[67]

> Right away, if somebody wishes to introduce a new function into analysis, I will ask him to make clear if he simply wishes to use it for real quantities (real values of the argument of the function), and at the same time will regard the imaginary values of the argument as an appendage [Gauss here spoke of a ganglion, Überbein], or if he accedes to my principle that in the domain of quantities the imaginary $a+b\sqrt{-1}=a+bi$ must be regarded as enjoying equal rights with the real. This is not a matter of utility, rather to me analysis is an independent science which, by slighting each imaginary quantity, loses exceptionally in beauty and roundness, and in a moment all truths that otherwise would hold generally, must necessarily suffer highly tiresome restrictions.

After suggesting that he thought Bessel already agreed with him on this point, he turned to the question of the value of a complex integral and wrote:

> The integral $\int \varphi x.dx$ will always have the same value along two different paths if it is never the case that $\varphi\mathrm{x}=\infty$ in the space between the curves representing the paths. This is a beautiful theorem whose not-too-difficult proof I will give at a suitable opportunity [...]. In any case this makes it immediately clear why a function arising from an integral $\int \varphi x.dx$ can have many values for a single value of x, for one can go round a point where $\varphi x=\infty$ either not at all, or once, or several times. For example, if one defines $\log x$ by $\int \frac{dx}{x}$, starting from $x = 1$, one comes to $\log x$ either without enclosing the point $x=0$ or by going around it once or several times; each time the constant $+2\pi i$ or $-2\pi i$ enters; so the multiple of logarithms of any number are quite clear.

Although Gauss never found the suitable opportunity, it is clear that he had obtained the first crucial insight into the integration of functions of a complex variable. Much remained to be done, and the credit for discovering and, no less importantly, publishing it, belongs with Cauchy, as we discuss in Chap. 3. Gauss's insight marks a decisive break with a purely formal theory of functions of a complex variable and a turn towards a more sophisticated, geometrical theory. Perhaps unfortunately, as we shall see in the subsequent chapters, this geometric theory had to wait some decades before being developed. Gauss did, however, publish some

[67] For the whole letter see *Werke* 10.1, 365–371, these quotations come from pp. 366–367.

hints about the utility of the geometric interpretation of complex numbers in his proofs of the fundamental theorem of algebra.

1.5.4 Gauss's Proofs of the Fundamental Theorem of Algebra

In this regard, the evolution of Gauss's proofs of the fundamental theorem of algebra is interesting. The first cryptic hint at a proof can be found in an entry of his diary on October 1797 (nr. 80): "Aequationes habere radices imaginarias methodo genuina demonstratum" (Proved with an appropriate method that the equations have imaginary roots). In the course of his life Gauss gave four proofs of this result, of which the first, third and fourth employ complex methods. The first was the subject of his doctoral thesis of 1799, where he devoted more than half of its pages to criticisms of earlier attempts. As Gauss noted, serious investigations of the general polynomial equation of arbitrary degree and with real coefficients seem to have begun with Euler, who wrote to Nicholas I Bernoulli on 1 October, 1742 to suggest that such equations have as many complex roots as their degree indicates. Bernoulli was not convinced and offered a flawed counter-example. Later that year, Euler wrote to Goldbach offering a proof valid for degrees 5 and 6. This time Goldbach was unconvinced, because the restriction on the degree seemed unnatural to him. After these attempts d'Alembert produced a long paper in 1746 in which he tried to prove the general theorem, then Euler tried again in 1749, only to find his argument criticised in 1759 by a student of Lagrange, Daviet François de Foncenex, and then by Lagrange himself in 1772, and their essays spawned criticisms and improvements in turn.[68]

Gauss's first criticism of d'Alembert's argument was that d'Alembert had assumed the existence of the roots and showed only that if they exist then they have to be complex. Gauss argued that d'Alembert should first have proved the roots actually exist and that he had not considered the possibility that the roots did exist but could not be manipulated like numbers; or indeed that the roots did not even exist. That said, Gauss agreed that this problem could be overcome. Gauss also criticised d'Alembert's use of infinite series in his proof and showed by means of an example that it was unsound, again politely admitting that it was perhaps capable of being re-cast in a more reliable form. This criticism notwithstanding, Gauss admitted that "the true strength of the proof seems to me not to have been weakened at all by all the objections", and he even stated that one could build a rigorous proof on that foundation (in *Werke* 3, 11).

Gauss then turned to the second of Euler's arguments, which reduced the problem to the factorisation of polynomials of degree a power of 2, it being generally agreed

[68]For an account of eighteenth century attempts on the FTA including an evaluation of d'Alembert's proof and Gauss' criticism of it, see Gilain (1991) and his introduction in d'Alembert (2007).

that polynomial equations of odd degree must have a real root. If factorisation can be assured, then the fundamental theorem of algebra follows immediately by induction. Gauss observed that not only did Euler's approach tacitly assume that polynomial equations have roots, the proof of factorisation replaced the original equation with a system of quadratic equations for which there is no guarantee that solutions exist, for it produces $2m-1$ equations in $2m-2$ unknowns. Gauss then commented that (Lagrange, 1774) had thoroughly resolved some of the objections to Euler's argument, but gaps in the proof remained, notably the assumptions that polynomial equations have roots and that the only problem is to show that they are complex numbers.[69]

Having thus swept aside all known proofs of the theorem, and considerably raised the standards by which attempts on the fundamental theorem of algebra are to be judged, Gauss then offered his own proof. He took a polynomial $f(z) = z^m + a_1 z^{m-1} + \cdots + a_m$ with real coefficients, and although he had promised in Sect. 2 of his essay not to use imaginary quantities he looked for its roots in an infinite plane whose points are specified by polar coordinates r, ϕ. This is to use complex numbers in polar form without saying so. Such roots will be the common points of the equation

$$\begin{aligned} U &= r^m \cos m\phi + a_1 r^{m-1} \cos(m-1)\phi + \cdots + a_{m-1} r \cos\phi + a_m = 0, \\ T &= r^m \sin m\phi + a_1 r^{m-1} \sin(m-1)\phi + \cdots + a_{m-1} r \sin\phi = 0. \end{aligned}$$

These are the curves defined by the real and imaginary parts of the equation $f(z) = 0$: $\mathrm{Re} f(z) = 0$ and $\mathrm{Im} f(z) = 0$, as z varies in the complex plane. Gauss now argued that outside a suitably large circle of radius R centred on the origin each of these curves meets a concentric circle of radius $r > R$ in two disjoint sets of $2m$ distinct points, so these curves consist of $2m$ arcs going off to infinity in the plane. Moreover, when R is suitably large the z^m is so dominant that the curves $\mathrm{Re} f(z) = 0$ and $\mathrm{Im} f(z) = 0$ meet the circle of radius R alternately. Gauss now set to work to show that these curves are each made up of m disjoint "parabola-shaped" pieces (our term, not Gauss's), and so they join up inside the circle and can only do so if a curve $\mathrm{Re} f(z) = 0$ crosses a curve $\mathrm{Im} f(z) = 0$ (a continuity argument was here tacitly assumed by Gauss). At such a crossing point, the equation has a root, and so the fundamental theorem of algebra is proved.

To that effect, Gauss argued that the curves are real algebraic curves, so they consist of m pieces that, as it were, come from and go to infinity; they cannot stop, break apart, or spiral to a point in the fashion of some transcendental curves (Gauss gave the example of $y = 1/\log(x)$). As he put it, if an algebraic curve enters a bounded region of the plane, it also leaves it (Sect. 21, footnote). Nobody had

[69]Remmert has aptly remarked that "the Gaussian objection against the attempts of Euler–Lagrange was invalidated as soon as Algebra was able to guarantee the existence of a splitting field for every polynomial" (in Ebbinghaus et al. 1990, 105). Remmert went on to point out that this had been already observed by Kneser (1888, 21).

raised doubts about this so far, Gauss added, but if someone requested he was "ready to provide a proof free of all doubts on another occasion". Here he limited himself to state that his argument was based "on the principles of the geometry of position (*geometriae situs*) that are no less valid than the principles of geometry of magnitudes" (Gauss, 1799, 28). Apparently, the promised occasion to offer a rigorous proof never occurred. Gauss's statement is true, but its proof requires topological arguments that are surely no easier to prove than the fundamental theorem of algebra itself, and to that extent Gauss's proof is also defective. On the other hand, his argument was systematic in that it dealt with polynomial equations of any degree. Moreover, the topological nature of Gauss's proof is attractive, and that was what Gauss saw as the heart of the matter. It was a remarkable insight for 1799, even if neither Gauss nor any one else could have provided a clear account of the completeness of the real numbers and a proper distinction between the real numbers and the rational numbers.

We pass over Gauss's second proof (dated late 1815), which does not illuminate his understanding of complex variables, and turn to his third proof (dated January 1816), which in many ways returns to the ideas of his first. On this occasion, Gauss wanted to show that there is a root of the polynomial $f(z)$ within a circle of suitably large radius R by considering the quantity the square of the modulus of $f(z)$, which is t^2+u^2, where $t = \mathrm{Re}f(z)$ and $u = \mathrm{Im}f(z)$. He introduced the derivatives t', u', t'', u'' and y, a rational function of these, and on the assumption that t and u do not simultaneously vanish, he evaluated the double integral $\int\int y dr d\varphi$ in two ways over the region $0 \leq r \leq R,\ 0 \leq \varphi \leq 2\pi$. Integrating first with respect to φ taken round a circle of constant radius r yields $\int y d\varphi = \frac{tu'-t'u}{r(t^2+u^2)}$, which is zero over the interval $0 \leq \varphi \leq 2\pi$, and therefore the double integral is zero. However, integrating first with respect to r as r goes from 0 to R leads to the integral $\int y dr = \frac{tt'+uu'}{t^2+u^2}$ and this, Gauss showed, implies that the double integral is a non-zero positive quantity. The only way out of this apparent contradiction was to deny its fundamental assumption that the integral may be evaluated in both orders and give the same result, and therefore Gauss concluded that there were points where $t^2+u^2=0$, and so t and u simultaneously vanish, and so the polynomial has a root. The proof therefore rests on the insight that when a double integral is replaced by a repeated integral the order of integration may matter when the integrand becomes infinite.

Gauss presented his fourth and last proof at a meeting of the Göttingen Gesellschaft der Wissenschaften on 16 July 1849 and published it the following year (Gauss, 1850). It was produced as a *Jubiläumschrift* on the occasion of the 50th anniversary of his first proof, which was marked by a celebration of Gauss's distinguished career. There Gauss declined to repeat his criticisms of the eighteenth century proofs. He limited himself to remarking that he had two aims in his 1799 paper: first, to show that all the proofs produced hitherto were "unsatisfactory and illusory", and second to produce a "new, completely rigorous proof" (Gauss, 1850, 73). After noting that Cauchy had given a new proof more recently, he explained that the form of the statement of the theorem he had given in 1799—any polynomial with real coefficients can be split into real factors of the first and the second degree—had been chosen by him deliberately in order "to avoid any intervention of imaginary quantities. Nowadays, since the concept of complex quantities is familiar to everybody, it seems appropriate to give that form up", and to state the theorem as the splitting of any polynomial with complex coefficients into linear factors over $\mathbb{C}$.[70] In this setting Gauss produced a proof by going back to the method he had followed in his first proof. Instead of the systems of curves $\mathrm{Re} f(z) = 0$ and $\mathrm{Im} f(z) = 0$, however, this time he considered the connected regions lying between these curves, where the functions $\mathrm{Re} f(z)$ and $\mathrm{Im} f(z)$ have constant sign. Thus, the consideration of only one system of curves, $\mathrm{Re} f(z) = 0$ say, was enough to conclude the proof. This did not prevent him from discussing the geometrical properties of the curves, and in this respect the proof was as flawed as his first one in 1799. However, as Ostrowski (1920) has shown, in the light of subsequent developments in the theory of conformal mapping and point set topology Gauss's idea of considering regions instead of their limiting Jordan curves was revealed to be a step of essential importance. In fact, starting from Gauss's regions one could produce a rigorous proof of the theorem, which makes the consideration of the curves within the circle $|z| = R$ completely superfluous. This, Ostrowski (1920, 9) commented, provided an unexpected confirmation of Gauss's cryptic statement that "Essentially, the proper content of the whole argument belongs to a higher field of general, abstract theory of magnitudes" that is independent of space, and whose object are "operations on connected quantities according to continuity" (Gauss, 1850, 79). With respect to his 1799 proof, this time Gauss dealt more carefully with the case of multiple roots and showed that all roots occur in the way described.[71] Eventually, he devoted the second part of the paper to the numerical evaluation of the roots of algebraic equations.

We could say in conclusions that Gauss's proofs, especially the third, could have provoked further reflections about complex functions and their integrals, but hints, like glances across a room, are not always noticed.

1.6 Elliptic Functions, Complex Functions

The simultaneous discovery of elliptic functions by Abel and Jacobi in the late 1820s transformed Legendre's already rich theory of elliptic integrals into a new branch of

[70]Remmert, perhaps sticking too closely to Gauss's words, commented on this that "until 1849 all proofs, including those found in the intervening period by Cauchy, Abel, Jacobi and others, dealt with real polynomials only. It was only in his fourth proof, which is a variant of the first, that Gauss in 1849, the time now being ripe for this step, allowed arbitrary complex polynomials" (in Ebbinghaus et al. 1990, 108). As we will see in Sect. 2.5, however, contrary to this claim the theorem was stated and proved for complex polynomials already by Argand (1806, and 1814–1815) not to mention Cauchy (1821a).

[71]For a modern account of Gauss's first and fourth proof, see Ostrowski (1920).

mathematics. By inverting the integrals and letting the variables become complex, they discovered numerous fascinating results that not only generated a new theory but also had implications for the solution of equations, number theory, and geometry. One can almost imagine that this dazzling display diverted mathematicians from two odd features of this work: that insofar as it was complex it was so in an entirely formal way; and that while some variables had gone complex the modulus was, for no compelling reason, still confined to being real. These oddities reflect a degree of hesitation among the mathematicians at the time over complex numbers. As Chap. 2 will describe in some detail, complex numbers and complex variables appeared in the mathematics of the eighteenth and early nineteenth centuries as artifices, aids to calculation that should not remain in the final answers. In the absence of a developed geometric theory, the only way to handle them was formally, and this formality was a comfortable place to rest for as long as mathematics was somehow essentially about real quantities.

The first great challenge of the work of Abel and Jacobi on elliptic functions was therefore that the new functions were *essentially* complex. Their complex character could not be asked to disappear without the new theory collapsing entirely. The functions had complex periods, their division properties only made sense if they took complex values. This is why elliptic functions were to be such a stimulus to the creation of complex function theory: it was quite clear that a proper elucidation of their theory would require a theory of complex functions.

The obstacles mathematicians faced in creating such a theory were many. As we shall see in Chap. 4, because the square root in the integrand of an elliptic integral made its value ambiguous—grievously so when the variables were complex—many mathematicians decided the situation was hopeless and the theory of elliptic functions would have to be re-created from scratch in some other way. There was no insight into the way a theory of functions (of some sort) from $\mathbb{C}$ to $\mathbb{C}$ could be carved out from a more general theory of maps from $\mathbb{R}^2$ to $\mathbb{R}^2$, and indeed Cauchy's own thorough delineation of a theory of continuity, integrability, and differentiability for real functions was itself a creation of the 1820s. And then, by historical accident, the man who took up the challenge of creating a theory of complex functions, Cauchy, had almost no interest in elliptic function theory, so, as we will see in Chap. 3, his many contributions were not intended to help in that area. Others had to step forward to rescue the foundations of elliptic function theory.

There could have been another history, but the man who saw further than anyone into the new domain chose to publish nothing about it. Gauss's interest in elliptic functions began as early as 1796, when he was 19, and from a deepening study of the lemniscatic integral he drew out more and more of a general theory. He defined the lemniscatic sine and cosine as complex functions of a complex variable, generalised their theory to elliptic functions with an arbitrary real modulus, discovered the connection to the hypergeometric equation, found the corresponding theta functions, and finally, perhaps in the 1820s, embarked on a general theory of elliptic functions with arbitrary periods (corresponding to a real or complex modulus). In so doing he gained insights into the distinction between a complex power series and a complex function, analytic continuation, and the integration of a complex function round a closed path. Unlike the rich and connected account he gave himself of elliptic functions these insights into general complex functions remained scattered, but they were the first fruits that the geometric theory of complex variables was to bear. For whatever reason, all this remained hidden from contemporary view except for slender hints left in his *Disquisitiones arithmeticae* and in his proofs of the fundamental theorem of algebra.

Appendix: Transformations and Complex Multiplication

We cannot do better than follow Houzel's lead in (Houzel 1978, Sects. 8, 13) and explain in modern terms what is going on. An elliptic function, f, has two distinct periods w_1 and w_2, say, which may both be complex and which are linearly independent over $\mathbb{R}$ (i.e. $w_1/w_2 \notin \mathbb{R}$). So :

$$f(z+w_1) = f(z) = f(z+w_2),$$

whence

$$f(z+m_1w_1+m_2w_2) = f(z)$$

for all integers m_1 and m_2. The set $\Lambda = \{m_1w_1+m_2w_2 \,|\, m_1,m_2 \in \mathbb{Z}\}$ of points in $\mathbb{C}$ is called a lattice, and the quotient space $\mathbb{C}/\Lambda$ is a torus. Since f is periodic it defines a function, also denoted f, on the quotient space, by

$$f(z \bmod \Lambda) = f(z + \text{ any element of} \Lambda) = f(z).$$

When these functions arose initially the torus was not considered, and the periods appeared as certain complete integrals. A transformation relates one elliptic function with its periods to another function with different periods, and since the periods in the classical approach were well known to be functions of the moduli, questions about periods became questions about the corresponding moduli too. The modern approach looks first at the periods, and asks, given lattices Λ and Λ', what holomorphic functions are there $h : \mathbb{C}/\Lambda \to \mathbb{C}/\Lambda'$?

One can lift any such function to $\tilde{h} : \mathbb{C} \to \mathbb{C}$ for which $\tilde{h}(z+w) - \tilde{h}(z) \in \Lambda'$ for all $w \in \Lambda$ and for all $z \in \mathbb{C}$. So $\tilde{h}(z+w) - \tilde{h}(z) \in \Lambda'$ is independent of z, since Λ' is discrete and $\mathbb{C}$ is connected. So $\frac{d\tilde{h}}{dz}(z+w) = \frac{d\tilde{h}}{dz}(z)$ and $\frac{d\tilde{h}}{dz}(z)$ is periodic with period lattice Λ. But it is holomorphic, and now, being periodic, it is bounded, so by Liouville's theorem it is constant. So $\frac{d\tilde{h}}{dz}(z) = a$, $\tilde{h}(z) = az+b$, and by a translation of the coordinates it is enough to consider $\tilde{h}(z) = az$. To avoid triviality, suppose $a \neq 0$, and look at the sub-lattice $\tilde{h}(\Lambda)$ of Λ'. We have $\tilde{h}(\Lambda) = a\Lambda$, so $a\Lambda$ is a sub-lattice of Λ', or, if you prefer, Λ is a sub-lattice of $\frac{1}{a}\Lambda'$. Lattices are abelian groups (abstractly they are just $\mathbb{Z} \oplus \mathbb{Z}$) so we can say Λ is a subgroup of $\frac{1}{a}\Lambda'$ of finite index, denoted n (where $n = a^2$). If we choose w'_1 and w'_2 as generators of Λ' then

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix},$$

where α, β, γ, and δ are integers and the determinant $\alpha\delta - \beta\gamma = n$. A suitable choice of bases in domain and codomain can always be found to diagonalise this matrix, when it becomes $\begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}$ where $d_1 d_2 = n$, and conversely from such a matrix it is easy to obtain a sub-lattice of Λ' of index n. Moreover, each divisor d_1 of n gives rise to d_1 distinct sub-lattices, so there are as many sub-lattices as the sum of the divisors of n, i.e. $\sigma_1(n) = \sum_{d|n} d$.

In classical language, since each sub-lattice corresponds to a transformation, there are $\sigma_1(n)$ transformations of order n. In particular, there are $p+1$ transformations of order a prime p. The set of transformations always includes the divisions that arise by multiplying one period by n and leaving the other period fixed, but other transformations also exist.

We would expect from the modern point of view that if one looks at the ratio $\frac{w_1}{w_2} = \tau$, it must be mapped to itself, so $\tau = \frac{\gamma\tau+\delta}{\alpha\tau+\beta}$, whence $\tau = \frac{\delta-\alpha\pm\sqrt{(\alpha+\delta)^2-4}}{2\beta}$ and $\alpha = \delta = a^2$. So occasionally lattices admit extra symmetries, which they do when τ takes particular complex values. This is the phenomenon of complex multiplication.

Springer

http://www.springer.com/978-1-4614-5724-4

Hidden Harmony—Geometric Fantasies
The Rise of Complex Function Theory
Bottazzini, U.; Gray, J.
2013, XVII, 848 p. 38 illus., 2 illus. in color., Hardcover
ISBN: 978-1-4614-5724-4