VIII.2

Gauss Sums

SAMUEL JAMES PATTERSON

The subject of this chapter is the determination of the sign of (quadratic) Gauss sums. The sums considered by Gauss are of the form

$$\sum_{j=0}^{n-1} \exp\Big(2\pi \mathrm{i} \frac{j^2}{n}\Big).$$

This determination of the sign, or rather of the argument, of this sum represents one of Gauss's most remarkable achievements. He discovered the underlying phenomenon experimentally while he was completing his work on the *Disquisitiones Arithmeticae*, shortly before it was finally published. He only discovered a formal proof some years later. This theorem and its ramifications have continued to fascinate over the last two hundred years – partly because of its intrinsic beauty and significance and partly because it goes considerably beyond what one should expect from class-field theory and the general theory of cyclotomic fields. There have been many very different proofs and many generalizations. Here we shall be concerned mainly with the proofs; a comprehensive survey of the generalizations would be too great a task to undertake within the confines of the present chapter. It should be remarked at this point that if n is square-free and odd then

$$\sum_{j=0}^{n-1} \exp\Big(2\pi i \frac{j^2}{n}\Big) = \sum_{x \pmod n} \Big(\frac{x}{n}\Big) \exp\big(2\pi \mathrm{i} \frac{x}{n}\big)$$

and therefore sums of the form $\sum_{x \pmod n} \chi(x) \exp\big(2\pi \mathrm{i} \frac{x}{n}\big)$, where χ is a Dirichlet character to the modulus n, are also often called Gauss sums, as are their local and finite-field analogues. This multiple use of the designation has led to some confusion.

1. Gauss and Gauss Sums

The determination of the sign of the Gauss sums, in the case of a prime modulus, is stated – with remarkable confidence – by Carl Friedrich Gauss at the end of art. 356 of the *Disquisitiones Arithmeticae*. We know from his mathematical diary that he discovered this result in the middle of May, 1801 and stated it there for general moduli. Specifically Gauss wrote:

> A fifth method of proving the fundamental theorem has presented itself thanks to a most elegant theorem from the division of the circle, namely[1]
>
> $$\begin{array}{lccccc|c|c|c|l} \sum & \left.\begin{array}{c}\sin\\ \cos\end{array}\right\} & \frac{nn}{a}\mathcal{P} & = & & \begin{array}{c}+\sqrt{a}\\ +\sqrt{a}\end{array} & \begin{array}{c}0\\ +\sqrt{a}\end{array} & \begin{array}{c}0\\ 0\end{array} & \begin{array}{c}+\sqrt{a}\\ 0\end{array} & \\ \text{according as} & & a & \equiv & & 0 & 1 & 2 & 3 & \pmod{4} \end{array}$$
>
> where for n are to be substituted all numbers from 0 to $a-1$.

Note that Gauss saw this theorem in the first place as another (fifth) proof of the law of quadratic reciprocity. Again, through the diary we know that he found the first *proof* of his 1801 entry only on August 30, 1805. Specifically:

> The proof of the most beautiful theorem mentioned above, May 1801, which we had been seeking for 4 years and more with all efforts, we have at last completed. *Comment[ationes] rec[entiores], I* [2]

The reference is to [Gauss 1811]; this only appeared in 1811 but it is not clear when it was written. Presumably this reference (and the underlining of this entry) were added later. Three entries of the diary intervening between the two entries quoted above deal with computational astronomy and entry 122 explains (for whom?) that the years 1802 to 1804 were spent doing astronomical calculations. Gauss wrote to Wilhelm Olbers about his proof on September 3, 1805. Let us quote the first half of this letter in full :

> I hope that it is only your too many duties, and not sickness or anything else unplesant, which account for the fact that I have not been made happy by a letter from you in such a long time. My recent occupations were also not such that they would have provided anything particular to communicate to the geometer, nor did events provide anything of interest for the sympathetic friend. Through various circumstances – partly through several letters from Le Blanc in Paris who studies my *Disquisitiones Arithmeticae* with true passion, has completely familiarised himself with them, and shared quite a few nice comments about them with me; partly because of the presence of a friend who is also studying that work and often asks me for advice – and partly also because of a sort of tedium, or at least fatigue from the dead, mechanical

1. [Gauss 1796–1814], entry 118; our translation of: *Methodus quinta theorema fundamentale demonstrandi se obtulit, adiumento theorematis elegantissimi theoriae sectionis circuli, puta ...*
2. [Gauss 1796–1814], entry 123; our translation of: *Demonstratio theorematis venustissimi supra 1801 Mai. commemorati quam per 4 annos et ultra omni contentione quaesiveramus, tandem perfecimus. Comment rec. I.*

> computations, I have been seduced to take a break from it for once and take up again my beloved arithmetical investigations. You may recall from our conversations in Bremen, in particular on that beautiful afternoon which we spent on the *Vahr*, that I have had for some time already a fair number of investigations, if not in my drawer, at least in my head, which would provide sufficient material for a second volume of the *Disquisitiones Arithmeticae*, and which – at least according to my own judgement – are just as remarkable as those contained in the first volume. But you may also recall my complaints about a theorem, which is partly interesting in itself, and partly serves as a foundation or keystone for a substantial part of those investigations, and which I have known for more than two years, but which confouded all my attempts to find an adequate proof. This theorem is already hinted at in the *Disquisitiones Arithmeticae*, p. 636,[3] or more precisely, only a special case of it, namely the one where n is a prime number, to which the others could be reduced. What is written there between *Quaecunque igitur radix etc.* and *valde sunt memoribilia*, is rigorously proved there, but what follows, i.e., the determination of the sign, is exactly what has tortured me all the time. This shortcoming spoiled everything else that I found; and hardly a week passed during the last four years where I have not made this or that vain attempt to untie that knot – especially vigorously during recent times. But all this brooding and searching was in vain, sadly I had to put the pen down again. Finally, a few days ago, it has been achieved - but not by my cumbersome search, rather through God's good grace, I am tempted to say. As the lightning strikes the riddle was solved; I myself would be unable to point to a guiding thread between what I knew before, what I had used in my last attempts, and what made it work. Curiously enough the solution now appears to me to be easier than many other things that have not detained me as many days as this one years, and surely noone whom I will once explain the material will get an idea of the tight spot into which this problem had locked me for so long. Now I cannot resist to occupy myself with writing up and elaborating on this material. However, my astronomical work should not be completely neglected all the same.[4]

Gauss's literary skill in this letter is abundantly clear. I would like to make a few comments about it:

Concerning the readers of the *Disquisitiones Arithmeticae* mentioned in the letter, the "Le Blanc" to whom Gauss refers was Sophie Germain. For more details about the exchange of letters between Gauss and Sophie Germain see [Leibrock 2001]. The "friend" in Braunschweig remains unclear. Although his identification is not that important it is of some interest as the number of people who studied the D.A. in detail in the years immediately following its publication – especially in Germany – was small.

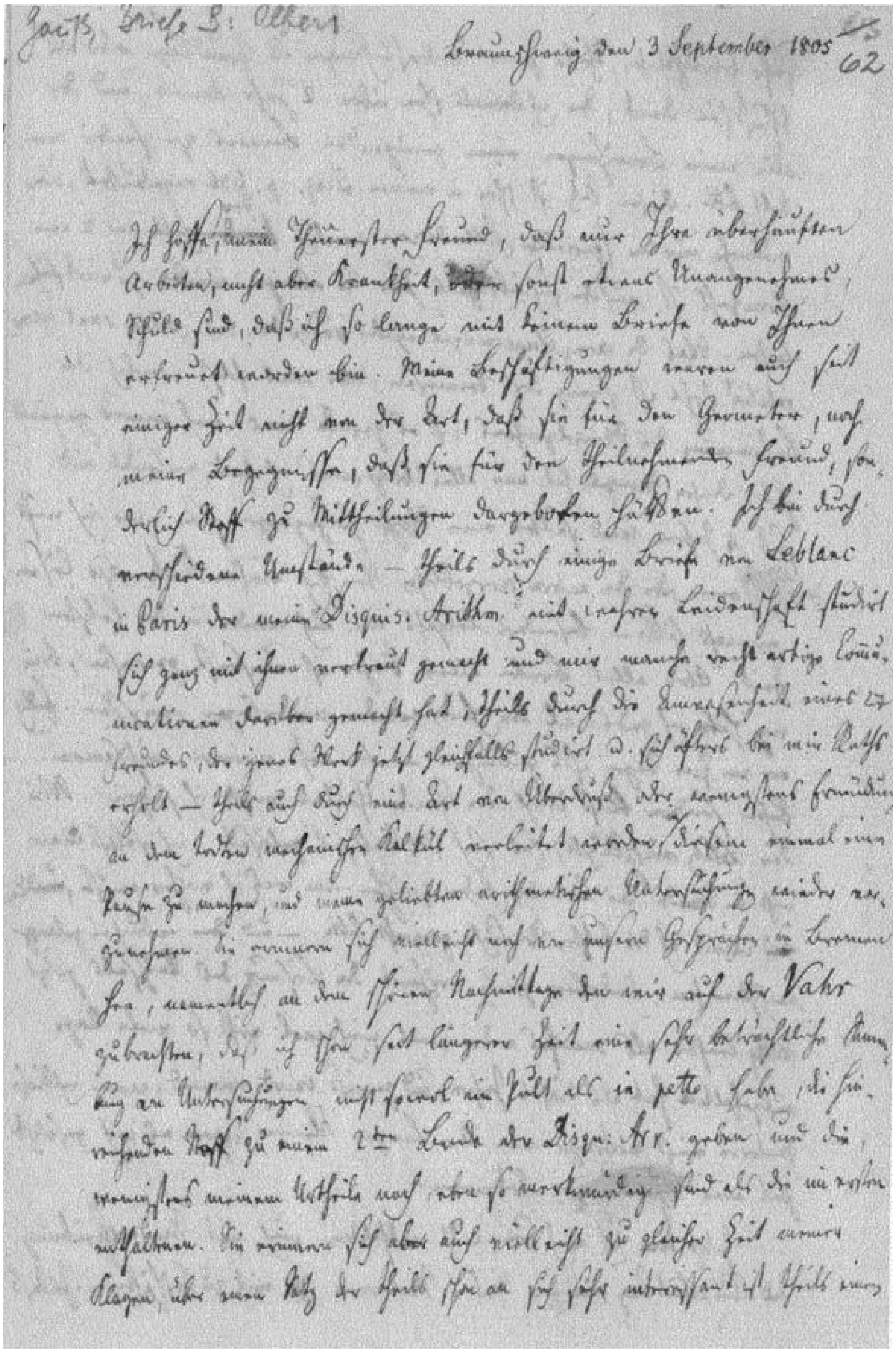
(*The main text continues on page 511.)*

3. Towards the end of art. 356.
4. Our translation. The original of this part of the letter is reproduced and transcribed on the following pages. The whole letter is published in [Gauss & Olbers 1900–1909], vol. 1, *Brief* 133, pp. 267–270. After the passage quoted here, Gauss goes over to an astronomical theme, which he also deals with in a letter to Bessel with the same date, and which he asks Olbers to give to Bessel.

Transcription of the first half of Gauss's letter to Olbers, September 3, 1805

Ich hoffe, mein theuerster Freund, daß nur Ihre überhäuften Arbeiten, nicht aber Krankheit, oder sonst etwas Unangenehmes, Schuld sind, daß ich so lange mit keinem Briefe von Ihnen erfreut worden bin. Meine Beschäftigungen waren auch seit einiger Zeit nicht von der Art, daß sie für den Geometer, noch meine Begegnisse, daß sie für den theilnehmenden Freund, sonderlich Stoff zu Mittheilungen dargeboten hätten. Ich bin durch verschiedene Umstände – theils durch einige Briefe von Le Blanc in Paris, der meine Disq. Arithm. mit wahrer Leidenschaft studirt, sich ganz mit ihnen vertraut gemacht und mir manche recht artige Communicationen darüber gemacht hat, theils durch die Anwesenheit eines Freundes, der jenes Werk jetzt gleichfalls studirt u[nd] sich öfters bei mir Raths erholt – theils auch durch eine Art von Überdruß oder wenigstens Ermüdung an dem todten mechanischen Kalkül verleitet worden, in diesem einmal eine Pause zu machen, und meine geliebten arithmetischen Untersuchungen wieder vorzunehmen. Sie erinnern sich vielleicht noch von unsern Gesprächen in Bremen her, namentlich an dem schönen Nachmittage den wir auf der Vahr zubrachten, daß ich schon seit längerer Zeit eine sehr beträchtliche Sammlung von Untersuchungen nicht sowohl im Pult als in petto habe, die hinreichenden Stoff zu einem 2$^{\text{ten}}$ Bande der Disq. Arr. geben und die, wenigstens meinem Urtheile nach, eben so merkwürdig sind als die im ersten enthaltenen. Sie erinnern sich aber auch vielleicht zu gleicher Zeit meiner Klagen, über einen Satz der
1|2 *theils schon an sich sehr interessant ist, theils einem | sehr beträchtlichen Theile jener Untersuchungen als Grundlage oder als Schlußstein dient, den ich damals schon über 2 Jahr kannte, und der alle meine Bemühungen einen genügenden Beweis zu finden, vereitelt hatte. Dieser Satz ist schon in meinen Disq. p. 636 angedeutet, oder vielmehr nur ein specieller Fall davon, nemlich der wo n eine Primzahl ist, auf den sich übrigens hier die übrigen würden zurückführen lassen. Was da von "Quaecunque igitur radix etc." bis "valde sunt memoribilia" steht ist streng dort bewiesen, aber was folgt nemlich die Bestimmung des Wurzelzeichens, ist es gerade was mich immer gequält hat. Dieser Mangel hat mir alles übrige, was ich fand, verleidet und seit 4 Jahren wird selten eine Woche hingegangen sein, wo ich nicht [den] einen oder den andren vergeblichen Versuch diesen Knoten zu lösen gemacht hätte – besonders lebhaft nun auch wieder in der letzten Zeit. Aber alles Brüten alles Suchen ist umsonst gewesen, traurig habe ich jedesmal die Feder wieder niederlegen müssen. Endlich vor ein paar Tagen ists gelungen – aber nicht meinem mühsamen Suchen sondern bloß durch die Gnade Gottes möchte ich sagen. Wie der Blitz einschlägt, hat sich das Räthsel gelöset : ich selbst wäre nicht im Stande den leitenden Faden zwischen dem was ich vorher wußte, dem womit ich die letzten Versuche gemacht hatte – und dem wodurch es gelang nachzuweisen. Sonderbar genug erscheint die Lösung des Räthsels jetzt leichter als manches andere, was mich wohl nicht so viele Tage aufgehalten hat als dieses Jahre, und gewiß wird niemand, wenn ich die Materie einst vortrage von der langem Klemme, worin es mich gesetzt hat, eine Ahndung bekommen.*

Jetzt kann ich mich nun nicht enthalten, mich mit Niederschreibung und Ausarbeitung einiger dieser Materien mit zu beschäftigen. Indeß sollen meine astronomischen Arbeiten darüber nicht ganz vernachlässigt werden.



Gauß, Briefe B: Olbers

Braunschweig den 3 September 1805

62

Fig. VIII.2A. The first page of Gauss's letter to Olbers, front.
(Courtesy of NSUB Göttingen)

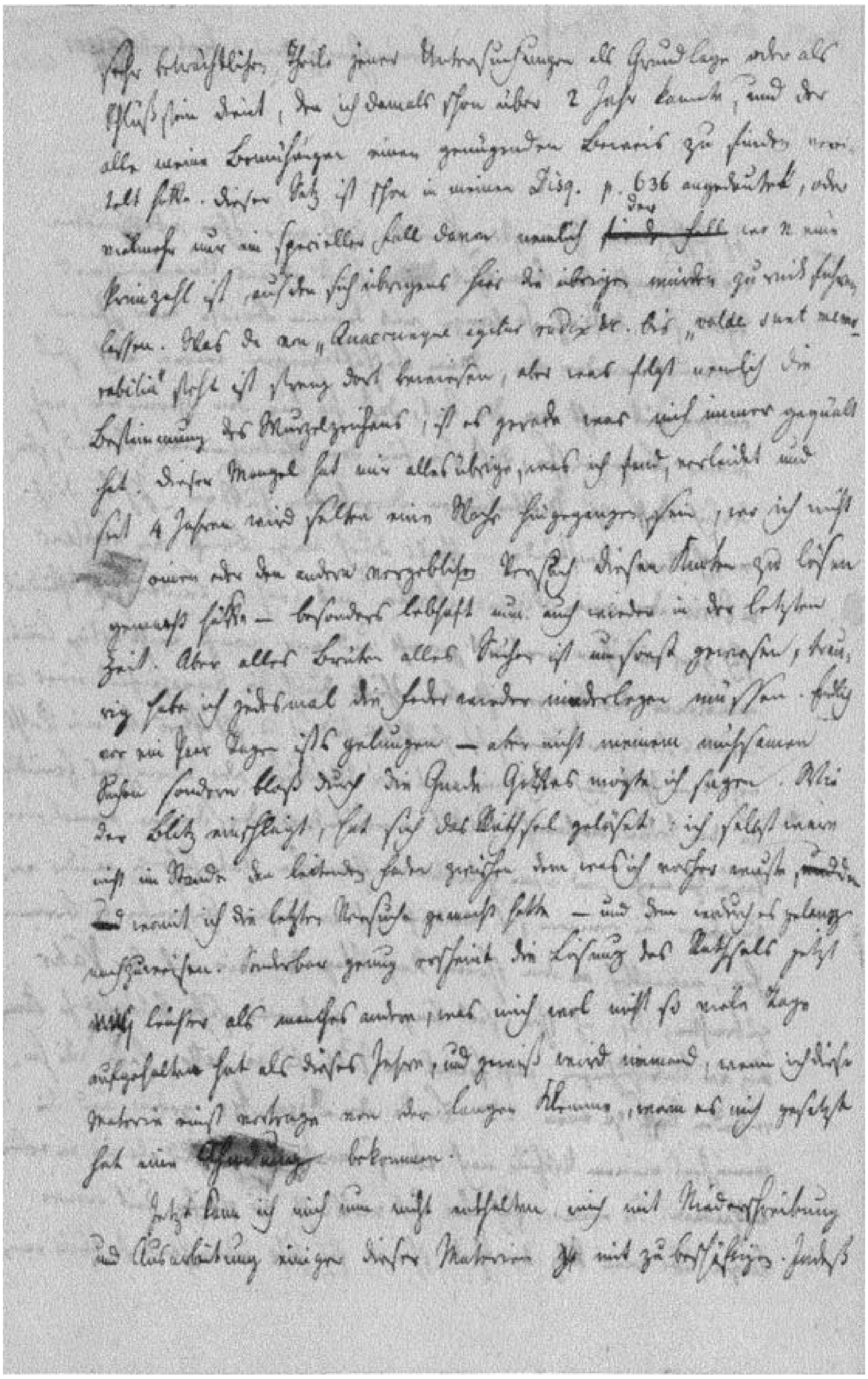


Fig. VIII.2B. The first page of Gauss's letter to Olbers, back.
(Courtesy of NSUB Göttingen)

The proof that Gauss found *was* easy; it is very much in the Euler tradition, especially Euler's work on theta functions.[5] Skill in this area is not common today, but it was for contemporaries of Gauss. More precisely, the proof is based on a q-analogue of the consequence of the binomial theorem $\sum_{j=0}^{n}(-1)^j \binom{n}{j} = 0$. For odd n, one has

$$(1-q)(1-q^3)\ldots(1-q^{n-2}) = 1 - \frac{(1-q^{n-1})}{(1-q)} + \frac{(1-q^{n-1})(1-q^{n-2})}{(1-q)(1-q^2)} - \cdots$$

which can be proved by a modification of the techniques used in working with combinatorial identities. From this it follows that if ζ is an n^{th} root of unity then, with $q = \zeta$, one obtains

$$\begin{aligned}(1-\zeta)(1-\zeta^3)\ldots(1-\zeta^{n-2}) &= 1 - \frac{(1-\zeta^{n-1})}{(1-\zeta)} + \frac{(1-\zeta^{n-1})(1-\zeta^{n-2})}{(1-\zeta)(1-\zeta^2)} \cdots \\ &= \sum_{j=0}^{n-1} \zeta^{-\frac{1}{2}j(j+1)}.\end{aligned}$$

If we take ζ to be $\exp(2\pi i/n)$, or better, $\exp(8\pi i/n)$, then the product is easily transformed into a multiple of a product of sines whose argument is easy to determine. From this Gauss's theorem follows easily.

The fact that the search for a proof *tortured* Gauss is perhaps a reflection that he had been so confident in the *Disquisitiones Arithmeticae* – he was honour-bound to find a proof, especially when others were making progress in studying the D.A. and could soon be asking questions. The part of the D.A. – Section 7 – which treats cyclotomy was studied with great interest, especially in France. The proof was, however, only published in 1811, and was, in fact, only taken up by others more than twenty years later.

Finally I would like to point out that this letter shows Gauss working on his image.[6] He naturally does so within the ethos of the time. One important element which marked those years was the aftermath of the sort of veneration of genius which had been a trademark of German literature and philosophy since the 1770s, but which underwent significant changes in the new postrevolutionary cultural and political reality. Some of these changes crystallized around the self-legitimizing heroic figure of Napoleon. Johann Wolfgang von Goethe, for instance, noted in his diary on August 8, 1806 that he had been musing about "new titles for Napoleon," about "subjective princes," and that he was able to interpret "Napoleon's deeds and practices" as a vindication of Fichte's theory.[7] The same kind of interpretation of Napoleon as a paradigm for autonomous, charismatic creativity underlied the original

5. See [Euler 1748], Cap. XVI, *De partitione numerorum*, [Euler 1783a] and [Euler 1783b].
6. Kurt R. Biermann has observed this in a different context in [Biermann 1991].
7. See [Schmidt 1985], p. 451.

dedication of Ludwig van Beethoven's Eroica Symphony to Bonaparte in 1804.[8]

But Napoleon's genius may not have been on Gauss's mind when he wrote the letter to Olbers.[9] Gauss's favourite German novelist, the bestselling Jean Paul,[10] had published between 1800 and 1803 his monumental novel *Titan* of which an important aspect is the balance of romantic genius and reality.[11] In his 1804 theoretical treatise "Vorschule der Ästhetik" he sees the decisive quality of true (literary) genius as the successful combination of subconscious instinct with superior talent. When describing this subconscious element, Jean Paul appears to be very close indeed to Gauss's description of his discovery (in Jean Paul's text preceding the following quote, there is also a reference to God somewhat similar to Gauss's letter):

> The instinct is the sense of the future; it is blind, but only as the ear is blind to light, and the eye deaf to sound. It signifies and contains its object in the same way as the cause contains the effect. If the secret were open to us as to how a given cause contains the effect in itself, despite the fact that the effect follows the cause in time, then we would also understand how the instinct claims, determines and knows its object and yet does without it.[12]

There was no reaction in print to Gauss's work [Gauss 1811] until about 1835 when Peter Gustav Lejeune-Dirichlet introduced his new, Fourier-theoretic methods. After this, by 1850 almost all of the general methods of proof had been found although over the 150 years since then many variants have been given. The main later innovation was the method of Hecke and I shall turn to this later. A second very novel proof is due to Issai Schur – see below.

2. Dirichlet and Poisson Summation

The first general method is Dirichlet's – see [Dirichlet 1835], [Dirichlet 1839–40]. It is perhaps worthy of note that Harold Davenport remarked:

---

8. See [Grove 1896], pp. 54–55. Cf. [George 1998]. For later, strongly contrasting views on this theme, we mention in passing Lev Tolstoi's considerations in *War and Peace* (1865–1868) and Thomas Carlyle's *Lectures on Heroes, Hero-Worship, and the Heroic in History* (1840).
9. The letter was written about three weeks before Napoleon's troops crossed the Rhine and more than a year before their advance would directly affect Gauss's life, first in Braunschweig, then in Göttingen.
10. "Jean Paul" was the pen name of Johann Paul Friedrich Richter (1763–1825). Gauss occasionally quoted romantic, witty, or sobering lines from various works of Jean Paul in letters and in conversation.
11. See [Schmidt 1985], pp. 433–446.
12. See Jean Paul, *Vorschule der Ästhetik*, 3rd *Programm*, § 13: *Der Instinkt … ist der Sinn der Zukunft; er ist blind, aber nur, wie das Ohr blind ist gegen Licht und das Auge taub gegen Schall. Er bedeutet und enthält seinen Gegenstand ebenso wie die Wirkung die Ursache; und wär' uns das Geheimnis aufgetan, wie die mit der gegebenen Ursache notwendig ganz und gar zugleich gegebene Wirkung doch in der Zeit erst der Ursache nachfolget, so verständen wir auch, wie der Instinkt zugleich seinen Gegenstand fodert, bestimmt, kennt und doch entbehrt.*

> The method used by Dirichlet in 1835 to evaluate G is probably the most satisfactory of all that are known. It is based on Poisson's Summation Formula, and it has the advantage that once the proof has been embarked upon, no special ingenuity is called for.[13]

Yet, this proof is the one that is least often reproduced; as far as I know, apart from Davenport, Martin Neil Huxley is the only writer to use it.[14]

Dirichlet's proof is as follows. We begin by proving a version of the Poisson Summation Formula. Let f be of bounded variation on $[0, 1]$. Then Dirichlet's Theorem on the representability of functions of bounded variation by their Fourier series yields

$$\frac{1}{2}\big(f(0+)+f(1-)\big) = \lim_{N\to\infty}\sum_{j=-N}^{N}\int_0^1 f(x)e^{-2\pi \mathrm{i} jx}\,\mathrm{d}x$$

Summing this over the intervals $[j, j+1]$ with $0 \le j < k$ gives Dirichlet's version of the Poisson Summation Formula for a continuous function of bounded variation on $[0, k]$:

$$\begin{aligned}\frac{1}{2}f(0)+f(1)+f(2)+\cdots+f(k-1)+\frac{1}{2}f(k)\\ = \lim_{N\to\infty}\sum_{j=-N}^{N}\int_0^k f(x)e^{-2\pi \mathrm{i} jx}\,\mathrm{d}x\end{aligned}$$

Now take $f(x) = \exp(2\pi \mathrm{i}x^2/k)$. The left-hand side is the Gauss sum whereas the right-hand side is

$$\begin{aligned}&= \lim_{N\to\infty}\sum_{j=-N}^{N}\int_0^k \exp(2\pi\mathrm{i}(x^2/k - jx))\mathrm{d}x\\ &= \lim_{N\to\infty}\sum_{j=-N}^{N}\int_0^k \exp\Big(\frac{2\pi\mathrm{i}(x-jk/2)^2}{k}\Big)\mathrm{d}x\ \exp\Big(\frac{2\pi\mathrm{i}j^2k}{4}\Big).\end{aligned}$$

We now separate the sum into those terms where j is odd and those where it is even. We find:

$$\begin{aligned}\lim_{N\to\infty}\sum_{j=-N,\ \text{even}}^{N}\Big\{\int_0^k \exp\big(\frac{-2\pi\mathrm{i}(x-jk/2)^2}{k}\big)\mathrm{d}x\,\exp\big(\frac{2\pi\mathrm{i}j^2k}{4}\big)\\ + i^{-k}\int_0^k \exp\big(\frac{2\pi\mathrm{i}(x-\frac{k}{2}-\frac{j-1}{2}k)^2}{k}\big)\mathrm{d}x\ \exp\big(-\frac{2\pi\mathrm{i}j^2k}{4}\big)\Big\}\\ = \sqrt{k}(1+\mathrm{i}^{-k})\int_{-\infty}^{\infty}\exp(2\pi\mathrm{i}x^2)\mathrm{d}x.\end{aligned}$$

13. See [Davenport 1967], p. 14.

14. See [Huxley 1996], § 5.4.

From $k = 1$ we get

$$\int_{-\infty}^{\infty} \exp(2\pi \mathrm{i} x^2)\mathrm{d}x = \frac{1+\mathrm{i}}{2}.$$

From this Gauss's theorem follows at once, and we have also evaluated the improper integral $\int_{-\infty}^{\infty} \exp(2\pi \mathrm{i} x^2)\mathrm{d}x$.

Note that the proof of the Poisson Summation Formula is not the usual one which is based on the Fourier synthesis of the periodic function $\sum_{n\in\mathbf{Z}} f(n+x)$ and was the method used by Siméon-Denis Poisson [Poisson 1827].[15] In fact, this method could have been used here again in conjunction with Dirichlet's theorem to give an alternative proof. It is worth noting here that Poisson had developed a theory of the representability of periodic functions by their Fourier series based on the use of the Poisson kernel and the summation method $\lim_{y<1, y\to 1} \sum_{n\in\mathbf{Z}} \hat{f}(n) y^{|n|} \exp(2\pi \mathrm{i} nx)$. This is valid for continuous functions. If the sum $\sum_{n\in\mathbf{Z}} \hat{f}(n) \exp(2\pi \mathrm{i} nx)$ converges then, by Abel's theorem,[16] the limit coincides with this sum. Although Abel's paper is almost contemporary with those of Poisson, Abel was interested in the hypergeometric function and made no mention of any application to Fourier series.

The same method was applied by Mathias Schaar and Angelo Genocchi to $\exp(2\pi \mathrm{i} px^2/q)$ to prove the reciprocity formula for Gauss sums, [Schaar 1848], [Genocchi 1852] and [Genocci 1854]. This formula states that, for p and q relatively prime positive integers one of which is even, one has

$$\frac{1}{\sqrt{q}} \sum_{j=0}^{q-1} \exp\Big(\pi \mathrm{i} \frac{pj^2}{q}\Big) = \exp\Big(\frac{\pi \mathrm{i}}{4}\Big) \frac{1}{\sqrt{p}} \sum_{k=0}^{p-1} \exp\Big(-\pi \mathrm{i} \frac{qk^2}{p}\Big).$$

This beautiful formula generalizes that of Gauss and yields the law of quadratic reciprocity as well. For this one only has to observe that, if p and q are odd primes, then an elementary transformation shows that

$$\sum_{j=0}^{q-1} \exp\Big(-\pi \mathrm{i} \frac{pj^2}{q}\Big) = \Big(\frac{-2p}{q}\Big) \sum_{j=0}^{q-1} \exp\Big(2\pi \mathrm{i} \frac{j^2}{q}\Big)$$

from which the assertion easily follows.

The standard argument that one uses today to prove this is the following. Let p and q be distinct odd primes; we shall consider the Gauss sum

$$\sum_{j=0}^{pq} \exp\Big(2\pi \mathrm{i} \frac{j^2}{pq}\Big).$$

15. I would like to take this opportunity to thank Catherine Goldstein for her generous help obtaining copies of Poisson's rather extensive papers. These are regrettably not easily accessible, although their importance in the development of analysis is manifest.

16. See [Abel 1826], Theorem IV.

In this we set $j = pj_1 + qj_2$; by the Chinese Remainder Theorem the set of such j will run through a set of residues (mod pq) if j_1 runs through a set of residues (mod q) and j_2 through a set of residues (mod p). If we replace j^2 by $(pj_1+qj_2)^2 = p^2j_1^2 + 2pqj_1j_2 + q^2j_2^2$ and then one finds that the sum above is equal to

$$\sum_{j_1=0}^{q} \exp\Big(2\pi\mathrm{i}\frac{pj_1^2}{q}\Big) \sum_{j_2=0}^{p} \exp\Big(2\pi\mathrm{i}\frac{qj_2^2}{p}\Big).$$

The first of these sums is

$$\Big(\frac{p}{q}\Big) \sum_{j_1=0}^{q} \exp\Big(2\pi\mathrm{i}\frac{j_1^2}{q}\Big)$$

as was used in the previous paragraph. The second is given by the analogous expression. Thus

$$\sum_{j=0}^{pq} \exp\Big(2\pi\mathrm{i}\frac{j^2}{pq}\Big) = \Big(\frac{p}{q}\Big)\Big(\frac{q}{p}\Big) \sum_{j_1=0}^{q} \exp\Big(2\pi\mathrm{i}\frac{j_1^2}{q}\Big) \sum_{j_2=0}^{p} \exp\Big(2\pi\mathrm{i}\frac{j_2^2}{p}\Big).$$

If we substitute Gauss's evaluation for the three Gauss sums here we obtain the law of quadratic reciprocity. Presumably Gauss's argument was similar.

A more modern and more precise formulation of these relationships is due to André Weil: Let (e_p) be a non-trivial character on the adèle ring $\mathbf{Q_A}$ of $\mathbf{Q}$, which is trivial on $\mathbf{Q}$. Then Weil's formula states that the Hilbert symbol at p – either a prime or ∞, i.e., representing the archimedean completion $\mathbf{R}$ of $\mathbf{Q}$ – satisfies

$$(x, y)_p = \frac{\gamma_p(xy)\gamma_p(1)}{\gamma_p(x)\gamma_p(y)} \tag{$*$}$$

where

$$\gamma_p(x) = \int_{\mathbf{Q}_p} e_p(x \cdot u^2)\mathrm{d}u,$$

(an improper integral) and, with a self-dual additive measure on the adèles, one has for $x \in \mathbf{Q}$ the product formula

$$\prod_p \gamma_p(x) = 1,$$

where the product is over all primes and ∞. The Hilbert-Furtwängler form of the reciprocity law is

$$\prod_p (x, y)_p = 1$$

which therefore follows immediately.

One of the major preoccupations around 1840–1860 was the generalization of the quadratic reciprocity law. Since the Hilbert symbol of order n is skew–symmetric it follows that for $n > 2$ there is no analogue to (*). The recognition that this was so was chiefly due to Eisenstein – see [Eisenstein 1850a]. Before this paper, Kummer had tried to formulate a version of the general reciprocity law in terms of "ideal numbers," i.e., ideals. For this to make sense he was forced to restrict his attention to the case of regular cyclotomic fields, that is, fields where the class number is coprime to n. In this, n was usually taken to be prime. Eisenstein recognised that this was not natural but that one could define a Legendre symbol $\left(\frac{A}{B}\right)$ of order n by copying the classical definition from the quadratic case. It is possible for B to be an ideal but not A. For the symbol to be well defined, B has to be coprime to n and A and B have to be coprime to one another. In [Eisenstein 1850a], Eisenstein posits a reciprocity law of the form

$$\left(\frac{A}{B}\right) = (A, B)\left(\frac{B}{A}\right)$$

under the assumption that A is also coprime to n. Here (A, B) should depend only on A, B modulo some power of n. By an argument using congruences and special choices of the A, B he shows that (A, B) is completely determined should the reciprocity law hold in this form.[17]

Although the notation used by Eisenstein is very similar to that used by Hilbert, this is probably a coincidence as Hilbert came to introduce the symbol named after him by a rather different line of argument. It apparently developed while he was writing the *Zahlbericht* [Hilbert 1897] where it is introduced in § 64 while he is discussing the theory of genera in quadratic fields. The product formula appears as a lemma (*Hilfssatz 14*) in [Hilbert 1897], § 69. It seems clear that Hilbert had not yet recognised the fundamental nature of his product formula; this he emphasises in [Hilbert 1899]. In [Hilbert 1897] he also introduced an analogue of the quadratic Hilbert symbol in the case of cyclotomic fields.[18] This is based on Kummer's approach to the general reciprocity law[19] ; Hilbert, like Kummer, proves it when n is a regular prime. At the end of [Hilbert 1897], § 161, Hilbert notes the relationship to Eisenstein's approach.

Once Hilbert had formulated the general reciprocity law in a general context, one of his students, Philipp Furtwängler[20] (1869–1940), took up the challenge of proving it. This he did in a series of papers over a long period[21] in which he removed the condition that n should be regular; n remained a prime, with 2 permitted. These papers use the methods developed by Kummer, Hilbert and Heinrich Weber and are

17. This argument is repeated in [Cassels, Fröhlich 1967] as Exercises 2.12 and 2.13, pp. 353–354.
18. See [Hilbert 1897], § 148.
19. See [Kummer 1859].
20. In fact, Furtwängler was formally a student of Felix Klein, but worked along the lines of Hilbert's *Zahlbericht*.
21. See [Furtwängler 1902], [Furtwängler 1909], [Furtwängler 1912], [Furtwängler 1913], and [Furtwängler 1928].

fairly technical. They do make use of the Hilbert symbol as introduced in [Hilbert 1897]. These papers became obsolete with the development of the new class–field theory in the 1920s by Emil Artin, Helmut Hasse, and others, for it was easy to deduce them from Artin's Reciprocity Law. The fact that this form of the reciprocity law was introduced by Hilbert and was proved in a large number of cases by himself and Furtwängler seems, at least to the author, to justify the designation Hilbert-Furtwängler Reciprocity Law.

A number of variants of the proofs using Gauss sums indicated above were found by about 1850. Thus both Victor Amédée Lebesgue [Lebesgue 1840] and Ferdinand Gotthold Max Eisenstein[22] gave very elegant derivations of Gauss's q-formula. On the other hand Augustin Louis Cauchy gave in [Cauchy 1840] a quite different proof of the identity

$$\sum_{j=0}^{n-1} \zeta^{j^2} = \left(\frac{-2}{n}\right) \prod_{k=1}^{(n-1)/2} (\zeta^k - \zeta^{-k}) \tag{†}$$

where n is odd and ζ is a primitive n^{th} root of unity. This is essentially equivalent to the formula derived by Gauss. Cauchy's proof consists essentially in showing that the quotient of the two sides is a root of unity in $\mathbf{Z}[\zeta]$ and then, at least when n is a prime, using a congruence argument to show that this root of unity is in fact 1. More or less the same proof was published by Leopold Kronecker in [Kronecker 1856].

A very interesting and suggestive proof of (†) was given by Issai Schur [Schur 1921]. He interpreted the left–hand side of the formula as the trace of the Fourier transformation over a finite field and investigated the matrix of this transformation. The eigenvalues are fourth roots of 1; the determinant is a Vandermonde determinant which gives enough information to determine the multiplicities.

Later, as complex analysis developed, it was possible to replace Fourier theory by complex analytic methods. This was first done by Leopold Kronecker [Kronecker 1889] but there were many variants found later. In particular Louis Joel Mordell developed these ideas[23] and they were taken up again by George Neville Watson in his work on Ramanujan's Mock Theta Functions.[24]

Likewise, real-analytic proofs were found, the best-known being Theodor Esterman's [Estermann 1945], but Edmund Landau's version in [Landau 1928] is both interesting and earlier. Gauss's formula is also a consequence (for almost all n) of the Euler-Maclaurin Summation Formula, or of van der Corput's Summation Formula, i.e., an approximate version of the Poisson Summation Formula with a truncated sum over the Fourier transform – see [Lehmer 1976], although this was presumably known earlier.[25] Although it may be pushing the point a bit, one can regard these analytic proofs as being variations on Dirichlet's theme.

---

22. In [Eisenstein 1844a], [Eisenstein 1844b], [Eisenstein 1844c], and esp. [Eisenstein 1844f].
23. See [Mordell 1918] and [Mordell 1933].
24. See [Watson 1936] and [Watson 1937].
25. A lengthy bibliography of the proofs of Gauss's theorem, with more emphasis on the more recent ones than here, is given in [Berndt, Evans, Williams 1998].

3. Theta Functions

The theory of quadratic Gauss sums runs parallel to the theory of theta functions especially in terms of the techniques used. However, the theory of theta functions, especially the transformations under the modular group, can be used to determine the Gauss sums. This was first discovered by Cauchy, [Cauchy 1840], the same paper where his proof of Gauss's theorem based on a congruence argument appeared. A little later Carl Gustav Jacob Jacobi [Jacobi 1848] indicated vaguely what form the general transformation law should have; the multiplier system has a Jacobi symbol as one of its major components. This was formulated precisely and proved by Charles Hermite in [Hermite 1858]. This emphasizes the relationship between the reciprocity formula for Gauss sums and the transformation theory for the theta functions. Many proofs have been given. Particularly noteworthy is that of Heinrich Weber in [Weber 1908], § 38, which is group-theoretic. Although it did not have any immediate influence the same idea was rediscovered by Tomio Kubota in 1964.[26] This can be combined with more modern techniques in the theory of automorphic functions to investigate Gauss sums of higher order. Elaborating on this would, unfortunately, take us too far afield.

In the attempts to discover and to prove a general reciprocity law in the 1840s, Gotthold Eisenstein used Gauss sums in a very sophisticated fashion.[27] It appears that the experience gained in these works was his major motivation for formulating the "general" reciprocity law [Eisenstein 1850a], that is, the Hilbert-Furtwängler reciprocity law. The Gauss sums that he used were, in contradistinction to Gauss and Ernst Eduard Kummer, those of the form

$$\sum_{x \pmod{n}} \chi(x) \exp\Big(2\pi \mathrm{i}\frac{x}{n}\Big),$$

but defined over the ring of integers of a cyclotomic field and where χ is a suitable Legendre-Jacobi symbol of higher order.[28] These sums are the basis of Eisenstein's reciprocity law, namely that

$$\left(\frac{A}{B}\right) = \left(\frac{B}{A}\right)$$

when either A or B is a rational integer.[29] It turns out that, as Eisenstein discovered, first in various special cases in 1844,[30] and then in the case of cyclotomic fields of

26. See [Kubota 1966b]. This was first indicated by Kubota in [Kubota 1966a].
27. See [Eisenstein 1844a], [Eisenstein 1844d], [Eisenstein 1844e], [Eisenstein 1844g], [Eisenstein 1844h], and [Eisenstein 1850b].
28. The description is a little anachronistic as the theory of algebraic numbers was just beginning to develop at this time. The notion of a character is generally introduced through artifice (for us) of the index with respect to a primitive root.
29. See [Eisenstein 1850b].
30. See [Eisenstein 1844c], [Eisenstein 1844d], [Eisenstein 1844e], [Eisenstein 1844g], and [Eisenstein 1844h].

prime order in [Eisenstein 1850b], certain powers of the Gauss sums can be given explicitly.

This was rediscovered by André Weil in 1952 in a more general setting, and extended by him further in 1974.[31] In the course of this, one needs some p-adic theory of Gauss sums, which is to say that one determines both the powers of those prime ideals dividing the Gauss sums and certain congruences that they satisfy. This was pioneered first by Jacobi and then by Eisenstein in increasing generality.[32] It was given in a general form by Ludwig Stickelberger in [Stickelberger 1890]. Gauss's investigations in the same direction were not published in his lifetime and for this reason we shall not consider them further.

Cauchy's "p-adic" proof of Gauss's formula led John William Scott Cassels to propose a kind of formula for cubic Gauss sums in terms of elliptic functions.[33] This, and a biquadratic analogue, were proved by Charles Russell Matthews.[34] In terms of individual Gauss sums, little more is known at the moment; to find a more general analogue remains a most challenging problem.

In the 1840s, Kummer proposed, rather tentatively, a statistical distribution of cubic Gauss sums.[35] This proposal turned out to be false in its original form; in the case considered by Kummer this was proved by David Roger Heath-Brown and the author, and in general by the author.[36] These results can be understood as asserting that there is no formula analogous to that of Gauss for Gauss sums of order greater than 2.

4. Hecke's Approach

Around 1900 the theory of the Riemann ζ-function was well-developed – the prime number theorem had been proved and the basis of the determination of the properties of the $L(s, \chi)$ had been laid by Hurwitz in 1882, although one sees, for example in Edmund Landau's *Handbuch*, that the proof still was considered difficult in 1909.[37] In algebraic number theory, and especially in class-field theory, a number of further ζ- and L-functions had been introduced – for example Dedekind's ζ-function of a number field dates to 1871[38] – the ideas are already implicit in Kummer's paper [Kummer 1859] where Kummer cites Dirichlet.[39] Moreover the relationship between $\lim_{s\to 1}(s-1)\zeta_k(s)$ and the class-number of the number field k was known; this was the main motivation behind its introduction. Moreover, one had studied various

31. See [Weil 1952] and [Weil 1974d].
32. See [Jacobi 1827], and the papers [Eisenstein 1844c], [Eisenstein 1844d], [Eisenstein 1844e], [Eisenstein 1844g], [Eisenstein 1844h], and [Eisenstein 1850b].
33. See [Cassels 1970], [Cassels 1977].
34. See [Matthews 1979a] and [Matthews 1979b].
35. See [Kummer 1842] and [Kummer 1846].
36. See [Heath-Brown, Patterson, 1979] and [Patterson 1987].
37. See [Landau 1909], § 103, § 124.
38. See [Dirichlet 1871], Supplement X, § 167 ff.
39. See [Kummer 1859], p. 138, and, perhaps more to the point, [Kummer 1850], introduction.

L-functions of the type $L_k(s, \chi)$, where

$$\chi(\mathfrak{a}) = \left(\frac{\delta}{N_{k/k_0}(\mathfrak{a})}\right)_n ;$$

again such L-series also go back at least to Kummer's paper. One also knew that $L_{\mathbf{Q}}(s, \chi)$, $\chi = (\frac{D}{\cdot})$, is essentially an L-function of the type considered by Hurwitz. But if we now look at Hilbert's 8th problem what we see is rather curious:

> But of no lesser interest, and perhaps even broader consequences, seems to me to be the task of transferring the results obtained about the distribution of rational prime numbers, to the theory of the distribution of prime ideals in a given algbraic number field k – a task which amounts to studying the function associated to the field
>
> $$\zeta_k(s) = \sum \frac{1}{n(\mathfrak{j})^s},$$
>
> where the sum extends over all ideals $\mathfrak{j}$ of the given number field k, and $n(\mathfrak{j})$ denotes the norm of the ideal $\mathfrak{j}$. [40]

Thus Hilbert stresses that the Dedekind ζ-functions could be important, but he *does not* conjecture that they have analytic continuation. One can only read into his formulation that he had at the back of his mind that this might be the case, but he was not prepared to stick his neck out. From a present-day perspective this seems strange; it appears that there was a barrier that had to be overcome. Also the Artin Hypothesis on the analytic properties of Artin L-functions and Hasse's question about the properties of the global L-function of an elliptic curve have so shaped our thinking that we can hardly imagine not putting such questions at the centre of our considerations. It is, however, helpful to realize that, at that time, one apparently felt that whereas the Riemann ζ-function is arithmetically significant, its construction appeared to be analytic and so its analytic properties seemed natural. The Dedekind ζ-function was arithmetic in its definition and so its analytic properties were not considered natural. It is the experience of the last hundred years that leads us to be confident of the analytic properties of such arithmetic functions.

Curiously it was Landau, just three years after Hilbert's Paris lecture, who proved that the Dedekind ζ-function has an analytic continuation into a narrow strip,[41] and he used methods which Hilbert had used in his paper on relative quadratic extensions.[42]

---

40. See [Hilbert 1900], pp. 309–310, the discussion of the 8th problem, *8. Primzahlprobleme*: *… Aber nicht von geringerem Interesse und vielleicht noch größerer Tragweite, erscheint mir die Aufgabe, die für die Verteilung der rationalen Primzahlen gewonnenen Resultate auf die Theorie der Verteilung der Primideale in einem gegebenen Zahlkörper k zu übertragen – eine Aufgabe, die auf das Studium der dem Zahlkörper zugehörigen Funktion* $\zeta_k(s) = \sum \frac{1}{n(\mathfrak{j})^s}$ *hinausläuft, wo die Summe über alle Ideale* $\mathfrak{j}$ *des gegebenen Zahlkörpers k zu erstrecken ist, und* $n(\mathfrak{j})$ *die Norm des Ideals* $\mathfrak{j}$ *bedeutet.*

41. See [Landau 1903a], p. 81.

42. See [Hilbert 1899], *Satz 31*; see also [Weber 1896], § 194.

For Landau, in contrast to Hilbert, the methods of complex function theory were entirely natural; he used his result in [Landau 1903b] to prove the prime ideal theorem, and later developed the method to deal with characters of the class group in [Landau 1907].

Ten years after this, Erich Hecke proved the analytic properties of the Dedekind ζ-function.[43] As one now sees he used a generalization of the known proofs, sharpening the technique used by Landau. The method also immediately extends to the analogues of Dirichlet L-series. Hecke [Hecke 1917] observed moreover that his results implied that the relation (for an abelian extension K/k of number fields, and all characters χ on the Galois group of K/k)

$$\zeta_K(s) = \zeta_k(s) \prod_{\chi \neq 1} L_k(s, \chi)$$

which one can prove for almost all Euler factors directly, has to be true for all factors. Then he remarked that, if one looks at the root number $W_k(\chi)$ appearing in the functional equation, and K/k is quadratic, the functional equation yields

$$W_K(\omega \circ N) = \prod W_k(\omega \cdot \chi),$$

where ω is a *Grössencharakter* for k. The case of a quadratic extension implies the determination of the sign of the quadratic Gauss sum – and its generalization to algebraic number fields.

This line of argument can be considered as a deduction of Gauss's theorem from the law of quadratic reciprocity, and this is what is novel about it. Hecke took this up in his book *Algebraische Zahlen* [Hecke 1923] and used it to prove a beautiful theorem on the different of k, namely that the class of the absolute different in the ideal class group is a square. This theorem – an analogue of the fact that the Euler characteristic of a Riemann surface is even – is the crowning moment (*coronidis loco*) in both Hecke's book[44] and André Weil's *Basic Number Theory*.[45] The same idea also leads to the Davenport-Hasse theorems [Davenport, Hasse 1935]. The idea has been in the modern theory of automorphic forms to deduce that a statement which has been proved at almost all places of an **A**-field is in fact valid at all places. This beautiful technique reflects the marvellous way in which the places of a number-field are intimately bound up with one another, as one already sees in Gauss's theorem.

In considering Gauss' theorem one is impressed over and over again with how fecund it has been. Although at first sight his theorem about Gauss sums may seem to be little more than a curiosity, what we have seen is that it has been at the root of several important developments in number theory over the last 200 years and that it has continued to inspire the leading practitioners of this part of mathematics.

---

43. Hecke refers no more than necessary to Landau; there is one obscurely formulated footnote. This is hard to understand, for five years earlier in the CV attached to his dissertation he had thanked "Herr Professor Landau, to whom I owe a large part of my mathematical education." (*Herrn Professor Landau, dem ich einen großen Teil meiner mathematischen Ausbildung verdanke.*) See [Hecke 1910], p. 58.

44. See [Hecke 1923], § 63, *Satz 176*.

45. See [Weil 1967], chap. XIII, § 12, Theorem 13.

References

ABEL, Niels Henrik. 1826. Untersuchung über die Reihe $1+\frac{1}{m}x+\frac{m.m-1}{2}x^2+\frac{m.m-1.m-2}{2.3}x^3+\cdots$. *Journal für die reine und angewandte Mathematik* 1, 331–339. French version: Recherches sur la série $1+\frac{1}{m}x+\frac{m.m-1}{2}x^2+\frac{m.m-1.m-2}{2.3}x^3+\cdots$. Repr. *Œuvres complètes*, vol. I, ed. L. Sylow, S. Lie, pp. 219–250. Christiana: Grondahl, 1881; repr. New York, London: Johnson, 1965. Amended German transl.: ed. A. Wangerin. Ostwalds Klassiker der exakten Wissenschaften 71. Leipzig: Wilhelm Engelmann, 1895.

BERNDT, Bruce C., EVANS, Ronald J., WILLIAMS, Kenneth S. 1998. *Gauss and Jacobi Sums*. Canadian Mathematical Society Series of Monographs and Advanced Texts. New York: John Wiley.

BIERMANN, Kurt R. 1991. Wandlungen unseres Gaußbildes, *Mitteilungen der Gauß-Gesellschaft Göttingen* 28, 3–13.

CASSELS, John William Scott. 1970. On Kummer sums. *Proceedings of the London Mathematical Society* (3) 21, 19–27.

———. 1977. Trigonometric sums and elliptic functions. In *Algebraic Number Theory (Kyoto International Symposium, Research Institute for Mathematical Sciences, University of Kyoto, Kyoto 1976)*, ed. S. Iyanaga, pp. 1–7. Tokyo: Japanese Society for the Promotion of Science.

CASSELS, John William Scott, FRÖHLICH, Albrecht (eds.). 1967. *Algebraic Number Theory. Proceedings of a Conference in Brighton, UK, 1965.* London, New York: Academic Press.

CAUCHY, Augustin Louis. 1840. Méthode simple et nouvelle pour la détermination complète des sommes alternées formées avec les racines primitives des équations binômes. *Journal de mathématiques pures et appliquées* 5, 154–168.

DAVENPORT, Harold. 1967. *Multiplicative Number Theory*. Chicago: Markham Publishing Company.

DAVENPORT, Harold, HASSE, Helmut. 1935. Die Nullstellen der Kongruenzzetafunktionen in gewissen zyklischen Fällen. *Journal für die reine und angewandte Mathematik* 172, 151–182.

DIRICHLET, Johann Peter Gustav LEJEUNE-. 1835. Über eine neue Anwendung bestimmter Integrale auf die Summation endlicher oder unendlicher Reihen. *Abhandlungen der Königlich preussischen Akademie der Wissenschaften* 1835, 391–407. Repr. in [Dirichlet 1899], pp. 239–256. Condensed French transl. *Journal für die reine und angewandte Mathematik* 17 (1837), 57–67. Repr. in [Dirichlet 1889], pp. 259–270.

———. 1839–1840. Reserches sur diverses applications de l'analyse infinitésimale à la théorie des nombres. *Journal für die reine und angewandte Mathematik* 19, 324–369; 21, 1–12 and 134–155. Repr. in [Dirichlet 1889], pp. 411–496.

———. 1871. *Vorlesungen über Zahlentheorie*, ed. with supplements R. Dedekind. 2nd ed. Braunschweig: Vieweg.

———. 1889. *Werke*, vol. 1, ed. L. Kronecker. Berlin: G. Reimer.

EISENSTEIN, Gotthold. 1844a. Über eine merkwürdige identische Gleichung. *Journal für die reine und angewandte Mathematik* 27, 105–106. Repr. in [Eisenstein 1975], vol. I, pp. 26–27.