Theta Functions, Gauss Sums and Modular Forms

Benjamin Moore

Thesis submitted for the degree of
Master of Philosophy in Pure Mathematics
at The University of Adelaide
Faculty of Engineering, Computer and Mathematical Sciences
School of Mathematical Sciences

THE UNIVERSITY of ADELAIDE

May 2020

An artist's conception of the action of $\Gamma_0(4)$ on the unit disk.

Drawing by Katarzyna Nurowska; used with permission.

Contents

Citation listing for included publications

[Moo20] B. Moore. "A proof of the Landsberg–Schaar relation by finite methods". In: *Ramanujan J.* (2020). DOI: 10.1007/s11139-019-00195-4.

Abstract

We present some results related to the areas of theta functions, modular forms, Gauss sums and reciprocity. After a review of background material, we recount the elementary theory of modular forms on congruence subgroups and provide a proof of the transformation law for Jacobi's theta function using special values of zeta functions. We present a new proof, obtained during work with Michael Eastwood, of Jacobi's theorem that every integer is a sum of four squares. Our proof is based on theta functions but emphasises the geometry of the thrice-punctured sphere.

Next, we detail some investigations into quadratic Gauss sums. We include a new proof of the Landsberg–Schaar relation by elementary methods, together with a second based on evaluations of Gauss sums. We give elementary proofs of generalised and twisted Landsberg–Schaar relations, and use these results to answer a research problem posed by Berndt, Evans and Williams. We conclude by proving some sextic and octic local analogues of the Landsberg–Schaar relation.

Finally, we give yet another proof of the Landsberg–Schaar relation based on the relationship between Mellin transforms and asymptotic expansions. This proof makes clear the relationship between the Landsberg–Schaar relation and the existence of a metaplectic Eisenstein series with certain properties. We note that one may promote this correspondence to the setting of number fields, and furthermore, that the higher theta functions constructed by Banks, Bump and Lieman are ideal candidates for future investigations of such correspondences.

Statement

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works. I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Acknowledgements

Firstly, thank you to my supervisors Mike Eastwood and Thomas Leistner, for a very straightforward two years. Thank you to Mike, who has put up with my constant habit of running off on tangents with not a hint of complaint, and for going out of his way to see me through the last academically hectic six months. Thank you to Thomas, for support throughout, and for some helpful comments on a draft. I am also grateful to the examiners, who corrected several mistakes and made a number of helpful remarks and criticisms.

Thank you to Nick Buchdahl. Without your guidance over the last five years I would have given up many times over. Thank you for not only opening my eyes to the wonders of complex geometry, but also for your inspiring lectures. I owe to you my appreciation for mathematics as well as my outlook on academic life. Thank you for always being available to discuss, in great detail, anything at all; you have been a terrific mentor.

Thank you to David Roberts, for always being available to chat about any and all areas of maths, providing insights on request on everything from bundles to theta functions. In addition, thank you for reading the rolling drafts of the thesis, catching a great number of blunders, and for suggesting a number of interesting references (in particular, the exact location of Bachet's editorial comment, in Latin, on sums of four squares!).

Thank you to Tim Trudgian, for making me feel welcome in the number theory community and for generous advice and support, always given whilst projecting a general air of great enthusiasm. Also, my gratitude to Peter Humphries for numerous bits of mathematical and academic advice.

I would also like to thank Professor Ben Brubaker and Professor Jeff Hoffstein for setting me straight (or at least firmly crooked) on a few points concerning metaplectic theta functions; and Michael Berg, for taking the trouble to send me a copy of one of his publications which had proved otherwise inaccessible.

Particular thanks is due to Professor Samuel Patterson, for patiently providing extremely detailed responses to a number of questions. Without his help, I would be even further out to sea than I already am.

Thank you also to Anthony, Joe and Paul, for promoting almost a decade of sensible and responsible behaviour, only very occasionally straying into the realms of the lunatic; and also to Jake, a somewhat more moderating influence.

Most of all, to my mother and grandmother, for your unwavering support through all these years. Without you both I would never have made it this far.

Introduction

This thesis is concerned with the relationships between modular forms, theta functions, Gauss sums and reciprocity laws. These areas of mathematics are deeply interconnected and have their roots in the number-theoretic investigations of antiquity. We do not focus on a single problem, but rather investigate several different aspects of the theory: we hope that the reader will forgive the accordingly extended exposition.

At a basic level, we are concerned with *perfect squares*. A classical problem, first explicitly stated as a editorial note by Bachet in his 1621 translation of Diophantus' *Arithmetica* [Dio21, Liber IIII, Quaestio XXXI], is to determine whether or not every integer may be represented as a sum of four squares. For example,

$$23 = 3^2 + 3^2 + 2^2 + 1^2,$$

and Bachet checked that all positive integers up to and including 325 may be represented in such a way (though he only explictly listed the paritions into squares of the integers up to 120). In 1770, Lagrange [Lag72] proved that every integer is indeed a sum of four squares; but the proof that concerns us is the 1829 contribution of Jacobi [Jac29], which gives a formula for the number of ways in which an integer may be thus represented in terms of its divisors. His proof rests on the remarkable properties of an analytic object now known as a *theta function*, and is an early and striking example of the principle that *certain analytic functions contain number-theoretic information*.

Jacobi's theta function is remarkable for a second reason: it is one of the oldest examples of a *modular form*. Investigated throughout the 19th century in various guises, a comprehensive theory of modular forms started to become available in the first half of the 20th century, rapidly establishing itself as a centerpiece of modern number theory. There is an enlightning modern proof of Jacobi's theorem using modular forms, which allows one to readily appreciate methods for attacking similar problems. In Chapter 3, we show that one may get away with less machinery, reducing the proof to its geometric essentials.

During the same era as Jacobi, Gauss was beginning to construct a solid foundation for the study number-theoretic phenomena. His book *Disquisitiones Arithmeticae* is a remarkable achievement in this direction: it includes, amongst other gems, a statement of the law of *quadratic reciprocity*. This miraculous theorem states that the solvability of the equations

$$x^2 = p \bmod q \quad \text{and} \quad x^2 = q \bmod p,$$

where p and q are distinct odd primes, are related. The problem of generalising this result to higher powers was partly responsible for the creation of the field of *algebraic number theory*, and has continued to shape the fortunes of mathematics up to the present day.

One may wonder whether there is a connection between theta functions and quadratic reciprocity, as both are concerned with squares, and it does turn out that this is the case. Many mathematicians of the 19th century seem to have had a hand in the process, but it was Schaar [Sch50] who first proved, in 1850, that the law of quadratic reciprocity may be deduced from the transformation law for Jacobi's theta function. His proof contains, as an interim step, a remarkable identity between two finite sums, which has become known as the *Landsberg–Schaar relation*, after the independent discovery of Landsberg [Lan93] in 1893. We are particularly concerned with the Landsberg–Schaar relation in Chapter 4, in which we give two elementary proofs and a number of generalisations.

The finite sums involved in Schaar's identity are known as *Gauss sums*, as the same objects were used prolifically by Gauss in his number-theoretic investigations. The Gauss sums have remarkable properties, and are almost by definition closely linked to exponential sums twisted by the quadratic residue symbol. Naturally, mathematicians search for cubic and higher analogues, and Kummer [Kum42] was the first to notice (in print) that, in stark contrast to the situation for Gauss sums, the sums formed with *cubes* obey no obvious pattern. Almost 150 years later, Patterson [Pat77a; Pat77b] proved that the arguments of such sums are equidistributed using the *higher theta functions* mentioned in Chapter 6. Remarkably, in 1979 Matthews [Mat79b], using modular forms, provided an evaluation (of sorts) of a *quartic* Gauss sum. No analogous result for higher powers has ever been found.

We turn to the contents of the thesis. In the first chapter, we review basic material from differential geometry, Fourier analysis and algebraic number theory. The second chapter recounts the theory of modular forms for congruence subgroups in enough detail to provide the "usual" proof of Jacobi's theorem [Jac29] that every positive integer may be expressed as a sum of four squares. In this chapter we also include a new proof of the transformation law for Jacobi's theta function, based on special values of zeta functions. This proof has the advantage that it does not require Fourier analysis, and instead makes use of the reduction, via purely combinatorial methods, of special values of the *Mordell–Tornheim zeta function* to special values of the Riemann zeta function.

In the third chapter, we present an original geometrical variation of the proof of Jacobi's theorem on sums of four squares, making use of the special properties of the quotient of the upper half plane by a certain congruence subgroup. The "usual proof", outlined elsewhere [DS05], makes no real use of the fact that we are interested in *four* squares, as opposed to six, eight or ten, but we enthusiastically take advantage of geometric simplifications possible in this special case. During the proof, we show that the weight two Eisenstein series G_2 is quasimodular by investigating a certain *projectively invariant* differential operator. This chapter represents joint work with Michael Eastwood.

In the fourth chapter, we outline some of our results on Gauss sums over $\mathbb{Q}$. We include our proof [Moo20] of the Landsberg–Schaar relation by elementary methods, followed by the details of a second proof based on the evaluation of Gauss sums. We present elementary proofs of the generalised and twisted Landsberg–Schaar relations, building on work of Guinand [Gui45] and Berndt [Ber73]. We also prove, again using elementary methods, a local quartic version of the Landsberg–Schaar relation, thus answering a question posed by Berndt, Evans and Williams [BEW98, Research Problem 8, pp. 496]. We conclude with some local sextic and octic analogues of the Landsberg–Schaar relation, incorporating the evaluations of cubic and quartic Gauss sums due to Matthews [Mat79a; Mat79b]. We believe these results to be new.

In the fifth chapter we give a brief overview of our attempts towards generalising the results of Chapter 4 to algebraic number fields. In this setting, one has a concrete law of quadratic reciprocity, and in recognition of Hecke's fundamental work, the analogues of the Gauss sums of Chapter 4 are called *Hecke sums.* In analogy with the Landsberg–Schaar relation, these sums enjoy an identity known as *Hecke reciprocity.* We prove, using theta functions, a generalised version of Hecke's reciprocity relation, together with a twisted version, both valid over totally real number fields.

Although our results are too restrictive at present to obtain local quartic analogues of Hecke reciprocity, we observe that most of the other results from Chapter 4 generalise immediately to this setting. However, we expect that the *Hecke theta functions* suffice to provide an analytic proof of the "correct" twisted relation; given this, it should be an easy matter to determine the local quartic version of Hecke reciprocity. We also advocate the evaluation of the Hecke sums, in parallel to Chapter 4, as has been proposed by Boylan and Skoruppa [BS13, pp. 111] and carried out by the same authors over quadratic number fields [BS10].

In the final chapter, we outline a natural avenue through which one might attempt to generalise the correspondence between theta functions, the Landsberg–Schaar relation and quadratic reciprocity. We advocate the viewpoint that the Landsberg–Schaar relation arises from the asymptotic expansion of Jacobi's theta function, which may be computed from information about the locations of the singularities of Dirichlet L-functions. We briefly recall the theory, due to Weil [Wei64], of metaplectic covers of $GL(2)$, which, when combined with Maaß's calculations of Eisenstein series, constructs the theta function *canonically* from the law of quadratic reciprocity.

We repeat the construction of the Eisenstein series in detail and compute its asymptotic expansion at the cusps, verifying that one obtains the Landsberg–Schaar relation. From this perspective, we note that the appearance of squares in the Landsberg–Schaar relation is governed by the fact that the Fourier coefficients of the Eisenstein series are Dirichlet series formed from quadratic characters.

Thus, one might be tempted to think that *higher-degree analogues* of the Landsberg–Schaar relation may be obtained by computing residues of Eisenstein series formed from higher reciprocity laws, but these considerations belong to a rather difficult corner of the theory of automorphic forms. Given a number field K containing the rth roots of unity, it is possible, following work of Kazhdan–Patterson [KP84], to construct Eisenstein series over $GL(n)$, but a result of Deligne [Del80] shows that hideous difficulties are encountered unless $r = n$ or $r = n + 1$.

Furthermore, only in the case $n = r$ does one obtain that the Fourier coefficients of the Eisenstein series are Dirichlet series twisted by characters of order n [BBL03]. Consequently, further investigations must take place in the language of *automorphic forms* rather than modular forms; this we defer for the moment. We do, however, review the construction of the quadratic theta function over $\mathbb{Q}(i)$, as treated by Hoffstein [Hof91], and the breakthrough computation, due to Bump and Hoffstein, of the Fourier coefficients of the cubic theta function on $GL(3)$ over $\mathbb{Q}\left(-\frac{1}{2} + \frac{\sqrt{3}}{2}i\right)$.

Chapter 1

Overview of differential geometry, Fourier analysis and number theory

In this first chapter we present the foundational material required for later chapters. Most of this material is well-known, and serves to accustom the reader to the gadgets which will take centre stage later on.

In the first section, we deal with geometry. In particular, following Diamond–Shurman and Shimura, we include the rather tedious proof of the existence of the structure of a complex manifold on the compactification of quotients of the upper half plane by congruence subgroups. These objects, known as *modular curves*, are of fundamental importance, as they are the natural habitat of *modular forms*. Afterwards, we temper this excess of explication with a more intuitive geometrical argument. We also pass very quickly over vector bundles and projective and conformal structures on manifolds. The main ideas at work in Chapter 3 are motivated by a perspective centred on the relationships between projective structures and modular curves; however, the reader may rest assured that no especially deep understanding of either topic is required to understand our arguments in that chapter: we simply use this projective perspective to take shortcuts.

In the second section, we present Fourier analysis on finite *abelian* groups and on $\mathbb{R}^n$. We will use the results stated there for markedly different purposes: the results on finite abelian groups will be employed to study *Dirichlet characters* and *Gauss sums*, in Chapters 4, 5 and 6, whereas the results for $\mathbb{R}^n$ — namely, the *Poisson summation formula* — will be used exclusively to prove functional equations for a very important class of modular forms known as *theta functions*, entities which will pervade all subsequent chapters.

In the final section, we collect together facts related to algebraic number theory. We recall the fundamental definitions for algebraic number theory in the first subsection, with an eye toward techniques related to the study of *ideal numbers*, as they appear in Hecke's theory of Gauss sums over number fields. We introduce *quadratic reciprocity* and the *Legendre symbol*, and we state some particular cases of *cubic* and *quartic* reciprocity, as the focus of Chapter 6 is on the construction of *higher theta functions* using these results. We define *Gauss sums* and make use of some of the Fourier-analytic identities stated earlier to prove results preparatory to Chapters 4 and 5.

Also included is a definition of the *Dedekind zeta functions* and *Dirichlet L-functions*, together with a catalogue, to be deployed in Chapter 6, of the positions of the poles and trivial zeros. A somewhat more esoteric inclusion is the subsection defining multiple zeta values and the Mordell–Tornheim zeta function. Here, we outline a series of results which allow for the evaluation of a particular Mordell-Tornheim zeta value without the use of Fourier analysis (by which we mean Poisson summation and Parseval's formula). In fact, the argument, one part of which is due to Euler, is combinatorial except for the well-known evaluations of $\zeta(4)$ and $\zeta(6)$, for which one may employ the product formula for the sine function. We will make use of these results in Subsection 2.2.2, where we prove the all-important functional equation for Jacobi's theta function by means of an integral transform and the evaluation of special values of exactly the zeta functions discussed above.

1.1 Manifolds and differential geometry

Manifolds are the fundamental objects of study in differential geometry. Interesting functions on manifolds often take values in objects called *bundles*. We present the minimum of these topics needed for subsequent chapters. The reader will find the necessary background material in texts on the theory of smooth manifolds [Lee12] and complex manifolds [Gun66; Wel86; Kod86].

Suppose that X is a second countable Hausdorff topological space. A *smooth* (resp. *complex*) *atlas* is a collection $\mathscr{A} = \{(U_\alpha, \phi_\alpha)\}$ where $\{U_\alpha\}$ is an open cover of X and ϕ_α is a homeomorphism from U_α onto an open subset of $\mathbb{R}^n$ (resp. $\mathbb{C}^n$) such that if $U_\alpha \cap U_\beta \neq \varnothing$, the map

$$\phi_\beta \circ \phi_\alpha : \phi_\alpha(U_\alpha \cap U_\beta) \to \phi_\beta(U_\alpha \cap U_\beta)$$

is diffeomorphism (resp. biholomorphism). An atlas is said to be *maximal* if it is not properly contained in any other atlas.

Definition 1.1.0.1. *A* smooth *(resp.* complex*)* manifold *of dimension n is a second countable Hausdorff topological space X together with a maximal smooth (resp. complex) atlas containing charts into $\mathbb{R}^n$ (resp. $\mathbb{C}^n$). A complex manifold of (complex) dimension 1 is called a* Riemann surface.

We defer a discussion of the basic differential geometry required for our narrative until Subsection 1.1.2, where we define vector bundles and connections.

1.1.1 Group actions on topological spaces

> *The first few pages of Shimura's book give all the details for a complete description of $\Gamma\backslash\mathscr{H}^*$. As this is exceedingly boring, we will not reproduce the arguments here...*
>
> *(Lang, Introduction to Modular Forms, pp. 26)*

The goal of this section is to prove that, for certain arithmetic subgroups of $SL(2,\mathbb{R})$ acting on the upper half plane $\mathscr{H}$ by Möbius transformations, the quotient $\Gamma\backslash\mathscr{H}$ may be compactified and endowed with the structure of a compact Riemann surface. We present a synthesis of the treatments of Shimura [Shi71, Chapter 1] and Diamond–Shurman [DS05, Chapter 2].

Let X be a topological space and G be a topological group. A continuous map

$$\mu : G \times X \to X,$$

often abbreviated as $\mu(g,x) = gx$, is a (left) G-action on X if

1. $g_1(g_2x) = (g_1g_2)x$ for all $g_1, g_2 \in G$ and $x \in X$;
2. $ex = x$ for all $x \in X$.

The set of orbits $G\backslash X = \{\Gamma x \mid x \in X\}$ is a topological space under the quotient topology induced by the projection $\pi : X \to G\backslash X$. We are typically interested in dealing with G and X such that the (topological) space of orbits $G\backslash X$ is Hausdorff. The next definition gives a condition on the action such that this is true.

Definition 1.1.1.1. *A topological group G acts* properly discontinuously *on a topological space X if for each pair of points x_1 and x_2 of X, there exist neighbourhoods U_1 of x_1 and U_2 of x_2 such that for all $\gamma \in G$, $\gamma U_1 \cap U_2 \neq \varnothing$ implies $\gamma(x_1) = x_2$.*

We now prove that the quotient $G\backslash X$ is Hausdorff.

Proposition 1.1.1.2. *Suppose that X is a topological space and that G acts continuously and properly discontinuously on X. Then $G\backslash X$ is Hausdorff.*

Proof. Let $\pi : X \to G\backslash X$ be the projection. One easily checks that for subsets U_1 and U_2 of X,

$$\pi(U_1) \cap \pi(U_2) = \varnothing \text{ in } G\backslash X \quad \text{if and only if} \quad GU_1 \cap U_2 = \varnothing \text{ in } X. \tag{1.1}$$

Now, let $\pi(x_1)$ and $\pi(x_2)$ be distinct points in $G\backslash X$. Select neighbourhoods of x_1 and x_2 according to Definition 1.1.1.1. By assumption, $\gamma x_1 \neq x_2$ for any $\gamma \in G$, so Definition 1.1.1.1 implies that $\gamma U_1 \cap U_2 = \varnothing$. Then by 1.1, $\pi(U_1)$ and $\pi(U_2)$ are disjoint supersets of $\pi(x_1)$ and $\pi(x_2)$. Since $G\backslash X$ has the quotient topology, π is an open mapping, so $\pi(U_1)$ and $\pi(U_2)$ are disjoint neighbourhoods of $\pi(x_1)$ and $\pi(x_2)$. ☐

Actually, we are particularly interested in G and X such that $G\backslash X$ is a smooth manifold. Ignoring pathological examples, groups which act properly discontinuously on Hausdorff spaces are discrete, and we will use the notation Γ in this section to denote a discrete group. The extent to which $G\backslash X$ may be endowed with useful geometric structure is mainly governed by the degree to which the action of Γ doesn't fix points of X.

Definition 1.1.1.3. *A topological group G acts freely on a topological space X if $gx = x$ for some $x \in X$ implies that $g = e$. The isotropy subgroup G_x of a point $x \in X$ is the subgroup of all $g \in G$ which fix x:*

$$G_x = \{g \in G \mid gx = x\}.$$

We set $G_X = \{g \in G \mid gx = x \text{ for all } x \in X\}$. Then G_X is normal in G, so we may define the reduced isotropy subgroup of x to be $\overline{G_x} = G_x/G_X$.

Obviously, G acts freely if and only if $G_x = 0$ for all $x \in X$.

A topological group G which also admits a smooth manifold structure so that the multiplication-and-inversion map $(g, h) \mapsto gh^{-1}$ is smooth is said to be a *Lie group*. A Lie group G acts smoothly on a manifold M if the action μ is a smooth map between manifolds. The next result states that if Γ is a discrete Lie group acting freely, smoothly and properly on a manifold M, then $\Gamma\backslash M$ is guaranteed to admit the structure of a smooth manifold.

Proposition 1.1.1.4. *[Lee12, Theorem 21.13] Let M be a connected smooth manifold and let Γ be a discrete Lie group acting smoothly, freely and properly*[1] *on M. Then the space of orbits $G\backslash M$ is a topological manifold with a unique smooth structure such that the canonical quotient $\pi : M \to \Gamma\backslash M$ is a* smooth normal covering map *(the definition [Lee12, pp. 163] of a smooth normal covering map is quite involved).*

Remark 1.1.1.5. *Proposition 1.1.1.4 holds with "smooth" replaced by "complex" in reference to G and M, and "smooth" replaced by "holomorphic" in reference to π and the implicit action map μ.*

Unfortunately, the groups Γ which are of interest to arithmeticians do not usually act freely, but nevertheless, $\Gamma\backslash X$ turns out to be a manifold. We shall be concerned with discrete subgroups Γ of $SL(2, \mathbb{Z})$ acting on the upper half plane:

$$\mathscr{H} = \{z \in \mathbb{C} \mid \mathrm{Im}(z) > 0\},$$

and the action of such a group on $\mathscr{H}$ will always be that of fractional linear transformations (also known as Möbius transformations):

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} z = \frac{az + b}{cz + d}.$$

We are going to prove a much stronger statement than the claim that $\Gamma\backslash\mathscr{H}$ admits a smooth manifold structure: we will prove that $\Gamma\backslash\mathscr{H}$ may be compactified to obtain a compact Riemann surface if the index of Γ in $SL(2, \mathbb{Z})$ is finite. We will have to investigate the structure of $SL(2, \mathbb{Z})$ more minutely before we are in a position to do this, but we may first deduce that $\Gamma\backslash\mathscr{H}$ is Hausdorff from the following more general result:

Proposition 1.1.1.6. *Let G be a locally compact Hausdorff group and K a compact subgroup acting on the right by group multiplication. Let $S = G/K$ and suppose that Γ is a discrete subgroup of G acting continuously on S. Then the action of Γ is properly discontinuous on S, so that $\Gamma\backslash G/K$ is Hausdorff.*

We require a few short lemmas before the proof.

Lemma 1.1.1.7. *Suppose that G is a locally compact group and K is a compact subgroup. Set $S = G/K$ and let $\pi : G \to S$ be the canonical projection. If B is a compact subset of S, then $\pi^{-1}(B)$ is compact.*

Proof. Since G is locally compact, we may select an cover of G consisting of open subsets with compact closure. The image of this collection under π is an open cover of B, so B admits a finite cover by open subsets of S with compact closure: $B \subseteq \cup_i V_i$. Therefore $\pi^{-1}(B) \subseteq \cup_i \overline{V_i} K$. Since $\pi^{-1}(B)$ is a closed subset of $\cup_i \overline{V_i} K$, and each $\overline{V_i} K$ is compact, $\pi^{-1}(B)$ is compact. □

Lemma 1.1.1.8. *In addition to the hypotheses and notation of Lemma 1.1.1.7, suppose that Γ is a discrete subgroup of G. Then for any two compact subsets B_1 and B_2 of S, $\{\gamma \in \Gamma \mid \gamma B_1 \cap B_2 \neq \varnothing\}$ is finite.*

[1]Here, Γ acts properly on M if the map $\Gamma \times M \to M \times M$ defined by $(g, m) \mapsto (gm, m)$ is proper in the usual sense. This is weaker than the requirement that the action $\Gamma \times M \to M$ is a proper map.

Proof. Let $D_1 = \pi^{-1}(B_1)$, $D_2 = \pi^{-1}(B_2)$. If $\gamma B_1 \cap B_2 \neq \varnothing$ then $\gamma D_1 \cap D_2 \neq \varnothing$, so $\gamma \in \Gamma \cap D_1 D_2^{-1}$. By Lemma 1.1.1.7, D_1 and D_2 are compact, so $D_1 D_2^{-1}$ is compact. Since Γ is discrete, $\Gamma \cap D_1 D_2^{-1}$ is compact and discrete, so it must be finite. □

Lemma 1.1.1.9. *In addition to the hypotheses and notation of Lemma 1.1.1.8, suppose that G is Hausdorff. Then for every $x \in S$, there exists a neighbourhood U of x such that $\{\gamma \in \Gamma \mid \gamma U \cap U \neq \varnothing\} = \{\gamma \in \Gamma \mid \gamma x = x\}$.*

Proof. Every quotient of a locally compact group is locally compact, so we may select a neighborhood V of x with compact closure. By Lemma 1.1.1.8, $\{\gamma \in \Gamma \mid \gamma V \cap V \neq \varnothing\}$ is a finite set, which we write $\{\gamma_1, \ldots, \gamma_k\}$. Let I be the set of indices i such that $\gamma_i x \neq x$. Since G is Hausdorff, K is closed, so S is Hausdorff. Therefore, for each $i \in I$, we may select neighbourhoods V_i of x and W_i of $\gamma_i x$ such that $V_i \cap W_i = \varnothing$. Now put $U = V \cap \left(\cap_{i \in I} V_i \cap \gamma_i^{-1} W_i\right)$. □

Proof of Proposition 1.1.1.6. We will show that if x_1 and x_2 are points of S which are not in the same orbit under Γ, then there exist neighbourhoods U_1 of x_1 and U_2 of x_2 such that $\gamma U_1 \cap U_2 = \varnothing$ for every $\gamma \in \Gamma$.

Indeed, let X_1 and X_2 be compact neighborhoods of x_1 and x_2 respectively. By Lemma 1.1.1.8,

$$\{\gamma \in \Gamma \mid \gamma X_1 \cap X_2 \neq \varnothing\}$$

is a finite set, which we write as $\{\gamma_1, \ldots, \gamma_k\}$. By assumption, S is Hausdorff, so we may find neighbourhoods U_i of $\gamma_i x_1$ and V_i of x_2 such that $U_i \cap V_i = \varnothing$. Then if we set $U = X_1 \cap g_1^{-1}(U_1 \cap \cdots \cap g_k^{-1}(U_k)$ and $V = X_2 \cap V_1 \cap \cdots \cap V_k$, the claim follows from Lemma 1.1.1.9. □

Corollary 1.1.1.10. *For any discrete subgroup Γ of $SL(2,\mathbb{R})$ acting continuously on $\mathscr{H}$, the quotient $\Gamma\backslash\mathscr{H}$ is a locally compact Hausdorff space.*

Proof. We apply Proposition 1.1.1.6, with $G = SL(2,\mathbb{R})$ and $K = SO(2,\mathbb{R})$. Since $SO(2,\mathbb{R})$ is the isotropy subgroup of i under the action of $SL(2,\mathbb{R})$ on $\mathscr{H}$ by Möbius transformations, $S = \mathscr{H}$. Note that G is locally compact, so its quotients are locally compact too. □

From now onwards, Γ is a discrete subgroup of $SL(2,\mathbb{Z})$ acting on $\mathscr{H}$ by Möbius transformations. We have already mentioned that we wish to compactify $\Gamma\backslash\mathscr{H}$. We will now define the compactification and its topology. It is necessary to define this compactification fairly explicitly in order to be able to define a complex structure later on. Let ∞ denote the usual complex infinity obtained by compactifying $\mathbb{C}$. Any subgroup of $SL(2,\mathbb{R})$ acting on $\mathscr{H}$ by Möbius transformations also acts on $\mathscr{H} \cup \{\infty\}$ in the obvious manner.

Definition 1.1.1.11. *The orbits of $\mathbb{Q} \cup \{\infty\}$ under the action of Γ are called cusps.*

We set

$$X(\Gamma) = \Gamma\backslash\left(\mathscr{H} \cup \mathbb{Q} \cup \{\infty\}\right) = (\Gamma\backslash\mathscr{H}) \cup \left(\Gamma\backslash\left(\mathbb{Q} \cup \{\infty\}\right)\right).$$

We now define a topology on $\mathscr{H} \cup \mathbb{Q} \cup \{\infty\}$. If $x \in \mathscr{H}$, we take a fundamental system of open neighbourhoods at x to be the usual one. In order to take the cusps into account we adjoin a neighbourhood base consisting of the collection

$$\{\gamma \cdot (N_t \cup \{\infty\}) \mid t > 0 \text{ and } \gamma \in SL(2,\mathbb{Z})\},$$

where $N_t = \{x \in \mathscr{H} \mid \mathrm{Im}(x) > t\}$. Under this topology, every $\gamma \in SL(2,\mathbb{Z})$ gives rise to a homeomorphism of $\mathscr{H} \cup \mathbb{Q} \cup \{\infty\}$. The space $X(\Gamma)$ with the quotient topology is the candidate for our compactification. We will show, in the following order, that $X(\Gamma)$ is Hausdorff, connected, and a Riemann surface. If Γ has finite index in $SL(2,\mathbb{Z})$, then we show that $X(\Gamma)$ is compact.

We should note that, with a correspondingly generalised definition of the cusps, $X(\Gamma)$ is a connected Riemann surface under the weaker assumption that Γ is merely a discrete subgroup of $SL(2,\mathbb{R})$. However, in this case the proofs become much more tedious, and there is very little to gain since we will only study subgroups of $SL(2,\mathbb{Z})$ anyway.

Proposition 1.1.1.12. *For Γ a discrete subgroup of $SL(2,\mathbb{Z})$, $X(\Gamma)$ is connected, second countable and Hausdorff.*

We require a lemma first.

Lemma 1.1.1.13. *For any γ in $SL(2,\mathbb{Z})$ and $x \in \mathscr{H}$,*

$$\mathrm{Im}(\gamma x) \leq \max\{\mathrm{Im}(x), 1/\mathrm{Im}(x)\}.$$

Proof. A straightforward calculation shows that

$$\mathrm{Im}(\gamma x) = \frac{\mathrm{Im}(x)}{|cx+d|^2},$$

where $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ as usual. If $c = 0$, then $ad = 1$, so $\mathrm{Im}(\gamma x) = \mathrm{Im}(x)$. If $c \neq 0$, then $|cx+d|^2 \geq |c|^2 \mathrm{Im}(x)^2$. Since $c \geq 1$, we have $\mathrm{Im}(\gamma x) \leq 1/\mathrm{Im}(x)$ in this case. □

Proof of Proposition 1.1.1.12. Connectedness and second countability are easier: if $\mathscr{H} \cup \mathbb{Q} \cup \{\infty\}$ is a disjoint union of open set U_1 and U_2, then upon taking the intersection with the connected set $\mathscr{H}$, we have $\mathscr{H} \subseteq U_1$ and $U_2 \subseteq \mathbb{Q} \cup \{\infty\}$. So if U_2 is open, it must be empty. Since $X(\Gamma)$ is a quotient of $\mathscr{H} \cup \mathbb{Q} \cup \{\infty\}$, it must be connected too. For second countability, we take

$$\{\text{open disks in } \mathscr{H} \text{ of rational radii}\} \cup \{\gamma \cdot (N_t \cup \{\infty\}) \mid \gamma \in \Gamma, t \in \mathbb{Q}\}$$

for a countable base.

Now we show that $X(\Gamma)$ is Hausdorff. We know that $\Gamma \backslash \mathscr{H}$ is Hausdorff by Corollary 1.1.1.10, so we need to show that we can separate cusps from interior points and Γ-inequivalent cusps from each other. First, we separate cusps from points of $\Gamma \backslash \mathscr{H}$. Suppose that $z_1 = \Gamma x_1$ and $z_2 = \Gamma x_2$ with $x_1 \in \mathscr{H}$ and $x_2 \in \mathbb{Q} \cup \{\infty\}$. Since $\Gamma \backslash \mathscr{H}$ is locally compact, we find a neighbourhood U_1 of x_1 with compact closure K. For some $\gamma \in SL(2,\mathbb{Z})$, $x_2 = \gamma\infty$. By Lemma 1.1.1.13, for t sufficiently large $SL(2,\mathbb{Z})K \cap N_t = \varnothing$. If we take $U_2 = \gamma(N_t \cup \{\infty\})$, then $\pi(U_1)$ and $\pi(U_2)$ are disjoint.

Now suppose that $z_1 = \Gamma x_1$ and $z_2 = \Gamma x_2$ are two cusps in different orbits of Γ. Then $x_1 = \gamma_1 \infty$ and $x_2 = \gamma_2 \infty$ for some $\gamma_1, \gamma_2 \in SL(2,\mathbb{Z})$. Set $U_1 = \gamma_1(N_2 \cup \{\infty\})$ and $U_2 = \gamma_2(N_2 \cup \{\infty\})$. Then U_1 and U_2 are disjoint: otherwise, $\kappa\gamma_1 n_1 = \gamma_2 n_2$ for some $\kappa \in \Gamma$ and $n_1, n_2 \in N_2$. Written differently, the element $\gamma_2^{-1}\kappa\gamma_1 \in SL(2,\mathbb{Z})$ takes n_1 to n_2, but by the proof of Lemma 1.1.1.13, unless the lower left entry of $\gamma_2^{-1}\kappa\gamma_1$ is zero, $\mathrm{Im}(\gamma_2^{-1}\kappa\gamma_1 n_1) \leq 1/\mathrm{Im}(n_1)$, which implies that $n_2 \notin N_2$. So there is some integer m such that

$$\gamma_2^{-1}\kappa\gamma_1 = \pm\begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix}.$$

All such matrices fix ∞, so $\kappa x_1 = x_2$, which contradicts our assumption that x_1 and x_2 are in different Γ-orbits. □

We remarked above that if Γ acts freely, then our assertions are very easy. Although the action of a discrete subgroup Γ of $SL(2,\mathbb{Z})$ on $\mathscr{H}$ is not generally free, the quotient $\Gamma \backslash \mathscr{H}$ admits a smooth manifold structure. The feature of the action which permits the existence of this structure on the quotient is revealed in the next result.

Lemma 1.1.1.14. *For each $x \in \mathscr{H}$, the isotropy subgroup G_x from Definition 1.1.1.3 is a finite cyclic group.*

Proof. Pick $\kappa \in SL(2,\mathbb{R})$ so that $\kappa i = x$, and note that the isotropy subgroup of $SL(2,\mathbb{R})$ at i is $SO(2)$. Then

$$\{\gamma \in \Gamma \mid \gamma x = x\} = \kappa SO(2)\kappa^{-1} \cap \Gamma.$$

Since Γ is discrete and $SO(2)$ is compact, this intersection must be finite. But $SO(2)$ is isomorphic to $\mathbb{R}/\mathbb{Z}$, which has the property that all its finite subgroups are cyclic. □

We use the following terminology: a point in $\mathscr{H}$ is called *elliptic* if its *reduced* isotropy subgroup is nontrivial. The concept of ellipticity is well-defined on orbits, so we may equally well speak of an elliptic point of $\Gamma \backslash \mathscr{H}$.

Proposition 1.1.1.15. *If Γ is a discrete subgroup of $SL(2,\mathbb{Z})$, then $X(\Gamma)$ is a Riemann surface.*

Proof. We know that $X(\Gamma)$ is Hausdorff, connected and second countable by Proposition 1.1.1.12. Recall that by Lemma 1.1.1.9, for every $x \in \mathscr{H}$, there exists a neighbourhood U of x such that $\{\gamma \in \Gamma \mid \gamma U \cap U \neq \varnothing\} = \Gamma_x$. If $x \in \mathscr{H}$ and we set U to be the U from Lemma 1.1.1.9, then $\pi\big|_U : U \to \Gamma_X \backslash U$ is a homeomorphism, so we include $\{U, \left(\pi\big|_U\right)^{-1}\}$ into our complex atlas. One easily checks that if x_1 and x_2 are two points such that $U_1 \cap U_2 \neq \varnothing$, then

$$\pi\big|_{U_2} \circ \left(\pi\big|_{U_1}\right)^{-1} : \pi(U_1 \cap U_2) \to \pi(U_1 \cap U_2)$$

is the identity map, so it is holomorphic.

If $x \in \mathscr{H}$ is an elliptic point, then more work is necessary. Let λ be a biholomorphism from $\mathscr{H}$ onto $\mathbb{D}$ such that $\lambda(x) = 0$. By Lemma 1.1.1.14, $\overline{G_x}$ is finite cyclic. Suppose it has order n and let $\overline{\gamma}$ be a generator. For each $m \in \mathbb{Z}$, $\left(\lambda\overline{\gamma}^m\lambda^{-1}\right)^n$ is the identity, so

$$\lambda\overline{G_x}\lambda^{-1} = \{w \mapsto e^{\frac{2\pi i m}{n}} w \mid m = 0, \ldots, n-1\}.$$

Define a map $p : G_x \backslash U \to \mathbb{C}$ by $(p \circ \pi)(x) = \lambda(x)^n$. It is easily checked that p is a homeomorphism onto an open subset of $\mathbb{C}$, so we include $(\Gamma_x \backslash U, p)$ in our complex atlas. One may check that this chart is compatible with the charts for non-elliptic points as well as charts for other elliptic points.

Finally, suppose that x is a cusp, and choose $\kappa \in SL(2,\mathbb{Z})$ such that $\kappa x = \infty$. Then since $SL(2,\mathbb{Z})_\infty$ is an infinite cyclic group, there is an integer $h > 0$ such that

$$G_X(\kappa\Gamma_\infty\kappa^{-1}) = G_X\left\langle\begin{pmatrix}1 & h\\ 0 & 1\end{pmatrix}\right\rangle.$$

We define a homeomorphism p from $\Gamma_x \backslash U$ into an open subset of $\mathbb{C}$ by $(p \circ \pi)(x) = e^{2\pi i\kappa x/h}$, and include the chart $(\Gamma_x \backslash U, p)$ in our complex structure. We leave to the reader the details of checking that charts for cusps are compatible with each other and with the existing charts [DS05, pp. 59–62]. □

Proposition 1.1.1.16. *Suppose that Γ is not only a discrete subgroup of $SL(2,\mathbb{R})$, but also has finite index in $SL(2,\mathbb{Z})$. Then $X(\Gamma)$ is compact.*

Proof. We require the concept of a fundamental domain. A fundamental domain $\mathscr{F}$ for a discrete subgroup Γ of finite index in $SL(2,\mathbb{Z})$ is a connected subset of $\mathscr{H}$ that meets each orbit of Γ at a single point. Fundamental domains have been studied extensively [DS05, pp. 54 and pp. 69; For29; Kna92, pp 228 and pp. 260; Kra72; Leh64] and it is well-known that the set

$$\mathscr{F} = \left\{x \in \mathscr{H} \mid \mathrm{Re}(x) < \frac{1}{2}, |x| > 1\right\} \cup \left\{-\frac{1}{2} + it \in \mathscr{H} \mid t > \frac{\sqrt{3}}{2}\right\} \cup \left\{e^{i\theta} \mid \frac{\pi}{2} \leq \theta \leq \frac{2\pi}{3}\right\}$$

is a fundamental domain for $SL(2,\mathbb{Z})$ (see also the diagram in 2.1.4.6). There is only one cusp under the action of $SL(2,\mathbb{Z})$ on $\mathscr{H}$, and it is easily checked that $\mathscr{F}^* = \mathscr{F} \cup \{\infty\}$ is compact under the topology on $\mathscr{H}^* = \mathscr{H} \cup \mathbb{Q} \cup \{\infty\}$. But

$$\mathscr{H}^* = SL(2,\mathbb{Z})\mathscr{F}^* = \cup_i \Gamma\gamma_i\mathscr{F}^*,$$

where γ_i are representatives of the distinct cosets of Γ in $SL(2,\mathbb{Z})$. If Γ has finite index in $SL(2,\mathbb{Z})$, then the union is finite, so

$$X(\Gamma) = \cup_i \pi\left(\gamma_i\mathscr{F}^*\right)$$

is compact. □

One may show that there are only two distinct elliptic points of $\mathscr{H}$ under the action of $SL(2,\mathbb{Z})$ (see Lemma 2.1.4.8. Since we noted above that there is only one cusp for $\mathscr{H}$ under the action of $SL(2,\mathbb{Z})$, it follows that for any subgroup Γ of $SL(2,\mathbb{Z})$ of finite index, there are only finitely many elliptic points and cusps.

In summary, we have proved:

Theorem 1.1.1.17. *Let Γ be a discrete subgroup of $SL(2,\mathbb{R})$ of finite index in $SL(2,\mathbb{Z})$, acting on $\mathscr{H}$ by Möbius transformations. Then $X(\Gamma)$ is a compact connected Riemann surface.*

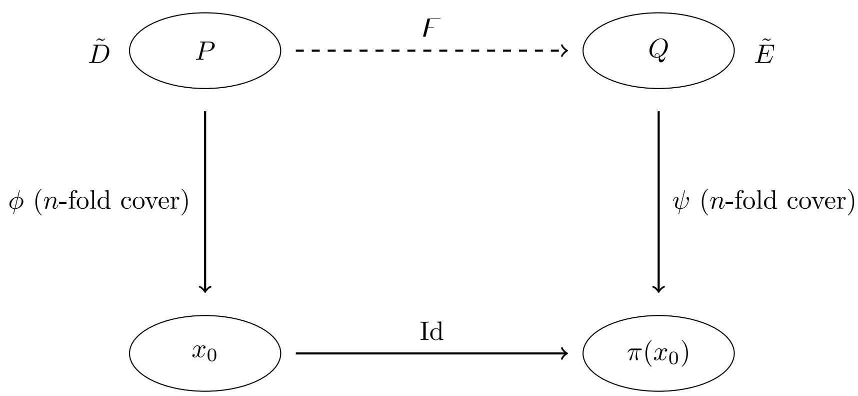
One final remark is in order concerning the proof of Theorem 1.1.1.17. The reader will agree that the proof sketched in this subsection was quite tedious. If the reader will take for granted the validity of some classical differential-geometric arguments, then we can get by with a much quicker, more geometrically intuitive proof.

Alternative proof of Theorem 1.1.1.17. An equivalent definition of a fundamental domain for the action of Γ on $\mathscr{H}$ is as follows: pick a point $x_0 \in \mathscr{H}$. Noting that any subgroup of $SL(2,\mathbb{R})$ preserves to hyperbolic metric $y^{-2}dxdy$ on $\mathscr{H}$, we set $\mathscr{F}$ to be the set of all points which have the property that their distance to x_0 is minimal compared to that of their translates under Γ.

The boundary of $\mathscr{F}$ is a finite sequence of piecewise smooth geodesics for $\mathscr{H}$: these are generalised arcs of circles, and straight vertical lines if we include the circles passing through the point at infinity. The action of Γ produces a tessellation of $\mathscr{H}$ by copies of $\mathscr{F}$. The identification of the boundary of $\mathscr{H}$ produces a Hausdorff topological space away from the cusps and elliptic points, and we will use the tessellation of $\mathscr{H}$ to produce charts.

As in the first proof of Theorem 1.1.1.17, it is easy to show that there are charts at the non-elliptic points. If one of these points happens to lie on the boundary of $\mathscr{H}$, then we may find a small enough neighbourhood so that its translates do not intersect each other.

At an elliptic point x_0, we know that the isotropy subgroup is finite cyclic, so we may select a small enough disk about x_0 such that the projection onto $\Gamma\backslash\mathscr{H}$ is an n-fold cover for some positive integer n. The map $z \to z^n$ is also an n-fold covering of the disk, so if we can show that there is a map F making the diagram below commute, we can infer the existence of a complex chart at x_0.



Pick a point P in $\tilde{D}$, and a point Q in $\tilde{E}$ such that $\phi(P) = \psi(Q)$. Then for a sufficiently small path γ_P through P, we may choose one of n small paths γ_Q through Q such that the compatibility condition $\phi(\gamma_P) = \psi(\gamma_Q)$ holds. If γ_P is then extended, there is a unique path extending γ_Q satisfying the compatibility condition. Therefore we may use the paths through P to define F. ☐

For more about branched coverings of Riemann surfaces, and in particular for the theorem that any compact Riemann surface admits a representation as a branched cover of $\mathbb{CP}^1$, we refer the reader to Gunning [Gun66, Section 10 (a)].

1.1.2 Vector bundles and connections

A motivation for the study of vector bundles on manifolds comes from physics, where "space" is formalised as a manifold and the classical notion of a "field" (a map associating measurements to points in the space) is formalised as a section of a vector bundle. Intuitively, a vector bundle is the assignment of a vector space to each point on the manifold in such a way that the vector spaces vary smoothly as the point varies smoothly. We could, for example, model the magnetic field generated by the Earth as a vector field on the two dimensional sphere, associating to each point on the sphere a vector in a 2-dimensional real vector space pointing in the direction a charged particle would follow for a short time if placed at that point, and whose magnitude represents the strength of the field (which determines how fast the particle would move). However, the map described in this way does not take values in $\mathbb{S}^2 \times \mathbb{R}^2$: there is no way to consistently identify the vector space attached to the North pole with the vector space attached to the South pole.

The object in which this map *does* take its values is called a *vector bundle*, and the vector field, as a smooth map from the manifold into the vector bundle, is an example of a *section*. The study of vector bundles is essential to physics, but is also of great importance in pure mathematics, in which vector bundles have proven particularly useful in "linearising" nonlinear phenomena. It is well-known (but by no means well understood) that deep results on the topology of manifolds may be obtained by studying vector bundles.

Vector bundles form an essential part of the study of modular forms, which may be viewed as sections of bundles over quotients of locally symmetric spaces, such as the space $X(\Gamma)$ constructed in Subsection 1.1.1. In the case that the underlying manifold is compact, the vector space of sections of a vector bundle satisfying some sort of positivity condition (such as being in the kernel of a linear elliptic differential operator), together with restrictions on the severity of the possible singularities of the sections, is finite dimensional [GS93; Wel86, Theorem 5.2]. The fact that the space of modular forms — to be stated more precisely in Subsection 2.1.1 — is finite dimensional, is of fundamental importance.

For a comprehensive treatment of vector bundles in geometry, there exist a number of well-known accounts [Lee12, Chapter 10; Har77, pp. 128, Chapters II, III and IV; Huy05, p. 72; Wel86, Chapter I, Chapter III].

In order to be able to apply tools from analysis to vector bundles with a minimum of fuss, we require that from now on, the fields underlying our vector spaces are $\mathbb{R}$ or $\mathbb{C}$. In the upcoming definition, k stands for $\mathbb{R}$ or $\mathbb{C}$.

Definition 1.1.2.1. *A continuous map $\pi : E \to X$ of Hausdorff spaces is called a k-vector bundle of rank r if the following conditions are satisfied:*

1. *For each $p \in X$, the fibre $E_p = \pi^{-1}(p)$ is a k-vector space of dimension r.*
2. *For each $p \in X$, there is a neighbourhood U of p and a homeomorphism $h : \pi^{-1}(U) \to U \times k^r$ such that $h(E_p) \subseteq \{p\} \times k^r$, and the composition*

$$E_p \underset{h}{\to} \{p\} \times k^r \xrightarrow[\text{proj.}]{} k^r$$

is an isomorphism of k-vector spaces. The pair (U, h) is called a local trivialisation of E at p.

We have just defined topological vector bundles. A vector bundle E is called a smooth vector bundle if E and X are smooth manifolds, π is a smooth map, and every h is a diffeomorphism. Similarly, E is a holomorphic vector bundle if E and X are complex manifolds, π is a holomorphic map and every h is a biholomorphism. Note that a holomorphic vector bundle must have even rank as an $\mathbb{R}$-vector bundle and each fibre inherts a canonical complex vector space structure; the phrase "complex vector bundles" refers to vector bundles with $k = \mathbb{C}$: they may be defined over topological manifolds and do not (necessarily) possess a complex structure.

Operations valid in the category of vector spaces tend to correspond to operations on vector bundles. Given two vector bundles E and F, then applying usual operations, we may form new vector bundles such as

1. $E \oplus F$, the direct sum bundle;
2. $E \otimes F$, the tensor product bundle;
3. $\mathrm{Hom}(E, F)$;
4. E^*, the dual bundle;
5. $\otimes^k E$, the tensor product of degree k;
6. $\wedge^k E$, the antisymmetric tensor product of degree k;
7. $\odot^k E$, the symmetric tensor product of degree k.

Each vector bundle is characterised by asserting that its fibres are obtained by applying the vector space operations to the fibres of the original bundles. For example, $(E \oplus F)_p = E_p \oplus F_p$, and $\left(\odot^k E\right)_p = \odot^k E_p$. As mentioned earlier, we are often interested in certain functions from manifolds into bundles. This prompts the following definition.

Definition 1.1.2.2. *Let $\pi : E \to X$ be a topological vector bundle over a topological manifold. A section $s : X \to E$ is a continuous map such that $\pi \circ s$ is the identity on X. For smooth or holomorphic vector bundles we define smooth or holomorphic sections in the obvious way.*

The set of sections of a fixed vector bundle E form a k-vector space, which we denote $\Gamma(X, E)$. The set of sections over an open subset U of X is defined in the obvious way, and is denoted $\Gamma(U, E)$. The assignment $U \mapsto \Gamma(U, E)$ gives rise to a sheaf of k-vector spaces on X. The correspondence between sheaves and vector bundles is quite subtle [Har77, pp. 128; Huy05, pp. 72].

Remark 1.1.2.3. *A useful shorthand notation for the space of global sections of a vector bundle E over X is $\mathscr{E}(X)$: in other words, we use Roman typeface for bundles and script forms for sections. When X is a complex manifold, there are two particularly important sheaves that comes with X: the sheaf of holomorphic functions, denoted $\mathscr{O}(X)$, and the sheaf of meromorphic functions (holomorphic functions $f : X \to \mathbb{CP}^1$), denoted $\mathscr{K}(X)$. For any holomorphic vector bundle E, we have the sheaf of meromorphic functions with values in E, obtained as $\mathscr{E}(X) \otimes_{\mathscr{O}(X)} \mathscr{K}(X)$. If E is the bundle of differential k-forms, then sheaf of meromorphic differential k-forms is denoted $\Omega^k(X)$. The notation is a tradition from algebraic geometry.*

Next, we list the bundles that interest us the most. Note that although we only speak of smooth manifolds and smooth vector bundles, the first three examples all have obvious holomorphic counterparts.

1. The trivial bundle. For an n-dimensional smooth manifold X, the product manifold $X \times \mathbb{R}^n$ together with the projection $(x, X) \mapsto x$ is a vector bundle on X. The sheaf of global sections of the trivial bundle is denoted $\mathscr{C}^\infty(X)$.

2. The tangent bundle. Suppose that X is a smooth manifold, and to each point $p \in X$ we associate the $\mathbb{R}$-vector space of derivations of germs of smooth functions at p. Each such vector space is called the tangent space of X at p: it is denoted T_pX and is the fibre at p of the tangent bundle TX.

   A fundamental fact is that, given a smooth map $f : X \to Y$ between two smooth manifolds, there is a smooth map $df : TX \to TY$. This generalises the ordinary notion of derivative from calculus.

3. The cotangent bundle. In the smooth and complex cases, this is characterised by asserting that $(T^*X)_p = T^*_pX$. We use the simpler notation $\wedge^1 X$ to refer to this bundle and $\Omega^1(X)$ to refer to the sheaf of (global) sections. Sections of this bundle may be written *locally* as $f_1(x)dx^1 + \cdots + f_n(x)dx^n$, where x is a local coordinate on X and n is the dimension of the manifold.

4. The bundle of differential k-forms. We let $\wedge^k X = \wedge^k T^*X$, and denote the sheaf of global sections of this bundle by $\Omega^k(X)$. A section of $\wedge^k X$ may be written locally as $\sum_{1 \le i_1 < \cdots < i_k} f_{i_1,\ldots,i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}$, and there are canonically defined $\mathbb{R}$-linear maps, all denoted d, such that

$$0 \longrightarrow \mathscr{C}^\infty(X) \xrightarrow{d} \Omega^1(X) \xrightarrow{d} \Omega^2(x) \xrightarrow{d} \ldots \xrightarrow{d} \Omega^n(X) \longrightarrow 0$$

   is a complex (that is, $d \circ d = 0$ for any two consecutive ds)

In preparation for the next subsection, we define connections, geodesics and Riemannian manifolds.

Definition 1.1.2.4. *Let $\pi : E \to X$ be a smooth vector bundle over a smooth manifold X, and write $\mathscr{E}$ for its sheaf of sections. A* connection *∇ on E is a smooth $\mathbb{R}$-linear map*

$$\nabla : \mathscr{E} \to \Omega^1(X) \otimes \mathscr{E}$$

such that, for all smooth function f on X and smooth sections $s \in \mathscr{E}$, the Leibnitz rule *holds:*

$$\nabla(fs) = f\nabla(s) + \nabla(f) \otimes s.$$

If ∇ is a connection on TX, then a smooth path $\gamma : [a, b] \subseteq \mathbb{R} \to X$ is said to be a *geodesic* for ∇ if $\nabla(Z, Z) = 0$, where Z is a smooth vector field[2] defined on a neighborhood of $\gamma([a, b])$ such that $Z_{\gamma(t)} = d\gamma\big|_t\left(\frac{d}{ds}\right)\big|_t$ for all $t \in [a, b]$ ($\frac{d}{ds}$ is the canonical vector field on $\mathbb{R}$)

Definition 1.1.2.5. *A Riemannian manifold X is a smooth manifold together with a smooth global section g of $\left(\odot^2 TX\right)^*$ — that is, a smoothly varying family of smooth maps $g_p : T_pX \odot T_pX \to \mathbb{R}$ — satisfying the* positive-definiteness *condition:*

$$g_p(X, X) \ge 0, \text{ with equality if and only if } X = 0.$$

[2]The vanishing (or not) of $\nabla(Z, Z)$ is independent of the choice of Z.

If ∇_E and ∇_F are connections on smooth vector bundles E and F over a smooth manifold X, then we obtain induced connections on all the vector bundles mentioned earlier which were functorially constructed from E and F. In particular, we from a connection on TX we obtain a connection on $\left(\odot^2 TX\right)^*$. If (X, g) is a Riemannian manifold then there exists a unique connection ∇ on TX, the *Levi-Civita connection*, satisfying

1. $\odot^2\nabla^* g = 0$, where $\odot^2\nabla^*$ is the induced connection on $\left(\odot^2 TX\right)^*$, and
2. $\mathrm{Proj}\big|_{\Omega^2(X)}(\nabla^*) = d$, where ∇^* is the induced connection on T^*X and Proj is the projection map from $\Omega^1(X)\otimes\Omega^1(X)$ to $\Omega^2(X)$.

1.1.3 Projective and conformal structures

In practice one deals not with plain smooth or complex manifolds, but with manifolds together with extra structure. A notable example of a class of interesting structures one may attach to manifolds is described by *Cartan geometry*. We will not describe the general theory, except to say that it is a sort of "global" continuation of Klein's Erlangen program, and to point out that there exist comprehensive studies [vS09, Chapter 4; Eas08; Kob95; OT05]

In Chapter 3, we will be dealing with two specific Cartan geometries: projective and conformal structures. We delve no further into either topic than is necessary two present the two results which we will require.

Definition 1.1.3.1. *Let X be an oriented smooth manifold. Two connections on TX are said to be* projectively equivalent *if they share the same unparameterised geodesics. A* projective structure *on X is a class of projectively equivalent connections on TX.*

Proposition 1.1.3.2. *A projective structure on an oriented complex 1-dimensional manifold X gives rise to a natural class of coordinates on X on which the group $SL(2,\mathbb{C})$ acts freely and transitively.*

Proof. Since X is oriented, we may select positive transition functions for $\wedge^1$. It is possible to build a bundle with the property that its positive transition functions are the square roots of those for $\wedge^1$. We denote this bundle by $\mathcal{E}(-1)$, and for integral $n > 0$ we define $\mathcal{E}(-n)$ to be the bundle whose transition functions are nth powers of the transition functions of $\mathcal{E}(-1)$. We define $\mathcal{E}(0)$ to be the trivial bundle, the sections of which are the complex-valued functions on X.

The usual exterior derivative d induces projectively invariant differential operators:

$$\Delta^{k+1} : \mathcal{E}(k) \to \mathcal{E}(-k-2),$$

so there is a projectively invariant differential operator $\mathscr{S} : \mathcal{E}(0) \to \mathcal{E}(-3)$ defined by $\mathscr{S}(f) = \Delta^2(\Delta(f))^{-\frac{1}{2}}$. One may easily check that

$$\mathscr{S}(f) = -4\Delta(f)^{-5/2}\left(2\Delta(f)\Delta^3(f) - 3(\Delta^2(f))^2\right).$$

So if $\mathscr{S}(f) = 0$, $2\Delta(f)\Delta^3(f) - 3(\Delta^2(f))^2 = 0$, and for any coordinate t on X, one checks that

$$f(t) = \frac{at+b}{ct+d}$$

for some $\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right) \in SL(2,\mathbb{C})$. In other words, the set of solutions to the operator $\mathscr{S}$ gives rise to a preferred class of coordinates for X, related by Möbius transformations, and conversely, a preferred class of coordinates related by Möbius transformations gives rise to a projective structure. □

The differential operator $\mathscr{S}$ arising in the proof of Proposition 1.1.3.2 is called the *Schwartzian*. It has many remarkable properties [OT05; OT09].

Definition 1.1.3.3. *Let X be a smooth manifold. Two Riemannian metrics g and $\hat{g}$ on TX are said to be* conformally equivalent *if there exists a smooth positive function $\Omega : X \to \mathbb{R}$ such that $g = \Omega^2\hat{g}$. A* conformal structure *on X is an equivalence class of conformally equivalent Riemannian metrics on X.*

Proposition 1.1.3.4. *Let X be a smooth oriented 2-manifold. Then X admits a complex structure if and only if it admits a conformal structure.*

Proof. Let X be a 2-manifold with a complex structure. For any $x \in X$, there is a trivialisation $\phi_x : T_xX \to \mathbb{C}$, which induces an inner product of T_xX defined by declaring 1 and i to be orthogonal and of unit length. Since X admits a complex structure, this inner product is independent of the choice of trivialisation and using a partition of unity, it extends smoothly to an inner product g on TX. Any other inner product for which 1 is orthogonal to i is of the form $\Omega^2 g$ for some smooth positive real-valued function Ω, so the complex structure generates a unique conformal structure.

Suppose that X is an oriented 2-manifold with a conformal structure. According to a theorem of Korn [Kor14] and Lichtenstein [Lic16], with a simpler proof by Chern [Che55], every smooth Riemannian 2-manifold admits local coordinates which are isothermal; that is, for every $x \in X$, there exists a chart (U, ϕ) about x such that $\phi^*(g) = f^2(dx^2 + dy^2)$, for some smooth positive real-valued f. We also stipulate that ϕ preserves the orientation of X. Indeed, we claim that our complex coordinates on U are given by $z = x + iy$.

If two of these isothermal coordinates, ϕ and ψ, are defined on overlapping open sets, then the composition $\psi \circ \phi^{-1}$ is a conformal orientation-preserving diffeomorphism between open subsets of $\mathbb{R}^2$. But any such map is a biholomorphism between open subsets of $\mathbb{C}$, so we have a complex structure. If we apply the same procedure to a Riemannian manifold conformally equivalent to the one we began with, then our f is rescaled, and so our conformal orientation-preserving diffeomorphism is rescaled by a smooth positive real-valued function. It follows that the complex structure is an invariant of the conformal class of X. □

An equivalent way to generate a unique complex structure from a conformal structure is to observe that a conformal structure gives rise to a unique *almost-complex structure* [Wel86, Chapter I, Section 3]. This almost-complex structure is integrable, since $\overline{\partial}\Omega_{\mathbb{C}}^{p,q} \subseteq \Omega_{\mathbb{C}}^{p,q+2} = 0$ for $p, q \geq 0$, so $\overline{\partial} = 0$. By the celebrated theorem of Newlander–Nirenberg [Hör73, Chapter 5, Section 7; NN57], there exists a unique complex structure inducing a given integrable almost-complex structure.

1.2 Fourier analysis

We begin by collecting the basic facts about Fourier analysis on finite abelian groups, which we will use throughout to investigate exponential sums over Dirichlet characters, and we also cover the most important aspects of Fourier analysis on $\mathbb{R}^n$. Indeed, Fourier analysis on $\mathbb{R}^n$ plays a fundamental role in the theory of *theta functions* and in the arithmetic of squares through a fundamental formula relating the sum of the values of a function at integer points to the sum of the values of its Fourier transform at integer points. The heart of the proof of the Poisson summation is in the representability of a function by its Fourier series, and variations of the Poisson summation abound, each with a different set of conditions on the function in question. We stick to a particularly common set of hypotheses on our function, namely that it be smooth and rapidly decreasing near infinity.

We subsequently define a fairly general class of theta functions and prove their all-important functional equations, anticipating the requirements of Chapters 3, 4 and 5. We allow the theta function of Chapter 2 the glory of its very own proof of the functional equation in Subsection 2.2.1, as we will be concerned in that chapter with the special relationship between this theta function and the Riemann zeta function.

1.2.1 Fourier analysis on finite abelian groups

We develop the minimum of harmonic analysis on finite abelian groups necessary to deal properly with Dirichlet characters. The presentation here is based on notes of Conrad [Con].

Let G be a finite abelian group. A character of G is a group homomorphism of G into $\mathbb{S}^1$, the complex numbers of unit modulus. The set of all characters of G themselves form a finite abelian group, denoted $\hat{G}$, with the group operation characterised by

$$(\chi + \chi')(n) = \chi(n)\chi'(n).$$

There is a natural Hermitian inner product on the complex vector space $\mathbb{C}[G]$ of complex-valued functions on G, defined by

$$\langle f_1, f_2 \rangle = \sum_{g \in G} f_1(g)\overline{f_2(g)}.$$

Lemma 1.2.1.1. *Suppose G is a finite abelian group and H is a subgroup. Then every character of H extends to a character of G.*

Proof. Without loss of generality, assume $H \subsetneq G$. Then pick some $a \in G \setminus H$, and let $k \geq 1$ be the smallest integer such that $a^k \in H$. Define a function $\tilde{\chi}$ on $\langle H, a\rangle$ by setting $\tilde{\chi}(a)$ equal to any of the complex numbers z such that $z^k = \chi(a^k)$, and letting $\tilde{\chi}(ha^i) = \chi(h)\tilde{\chi}(a)^i$ for any ha^i in H.

We must check that this function is well-defined: if $ha^i = ha^{i'}$, then $a^{i-i'} \in H$, so $i = i' \bmod k$. Write $i = i' + kq$, and note that $h = h'a^{i'-i} = h'a^{kq}$. Then

$$\begin{aligned}\chi(h'a^{i'}) &= \chi(h')\tilde{\chi}(a)^{i'}\\ &= \chi(h')\tilde{\chi}(a)^{i}\tilde{\chi}(a^{kq})\\ &= \chi(h'a^{kq})\tilde{\chi}(a)^{i}.\end{aligned}$$

It's clear that $\tilde{\chi}$ is a homomorphism, so $\tilde{\chi}$ is a character on $\langle H, a\rangle$ which restricts to χ on H. But $[G : \langle H, a\rangle] < [G : H]$, so the claim follows by induction on the index. □

The point of this subsection is the next result, known as the Fourier inversion formula.

Proposition 1.2.1.2. *For any function $f \in \mathbb{C}[G]$, we have*

$$f(g) = \frac{1}{|G|}\sum_{\chi\in\hat{G}}\left(\sum_{h\in G} f(h)\overline{\chi(h)}\right)\chi(g). \tag{1.2}$$

Proof. We may interchange the order of summation, as both groups are finite. The claim then follows if we can prove that

$$\sum_{\chi\in\hat{G}}\chi(g-h) = \begin{cases}|G| & g = h\\ 0 & g \neq h.\end{cases} \tag{1.3}$$

We need to prove that if $g \neq 0$, then $\sum_{\chi\in\hat{G}}\chi(g) = 0$. Let H be the cyclic subgroup of G generated by g. Then there is a non-canonical isomorphism of H with $\mu_n^\times \subset \mathbb{S}^1$, and this isomorphism gives rise to a Dirichlet character of H which doesn't map g to 1. By Lemma 1.2.1.1, there is a Dirichlet character χ_0 of G which doesn't map g to 1.

So

$$\chi_0(g) + \sum_{\chi\in\hat{G}}\chi(g) = \sum_{\chi\in\hat{G}}(\chi_0+\chi)(g) = \sum_{\chi\in\hat{G}}\chi(g),$$

and since $\chi_0(g) \neq 1$, we must have $\sum_{\chi\in\hat{G}}\chi(g) = 0$. □

We now apply Proposition 1.2.1.2 to $G = (\mathbb{Z}/n\mathbb{Z})^\times$ and the function $\delta_{k,n}$ on $(\mathbb{Z}/n\mathbb{Z})^\times$ defined by

$$\delta_{k,n}(m) = \begin{cases}1 & m = k \bmod n\\ 0 & m \neq k \bmod n.\end{cases}$$

The characters of $(\mathbb{Z}/n\mathbb{Z})^\times$ are exactly the Dirichlet characters of order n, and each such character may be extended to a function on $\mathbb{Z}$ by the decree that if $(k, n) > 1$, then $\chi(k) = 0$. We remind the reader that when $n = 1$, the trivial Dirichlet character has $\chi(1) = 1$, but when $n > 1$, $\chi(n) = 0$ for all χ. In any case, when $(k, n) = 1$ we have

$$\delta_{k,n}(m) = \frac{1}{\phi(n)}\sum_{|\chi|=n}\overline{\chi(k)}\chi(m),$$

where the sum is over all Dirichlet characters of order n.

1.2.2 Poisson summation

The essence of Poisson summation is the following identity:

$$\sum_{n\in\mathbb{Z}^m} f(n) = \sum_{n\in\mathbb{Z}^m}\hat{f}(n), \tag{1.4}$$

where $\hat{f}$ is the Fourier transform of f, described at 1.5 below. However, 1.4 is only valid for certain classes of functions f. In this subsection, we prove 1.4, without extracting too many rabbits from our hat.

Definition 1.2.2.1. *Let $f : \mathbb{R} \to \mathbb{C}$ be a function. For each $I \subseteq \mathbb{R}$, let*

$$\mathscr{P} = \{P = \{x_0, \ldots, x_{n_P}\} \mid P \subseteq I, x_i \leq x_{i+1} \text{ for } 0 \leq i < n_P\}$$

be the collection of partitions of I. Set

$$V_f(I) = \sup_{P \in \mathscr{P}} \sum_{i=0}^{n_P - 1} |f(x_{i+1}) - f(x_i)|.$$

Then we say that f is of bounded variation *if $V_f(I) < \infty$ for all closed intervals $I \subset \mathbb{R}$.*

Lemma 1.2.2.2. *[Zyg88, Theorem 8.14] Suppose that $f : \mathbb{R} \to \mathbb{C}$ is 1-periodic and of bounded variation. Suppose also that for all $x \in \mathbb{R}$, $f(x) = \frac{1}{2} \lim_{\epsilon \to 0} (f(x+\epsilon) + f(x-\epsilon))$. Then the Fourier series of f converges pointwise to f:*

$$f(x) = \sum_{n \in \mathbb{Z}} c_n(f) e^{2\pi i n x},$$

where

$$c_n(f) = \int_0^1 f(x) e^{-2\pi i n x} dx.$$

Definition 1.2.2.3. *Let $(\alpha) \in \mathbb{N}^m$. For $x \in \mathbb{R}^m$, define $|(\alpha)| = \alpha_1 + \cdots + \alpha_m$, $x^\alpha = x_1^{\alpha_1} \ldots x_m^{\alpha_m}$ and for $f : \mathbb{R}^m \to \mathbb{C}$ sufficiently differentiable, define*

$$(D^{(\alpha)} f)(x) = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_m}}.$$

Then the Schwartz space of functions on $\mathbb{R}^m$ is defined to be

$$S(\mathbb{R}^m) = \left\{ f \in \mathscr{C}^\infty(\mathbb{R}^m) \mid \sup_{x \in \mathbb{R}^m} |x^{(\alpha)} (D^{(\beta)} f)(x)| < \infty \text{ for all } (\alpha), (\beta) \in \mathbb{N}^m \right\}.$$

Let $f \in S(\mathbb{R}^m)$. If we hold all but one of the variables constant, then f satisfies the hypotheses of Lemma 1.2.2.2 as a function of one variable, and upon repeated application of Lemma 1.2.2.2 it follows that we may represent f as a Fourier series of several variables:

$$f(x) = \sum_{n \in \mathbb{Z}^m} c_n(f) e^{2\pi i \langle n, x \rangle},$$

where the sum converges absolutely on compact subsets and

$$c_n(f) = \int_{[0,1]^m} f(x) e^{-2\pi i \langle n, x \rangle} dx.$$

We now present the Fourier transform of a function $f : \mathbb{R}^m \to \mathbb{C}$, traditionally denoted $\hat{f}$:

$$\hat{f}(x) = \int_{\mathbb{R}^m} f(t) e^{2\pi i \langle x, t \rangle} dt. \tag{1.5}$$

The Fourier transform exists (in the sense that the integral defining $\hat{f}$ converges) whenever, for example, $f \in L^1(\mathbb{R}^m)$.

Proposition 1.2.2.4 (Poisson summation)**.** *Let $f \in S(\mathbb{R}^m)$. Then*

$$\sum_{n \in \mathbb{Z}^m} f(n + x) = \sum_{n \in \mathbb{Z}^m} \hat{f}(n) e^{2\pi i \langle n, x \rangle}. \tag{1.6}$$

Proof. Set $g(x) = \sum_{k \in \mathbb{Z}^m} f(x + k)$. Then since f is a Schwartz function, g is smooth. By 1.2.2.2, the Fourier series of g converges absolutely and pointwise for any $x \in \mathbb{R}$:

$$\sum_{k \in \mathbb{Z}^m} f(x + k) = g(x) = \sum_{k \in \mathbb{Z}^m} a_k(g) e^{2\pi i \langle k, x \rangle}. \tag{1.7}$$

Next, we write the Fourier coefficients of g in terms of f:

$$\begin{aligned} a_k(g) &= \int_{[0,1]^m} g(t)e^{-2\pi i\langle k,t\rangle}dt \\ &= \sum_{n\in\mathbb{Z}^m}\int_{[0,1]^m} f(n+t)e^{-2\pi i\langle k,t\rangle}dt \\ &= \int_{\mathbb{R}^m} f(t)e^{-2\pi i\langle k,t\rangle}dt \\ &= \hat{f}(k), \end{aligned} \tag{1.8}$$

where we used Fubini's theorem to arrive at the second line: since f is Schwartz,

$$\int_{[0,1]^m}\sum_{n\in\mathbb{Z}^m}|f(n+t)e^{-2\pi i\langle n,x\rangle}|dt = \int_{[0,1]^m}\sum_{n\in\mathbb{Z}^m}|f(n+t)|dt < \infty.$$

The expression 1.6 follows upon substituting 1.8 into 1.7. □

1.2.3 Functional equations for theta functions

In this subsection, we use Poisson summation (Proposition 1.2.2.4) to prove some functional equations for classes of theta functions which we will investigate in more detail later. The compulsory comment to make at this point is that the treatment of theta functions with Poisson summation is motivated by the fact that quadratic Gaussian functions tend to be "eigenfunctions" of the Fourier transform. Indeed, if $f_\tau(x) = e^{2\pi i\tau x^2}$, then

$$\hat{f}_\tau(x) = \int_{-\infty}^{\infty} e^{2\pi i\tau t^2 - 2\pi itx}dt = \sqrt{\frac{i}{2\tau}}e^{-2\pi ix^2/4\tau}.$$

Of course, any function of the form $f(x) + \hat{f}(x)$ is invariant under the Fourier transform, but quadratic Gaussians (parametrised by τ) are essentially unique in that their Fourier transform is "not too far" from being periodic in τ. Consequently, if we "average" over a collection of quadratic Gaussians, we obtain a function which is periodic *and* transforms simply under the Fourier transform. This observation has profound consequences.

For any quadratic form represented by an $m\times m$ positive definite real symmetric matrix Q, set

$$\theta_Q(\tau) = \sum_{n\in\mathbb{Z}^m} e^{2\pi i n^tQn\tau}.$$

Then θ_Q may be rewritten as a proper Fourier series:

$$\theta_Q(\tau) = \sum_{n\in\mathbb{Z}}^{\infty} r_Q(n)e^{2\pi in\tau},$$

where $r_Q(n) = \#\{k\in\mathbb{Z}^m \mid k^tQk = n\}$ is the number of representations of n by the quadratic form associated to Q. We will consider a slightly more general class of theta functions, known as *Jacobi forms*. They are defined by

$$\theta_Q(\tau,\omega) = \sum_{n\in\mathbb{Z}^m} e^{2\pi i(n^tQn\tau + n^t\omega)},$$

where Q is a positive definite symmetric $m\times m$ matrix as above, $\tau\in\mathscr{H}$ and $\omega\in\mathbb{C}^n$.

These are actually the functions which were investigated by Jacobi by the first place, and it is possible to view them as seeds for generalisations of modular forms: the *Jacobi forms* have been studied extensively [CS17, Chapter 15, Section 2; Jac29; MM97, Chapter 3].

Applying the Poisson summation formula, we obtain the critical transformation law for $\theta_Q(\tau,\omega)$.

Proposition 1.2.3.1. *Let Q be a positive definite real symmetric matrix. Then for all $\tau\in\mathscr{H}$,*

$$\theta_Q(\tau,\omega) = \frac{e^{-2\pi i\omega^tQ^{-1}\omega/4\tau}}{\sqrt{(-2i\tau)^m|\det(Q)|}}\theta_{Q^{-1}}\left(-\frac{1}{4\tau}, -2Q^{-1}\omega\right),$$

where the branch cut is taken along the negative imaginary axis.

Proof. Since Q is a positive definite real symmetric matrix, there exists an orthogonal matrix U such that

$$U^t Q U = \mathrm{diag}(\lambda_1, \dots, \lambda_m), \ \ \lambda_1, \dots, \lambda_n > 0.$$

We set $D = \mathrm{diag}(\lambda_1, \dots, \lambda_m)$, $\sqrt{D} = \mathrm{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_m})$, and $B = \sqrt{D}U$. Then $Q = B^t B$. We also define

$$f_{\tau,\omega}(x) = e^{2\pi i (x^t Q x \tau + x^t \omega)}.$$

Since $m^t Q m = O\left(||m||^2\right)$, $f_{\tau,\omega}(x)$ is Schwartz in x. Upon making the linear change of variables $y = Bx$,

$$\hat{f}_{\tau,\omega}(z) = \int_{\mathbb{R}^m} e^{2\pi i (x^t Q x \tau + x^t(\omega - z))} dx = \frac{1}{\sqrt{|\det(Q)|}} \int_{\mathbb{R}^m} e^{2\pi i (y^t y \tau + y^t B^{-t}(\omega - z))} dy. \tag{1.9}$$

Now write $B^{-t}(\omega - z) = b = (b_j)$ and $y = (y_j)$: then 1.9 separates into a product of one-dimensional integrals, which are easily evaluated:

$$\frac{1}{\sqrt{|\det(Q)|}} \prod_{j=1}^{m} \int_{\mathbb{R}} e^{2\pi i (y_j^2 \tau + y_j b_j)} dy_j = \frac{1}{\sqrt{(-2i\tau)^m |\det(Q)|}} e^{-2\pi i b^t b / 4\tau}.$$

But $b^t b = (\omega^t - z^t) Q^{-1} (\omega - z)$, so

$$\sum_{n \in \mathbb{Z}^m} \hat{f}_{\tau,\omega}(n) = \frac{e^{-2\pi i \omega^t Q^{-1} \omega / 4\tau}}{\sqrt{(-2i\tau)^m |\det(Q)|}} \sum_{n \in \mathbb{Z}^m} e^{-2\pi i \left(n^t Q^{-1} n / 4\tau - 2 n^t Q^{-1} \omega / 4\tau\right)},$$

and upon applying Poisson summation (Proposition 1.2.2.4), we find that

$$\theta_Q(\tau, \omega) = \sum_{n \in \mathbb{Z}^m} f_{\tau,\omega}(n) = \sum_{n \in \mathbb{Z}^m} \hat{f}_{\tau,\omega}(n) = \frac{e^{-2\pi i \omega^t Q^{-1} \omega / 4\tau}}{\sqrt{(-2i\tau)^m |\det(Q)|}} \theta_{Q^{-1}} \left(-\frac{1}{4\tau}, \frac{Q^{-1}\omega}{\tau} \right),$$

as claimed. □

1.3 Algebraic number fields and zeta functions

Consider the following classical number-theoretic problem, first proposed by Ramanujan and solved by Nagell:

Theorem 1.3.0.1. *The only solutions of the equation*

$$x^2 + 7 = 2^n$$

in integers are the pairs (x, n) *such that* $(\pm x, n) = (1, 3), (3, 4), (5, 5), (11, 7), (181, 15)$.

Although one requires only integers to state the problem, the most natural solution [ST87, Theorem 4.21] involves the arithmetic of

$$\mathbb{Q}(\sqrt{-7}) = \left\{ x + \sqrt{-7} y \mid x, y \in \mathbb{Q} \right\}.$$

The object $\mathbb{Q}(\sqrt{-7})$ is an example of an *algebraic number field*. Algebraic number fields turn out to be the correct setting for investigating all manner of classical problems. We will manage to get by without number fields in Chapters 2, 3 and 4, but the language of algebraic number fields is absolutely essential for Chapter 5, and pervades Chapter 6.

We use this section as an opportunity to introduce some results which, although we will not require the most general forms, expressed in the language of algebraic number theory, belong to the same general circle of ideas. These are the definitions and basic properties of Dirichlet characters and Gauss sums, which are essential to Chapters 4 and 5; the locations of poles and zeros of Dirichlet L-functions, which we will make use of in Subsection 6.1.2, and finally the definition of Mordell–Tornheim zeta functions and multiple zeta functions, together with the recent result [BZ10] that the special values of the Mordell–Tornheim zeta functions may be expressed as rational linear combinations of special values of multiple zeta functions.

A fleeting treatment of reciprocity for the nth power residue symbol is also provided, as this underlies the construction of higher theta functions, referred to extensively in Chapter 6.

1.3.1 The arithmetic of algebraic number fields

In this subsection, we present all the elementary definitions and results which we will need at various points, most especially in Chapter 5, and more covertly in Chapter 6.

Definition 1.3.1.1. *An* algebraic number field *K is a field extension of $\mathbb{Q}$ of finite degree. The degree is denoted $[K:\mathbb{Q}]$.*

The unwieldy term "algebraic number field" is henceforth shortened to "number field". A number field is meant to be an object which captures the essential aspects of $\mathbb{Q}$, but is flexible enough to provide a setting in which one may understand connections between the classical methods used to approach problems involving integers. Part of the package of such a theory is the determination of the object which is to play the role of $\mathbb{Z}$. The right definition turns out to be as follows.

Definition 1.3.1.2. *Let K be a number field. The set $\mathcal{O}_K$ of* algebraic integers *of K is the set of all x in K such that there exists a monic polynomial $f(X) \in \mathbb{Z}[X]$ with the property that $f(x) = 0$.*

More generally, for an arbitrary integral domain A contained in a field L, the set of elements $x \in L$ for which there exists a monic polynomial $f(X) \in A[X]$ such that $f(x) = 0$ are called *integral over A*. One expects that $\mathcal{O}_K$ is a subring of k: this is true, but not so easy to prove as one might hope.

Proposition 1.3.1.3. *Let K be a number field. Then*

1. *$\mathcal{O}_K$ is a subring of K with the property that its field of fractions is K*
2. *$\mathcal{O}_K$ is a free $\mathbb{Z}$-module of rank $[K:\mathbb{Q}]$.*

It follows from the second statement that the ring $\mathcal{O}_K/\mathfrak{a}$ is finite for every nonzero ideal $\mathfrak{a}$.

The cleanest way to prove the most basic results for algebraic number fields is to note that the ring of integers of an algebraic number field is an example of a *Dedekind domain*:

Definition 1.3.1.4. *A Dedekind domain A is an integral domain such that*

1. *A is Noetherian,*
2. *A is integrally closed,*
3. *Every nonzero prime ideal of A is maximal.*

The second condition is the statement that if $x \in K$ and x is integral over A, then $x \in A$. As promised, we have:

Proposition 1.3.1.5. *The ring of integers $\mathcal{O}_K$ of a number field K is a Dedekind domain.*

We will now outline the structure of the set of ideals of the ring of integers of a Dedekind domain A, obtaining by dint of Proposition 1.3.1.5 information on the ideal structure of rings of integers of number fields.

If $\mathfrak{a}$ and $\mathfrak{b}$ are ideals of a commutative ring A, then we say that $\mathfrak{a} \mid \mathfrak{b}$ if there exists an ideal $\mathfrak{c}$ such that $\mathfrak{a} = \mathfrak{b}\mathfrak{c}$. An ideal $\mathfrak{p}$ is said to be *prime* if $\mathfrak{p} \mid \mathfrak{a}\mathfrak{b}$ implies $\mathfrak{p} \mid \mathfrak{a}$ or $\mathfrak{p} \mid \mathfrak{b}$. For any nonzero ideal $\mathfrak{a}$ of a Dedekind domain A, the quotient ring $A/\mathfrak{a}$ is finite, and we have:

Proposition 1.3.1.6. *The* ideal norm *of a nonzero ideal $\mathfrak{a}$ of A is*

$$\mathfrak{N}(\mathfrak{a}) = |A/\mathfrak{a}| .$$

The ideal norm is multiplicative: if $\mathfrak{a}$ and $\mathfrak{b}$ are nonzero ideals of A, then

$$\mathfrak{N}(\mathfrak{a}\mathfrak{b}) = \mathfrak{N}(\mathfrak{a})\mathfrak{N}(\mathfrak{b}).$$

A famous observation concerning factorisation in number fields is that although elements of $\mathbb{Z}$ may be factorised uniquely into products of powers of primes, the analogous statement is not true for arbitrary number fields. For example, in $\mathbb{Q}(\sqrt{-5})$ we have

$$6 = 2 \cdot 3 = (1+\sqrt{-5})(1-\sqrt{-5}),$$

where $2, 3, 1+\sqrt{-5}$ and $1-\sqrt{-5}$ are all irreducible in the ring of integers $\mathbb{Z}\left[\frac{1+\sqrt{-5}}{2}\right]$ of $\mathbb{Q}(\sqrt{-5})$. This unhappy situation was brightened by Dedekind, who proved that unique factorisation holds for number fields at the level of *ideals*.

Proposition 1.3.1.7. *[ST87, Theorem 5.5] Every nonzero proper ideal of a Dedekind domain A has a unique factorisation as a product of prime ideals.*

We will now introduce *fractional ideals* for a Dedekind domain A. We write K for the field of fractions of A; this is perfectly consistent if A happens to be the ring of integers of a number field K by Proposition 1.3.1.5. From now on, we will use the phrase *integral ideal* to mean an ideal of A in the usual sense, and *ideal* to mean a fractional ideal (but *prime ideal* always means an integral prime ideal). It is not strictly necessary to introduce fractional ideals if one only wishes to discuss zeta functions of number fields at the most superficial level (as we do in the present section), but it is essential for our treatment of Gauss sums over number fields in Chapter 5.

Definition 1.3.1.8. *A fractional ideal $\mathfrak{a}$ of a Dedekind domain A is a nonzero finitely generated A-submodule of K.*

Given a nonzero fractional ideal $\mathfrak{a}$ of A, we define its inverse $\mathfrak{a}^{-1}$ by

$$\mathfrak{a}^{-1} = \{x \in K \mid x\mathfrak{a} \subseteq A\}.$$

One may prove that, with A as an identity element, the set of fractional ideals may be equipped with the structure of an abelian group. The structure of this group is provided by:

Proposition 1.3.1.9. *The set of fractional ideals is the free abelian group on the set of nonzero prime ideals of A.*

The expected results hold for fractional ideals: if $\mathfrak{a}$ is an ideal and $\mathfrak{b}$ and $\mathfrak{c}$ are nonzero ideals, then $\mathfrak{ab} = \mathfrak{ac}$ if and only if $\mathfrak{b} = \mathfrak{c}$; and $\mathfrak{b}$ and $\mathfrak{c}$ are coprime (in the sense of divisibility of ideals) if $(\mathfrak{b}, \mathfrak{c}) = \mathfrak{b} + \mathfrak{c} = 1$. Furthermore, the ideal norm may be extended to nonzero fractional ideals by multiplicativity.

The next three results, particularly Lemma 1.3.1.12, will come in handy in Chapter 5.

Proposition 1.3.1.10. *Let A be a Dedekind domain and $\mathfrak{a} \neq 0$ an integral ideal in A. Then every ideal in $A/\mathfrak{a}$ is principal.*

Proof. By Proposition 1.3.1.7, we may write $\mathfrak{a}$ as a product of powers of prime ideals. We localise to obtain

$$A/\mathfrak{a} \cong \prod A/\mathfrak{p}_i^{e_i} \cong \prod A_{\mathfrak{p}_i}/\mathfrak{p}_i^{e_i} A_{\mathfrak{p}_i}.$$

But every local ring $A_{\mathfrak{p}_i}$ is a principal ideal domain, so $A/\mathfrak{a}$ is a principal ideal ring. □

Lemma 1.3.1.11. *Let A be a Dedekind domain, $\mathfrak{a}$ a nonzero ideal of A. For each $x \in \mathfrak{a}$, there exists an ideal $\mathfrak{a}_x$ such that $(x) = \mathfrak{a}\mathfrak{a}_x$, and one may choose x so that $(\mathfrak{a}, \mathfrak{a}_x) = 1$.*

Proof. Since A is a Dedekind domain, for each x there exists an ideal $\mathfrak{a}_x$ such that $(x) = \mathfrak{a}\mathfrak{a}_x$ [Mil72, pp. 9]. It remains to show that x can be chosen so that $(\mathfrak{a}, \mathfrak{a}_x) = 1$. By Lemma 1.3.1.11, the ideal $\mathfrak{a}/\mathfrak{a}^2$ of $A/\mathfrak{a}^2$ is principal. Upon picking $x \in \mathfrak{a} \setminus \mathfrak{a}^2$, it follows that

$$\mathfrak{a} = ((x), \mathfrak{a}^2)) = (\mathfrak{a}\mathfrak{a}_x, \mathfrak{a}^2)).$$

Therefore $(\mathfrak{a}, \mathfrak{a}_x) = 1$. □

Lemma 1.3.1.12. *Suppose $\mathfrak{a}_1$ and $\mathfrak{a}_2$ are two coprime integral ideals of a Dedekind domain A. Then there exist auxiliary integral ideals $\mathfrak{c}_1$ and $\mathfrak{c}_2$ such that*

$$\mathfrak{a}_1\mathfrak{c}_1 = (\alpha_1) \quad \textit{and} \quad \mathfrak{a}_2\mathfrak{c}_2 = (\alpha_2), \quad \textit{where } \alpha_1, \alpha_2 \in A,$$

and $(\mathfrak{a}_1\mathfrak{a}_2, \mathfrak{c}_1\mathfrak{c}_2) = 1$.

Proof. We use the ideal norm, defined in Proposition 1.3.1.6. By Lemma 1.3.1.12, we may choose $\beta_1 \in \mathfrak{a}_1(\mathfrak{N}(\mathfrak{a}_2))$ and $\beta_2 \in \mathfrak{a}_2(\mathfrak{N}(\mathfrak{a}_1))$ such that the integral ideals

$$\mathfrak{c}_1 = (\beta_1)[\mathfrak{a}_1(\mathfrak{N}(\mathfrak{a}_2))]^{-1} \quad \text{and} \quad \mathfrak{c}_2 = (\beta_2)[\mathfrak{a}_2(\mathfrak{N}(\mathfrak{a}_1))]^{-1}$$

satisfy

$$(\mathfrak{a}_1(\mathfrak{N}(\mathfrak{a}_2)), \mathfrak{c}_1) = 1 \quad \text{and} \quad (\mathfrak{a}_2(\mathfrak{N}(\mathfrak{a}_1)), \mathfrak{c}_2) = 1.$$

and therefore, by multiplicativity of the ideal norm and the fact that $\mathfrak{a} \mid \mathfrak{N}(\mathfrak{a})$, we have

$$(\mathfrak{c}_1, \mathfrak{a}_1\mathfrak{a}_2) = 1 \quad \text{and} \quad (\mathfrak{c}_2, \mathfrak{a}_1\mathfrak{a}_2),$$

and the coprimality claim follows. □

Having introduced all the necessary supporting material which can be profitably presented in terms of Dedekind domains, and we now return to number fields. We will define the (element) norm, trace, different and discriminant. These objects are taken for granted over $\mathbb{Q}$; however, they are indispensable in the study of number fields. In fact, it is impossible to generalise Gauss sums and zeta functions without them.

Definition 1.3.1.13. *Let K be a number field of degree n, and let $\sigma_1, \ldots, \sigma_n$ be the n distinct field embeddings of K into $\mathbb{C}$. We define*

1. *The norm of $x \in K$:*
$$\mathrm{N}_K(x) = \prod_{i=1}^{n} \sigma_i(x).$$
2. *The trace of an element $x \in K$:*
$$\mathrm{Tr}_K(x) = \sum_{i=1}^{n} \sigma_i(x).$$
3. *The different ideal*
$$\mathfrak{d}_K^{-1} = \{x \in K \mid \mathrm{Tr}(x\mathcal{O}_K) \subseteq \mathcal{O}_K\}.$$
4. *The discriminant:*
$$d_K = \left(\det(\sigma_i(x_j))\right)^2,$$
where $\{x_1, \ldots x_n\}$ is a basis of K as a vector space over $\mathbb{Q}$ consisting of integers. *One must prove that such a basis exists and that the discriminant is well-defined.*

We sometimes omit the subscript referring to the field, if there is no chance of confusion. The reader should keep in mind, however, that for a fixed element $x \in K$, the trace, norm, different and discriminant may change if K is replaced by an extension field.

Proposition 1.3.1.14. *We list some well-known properties of the trace, norm and different [Neu99, Chapter 1, Sections 2 and 3; ST87, Chapter 2, Section 5, Theorem 5.8].*

1. *The norm and trace are rational integers, and for $x, y \in k$, $a, b \in \mathbb{Q}$, we have*
$$\mathrm{N}(xy) = \mathrm{N}(x)\mathrm{N}(y) \quad \text{and} \quad \mathrm{Tr}(ax + by) = a\mathrm{Tr}(x) + b\mathrm{Tr}(y).$$
2. *If (x) is a principal ideal of $\mathcal{O}_K$, then $\mathfrak{N}((x)) = |\mathrm{N}(x)|$.*
3. *The different is an integral ideal, and $\mathfrak{N}(\mathfrak{d}) = |d_K|$.*

The last few items we need to state are concerned with bases for ideals. Any nonzero ideal $\mathfrak{a}$ has a $\mathbb{Z}$-*basis* $\{\alpha_1, \ldots, \alpha_n\}$: that is, $\mathfrak{a}$ is equal to the $\mathbb{Z}$-span of $\{\alpha_1, \ldots, \alpha_n\}$. We define the *discriminant* of such a basis (not to be confused with the field discriminant) to be

$$\Delta[\alpha_1, \ldots \alpha_n] = \left(\det(\sigma_i(\alpha_j))\right)^2.$$

Proposition 1.3.1.15. *[Hec81, Theorem 101, Theorem 102] Let $\mathfrak{a}$ be a nonzero ideal of $\mathcal{O}_K$ with a $\mathbb{Z}$-basis $\{\alpha_1, \ldots, \alpha_n\}$ where $n = [K : \mathbb{Q}]$. Then*

1. *We have a relationship between the ideal norm and the discriminant of $\mathfrak{a}$:*
$$\sqrt{|\Delta[\alpha_1, \ldots \alpha_n]|} = \mathfrak{N}(\mathfrak{a})\sqrt{|d_K|}.$$
2. *Define n elements of K by the implicit relations*
$$\sum_{i=1}^{n} \sigma_i(\beta_j \alpha_k) = \begin{cases} 1 & j = k \\ 0 & j \neq k. \end{cases}$$
Then $\{\beta_1, \ldots, \beta_n\}$ form a $\mathbb{Z}$-basis for $\mathfrak{a}^{-1}\mathfrak{d}^{-1}$.

1.3.2 The Hilbert symbol and reciprocity laws

Let p be an odd prime and n an integer. Define the *Legendre symbol* by

$$\left(\frac{n}{p}\right) = \begin{cases} 0 & p \mid n, \\ 1 & (n,p) = 1 \text{ and } n = x^2 \bmod p \text{ for some integer } x, \\ -1 & (n,p) = 1 \text{ and } n \neq x^2 \bmod p \text{ for any integer } x. \end{cases}$$

Then we have the following theorem, stated by Euler, partially proved by Legendre, and first proved completely by Gauss:

Theorem 1.3.2.1 (Quadratic reciprocity [Gau01a; Gau01b; Gau08; Gau11])**.** *Let p and q be distinct odd primes. Then*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4}.$$

We will explore the relationship between the Legendre symbol and certain theta functions in Chapter 4.

One can extend quadratic reciprocity to an arbitrary number field, by defining, for an odd prime ideal $\mathfrak{p}$ and $n \in \mathcal{O}_K$,

$$\left(\frac{n}{\mathfrak{p}}\right) = \begin{cases} 0 & \mathfrak{p} \mid n, \\ 1 & (n,\mathfrak{p}) = 1 \text{ and } n = x^2 \bmod \mathfrak{p} \text{ for some integer } x, \\ -1 & (n,\mathfrak{p}) = 1 \text{ and } n \neq x^2 \bmod \mathfrak{p} \text{ for any integer } x. \end{cases}$$

Then the law of quadratic reciprocity takes the following form:

Theorem 1.3.2.2 (Quadratic reciprocity over number fields)**.** *Let p and q be distinct odd primes of $\mathcal{O}_K$, at least one of which is congruent to a square modulo 4. Let $\sigma_1, \ldots, \sigma_{r_1}$ be the* real *embeddings of K. Then*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\sum_{j=1}^{r_1} (\operatorname{sgn} p^{(j)} - 1)(\operatorname{sgn} q^{(j)} - 1)/4}.$$

In Chapter 5 we will see that many of the relationships between theta functions and Legendre symbols generalise to the number field case.

The desire to generalise Theorem 1.3.2.2 to *higher powers* sparked the study of class field theory, and was fulfilled by the discovery of *Hilbert reciprocity*. We will not attempt to treat Hilbert reciprocity with any degree of completeness, as it suffices for the purposes of Chapter 6 for the reader to know that reciprocity laws for higher powers exist; however, we will be able to at least explain some of the notation. The reader may consult [Neu99, Chapter V, Section 3, Chapter VI, Section 8] for the full story.

Definition 1.3.2.3. *Let K be a field. A* valuation *on K is a map $|\cdot| : K \to \mathbb{R}$ such that:*

1. *For all $x \in K$, $|x| \geq 0$, with equality if and only if $x = 0$,*
2. *For all $x, y \in K$, $|xy| = |x||y|$,*
3. *For all $x, y \in K$, $|x + y| \leq |x| + |y|$.*

A valuation induces a topology on K via the metric $d(x,y) = |x - y|$, and two valuations are equivalent *if they induce the same topology on K. A valuation is* trivial *if $|x| = 1$ for all $x \in K$, and is* non-archimedean *if the ultrametric inequality*

$$|x + y| \leq \max\{|x|, |y|\}$$

is satisfied for all $x, y \in K$.

Definition 1.3.2.4. *Let K be an algebraic number field. A* place $\mathfrak{p}$ *of K is an equivalence class of valuations on K. Each place of K belongs to exactly one of the following collections:*

1. *Primes $\mathfrak{p}$ of $\mathcal{O}_K$ ("finite primes"), corresponding to the valuations $|x|_{\mathfrak{p}} = \mathfrak{N}(\mathfrak{p})^{-\operatorname{ord}_{\mathfrak{p}}(x)}$, where $\operatorname{ord}_{\mathfrak{p}}(x)$ is the power of $\mathfrak{p}$ appearing in the factorisation of the ideal (x) into powers of prime ideals (c.f. 1.3.1.9);*
2. *Real embeddings $\sigma : K \to \mathbb{R}$, corresponding to the valuations $|x|_\sigma = |\sigma(x)|$;*

3. Conjugate pairs of complex embeddings $\sigma : K \to \mathbb{C}$, corresponding to the valuations $|x|_\sigma = |\sigma(x)|^2$.

Definition 1.3.2.5. *A* local field *is a field K together with a nontrivial valuation ν, such that the induced topology on K is locally compact. A local field is* non-archimedean *if ν is non-archimedean. The subring*

$$\mathcal{O}_K = \{x \in K \mid |x| \leq 1\}$$

is a local ring, and the maximal ideal associated to $\mathcal{O}_K$ is

$$\mathfrak{p}_K = \{x \in \mathcal{O}_K \mid |x| < 1\}.$$

Most importantly, if K is a number field and $\mathfrak{p}$ a place of K, then the completion $K_\mathfrak{p}$ of K with respect to the equivalence class of valuations corresponding to $\mathfrak{p}$ is a local field. The key to Hilbert reciprocity is the following result:

Theorem 1.3.2.6. *For every finite Galois extension L/K of local fields there is a canonical isomorphism, known as the* local Artin reciprocity law*:*

$$\mathrm{Gal}(L/K)^{\mathrm{ab}} \xrightarrow{\sim} K^*/\mathrm{N}_{L/K}L^*,$$

where G^{ab} denotes the abelianisation of G (the quotient of G by its commutator subgroup) and $\mathrm{N}_{L/K}$ is the map defined by

$$x \mapsto \prod_{\sigma \in \mathrm{Gal}(L/K)} \sigma(x).$$

Inverting the isomorphism gives rise to a surjection

$$(\quad, L/K) : K^* \to \mathrm{Gal}(L/K)^{ab}.$$

Now we may define the local Hilbert symbols.

Definition 1.3.2.7. *Let K be a local field containing the group μ_n of nth roots of unity, where n is a natural number relatively prime to the characteristic of K. Let $\mathfrak{p}$ be the maximal ideal associated to the local ring $\mathcal{O}_K$. Let $a, b \in K^*$. Then the* local Hilbert symbol $\left(\frac{a,b}{\mathfrak{p}}\right)$ *is the element of μ_n defined by*

$$\left(a, K\left(b^{1/n}\right)/K\right) b^{1/n} = \left(\frac{a,b}{\mathfrak{p}}\right) b^{1/n}.$$

It remains to define the global Hilbert symbols and state Hilbert reciprocity. Let K be a number field containing the nth roots of unity, and let $\mathfrak{p}$ be a prime with $(n) \nmid \mathfrak{p}$. Let $\alpha \in \mathcal{O} \setminus \mathfrak{p}$, so that

$$\alpha^{\mathfrak{N}(\mathfrak{p})-1} = 1 \mod \mathfrak{p}$$

by Fermat's little theorem. Then one characterises the *nth power residue symbol* $\left(\frac{\alpha}{\mathfrak{p}}\right)$ by

$$\left(\frac{\alpha}{\mathfrak{p}}\right)_n = \alpha^{(\mathfrak{N}(\mathfrak{p})-1)/n} \bmod \mathfrak{p}.$$

One easily checks that when $n = 2$, the nth power residue symbol reduces to the Legendre symbol, and that $\left(\frac{\alpha}{\mathfrak{p}}\right)_n = 1$ if and only if α is an nth power modulo $\mathfrak{p}$. We extend the nth power residue symbol to ideals $\mathfrak{b} = \prod_{\mathfrak{p} \nmid (n)} \mathfrak{p}^{\mu_p}$ relatively prime to n by

$$\left(\frac{\alpha}{\mathfrak{b}}\right) = \prod_{\mathfrak{p} \nmid (n)} \left(\frac{\alpha}{\mathfrak{p}}\right)^{\mu_p},$$

and when $\mathfrak{b} = (\beta)$ is principal we write $\left(\frac{\alpha}{\beta}\right)$ for $\left(\frac{\alpha}{\mathfrak{b}}\right)$. Finally, we state the analogue of Theorems 1.3.2.1 and 1.3.2.2:

Theorem 1.3.2.8 (Hilbert reciprocity, [Neu99, Chapter VI, Section 8])**.** *Suppose that K is a field containing the nth roots of unity, and let α and β be nonzero elements of K coprime to each other and to n. Then*

$$\left(\frac{\alpha}{\beta}\right)_n \left(\frac{\beta}{\alpha}\right)_n^{-1} = \prod_{\mathfrak{p} \mid n\infty} \left(\frac{\alpha, \beta}{\mathfrak{p}}\right),$$

where $\left(\frac{\alpha,\beta}{\mathfrak{p}}\right)$ is the local Hilbert symbol associated to the local field $K_\mathfrak{p}$, and the product runs over the infinite places of K as well as the finite primes dividing n.

In the special cases of cubic and quartic reciprocity, we have the following elegant statements (for which one may consult Cox [Cox89, Section 4]):

Theorem 1.3.2.9. *Let $K = \mathbb{Q}\left(-\frac{1}{2} + \frac{\sqrt{3}}{2}\right)$, so that $\mathbb{Z}\left[-\frac{1}{2} + \frac{\sqrt{3}}{2}\right]$ is its ring of integers. A prime p in $\mathbb{Z}\left[-\frac{1}{2} + \frac{\sqrt{3}}{2}\right]$ is said to be* primary *if $p = \pm 1 \bmod 3$. Then if p and q are primary primes of unequal norm,*

$$\left(\frac{p}{q}\right)_3 = \left(\frac{q}{p}\right)_3.$$

Theorem 1.3.2.10. *Let $K = \mathbb{Q}(i)$ and $\mathbb{Z}[i]$ be its ring of integers. A prime q of $\mathbb{Z}[i]$ is* primary *if $p = 1 \bmod (2+2i)$. If p and q are distinct primary primes, then*

$$\left(\frac{p}{q}\right)_4 = \left(\frac{q}{p}\right)_4 (-1)^{(\mathrm{N}(p)-1)(\mathrm{N}(q)-1)/16}.$$

1.3.3 Dirichlet characters and Gauss sums

In this subsection we will introduce *Dirichlet characters* and *Gauss sums*, objects of central importance in Chapters 4, 5 and 6. The term "Gauss sum" is most commonly used, for certain multiplicative functions $g(n)$ of period m, to denote an expression of the form

$$\sum_{n \bmod m} g(n) \exp\left(\frac{\pi i n}{m}\right),$$

or for variations of the above with multiplicative factors present in the numerator of the fraction. The function $g(n)$ is usually a Dirichlet character of modulus m, or, in the context of generalisations to number fields, the reciprocity symbol to a fixed denominator.

We begin by defining Dirichlet characters over $\mathbb{Q}$.

Definition 1.3.3.1. *A* Dirichlet character *χ of modulus m is a ring homomorphism from $(\mathbb{Z}/m\mathbb{Z})^\times$ to $\mathbb{S}^1$, the multiplicative group of complex numbers of unit norm.*

A Dirichlet character χ is called primitive *if it does not arise as a composite*

$$(\mathbb{Z}/m\mathbb{Z})^\times \longrightarrow (\mathbb{Z}/m'\mathbb{Z})^\times \xrightarrow{\chi'} \mathbb{S}^1$$

for some Dirichlet character χ' modulo m', where m' is a proper divisor of m. The principal *Dirichlet character modulo m is the map which sends all elements of $(\mathbb{Z}/\mathbb{Z})^\times$ to 1.*

Furthermore, χ is called even *or* odd *depending on whether $\chi(-1)$ is $+1$ or -1.*

Definition 1.3.3.2 (The Gauss sum of a Dirichlet character)**.** *Let χ be a Dirichlet character modulo m. Then the* Gauss sum *$G(\chi)$ of χ is*

$$G(\chi) = \sum_{n \bmod m} \chi(n) \exp\left(\frac{2\pi i n}{m}\right).$$

Lemma 1.3.3.3. *Let χ be a primitive character modulo m. Then*

$$\sum_{k=0}^{m-1} \chi(k) e^{2\pi i k n/m} = \overline{\chi(n)} G(\chi). \tag{1.10}$$

Proof. If $(n, m) = 1$, then we can find an inverse for n modulo m and the claim is clear. If $(n, m) = d > 0$, we can show that both sides are zero. This is because χ is primitive; we may choose $a = 1 \bmod m/d$ so that $a \neq 1 \bmod m$ and $\chi(a) \neq 1$. It follows that

$$\chi(a) \sum_{k=0}^{m-1} \chi(k) e^{2\pi i k n/m} = \sum_{k=0}^{m-1} \chi(k) e^{2\pi i k n/m}.$$ □

The study of Dirichlet characters and Gauss sums often comes down to Fourier analysis on $\mathbb{Z}/m\mathbb{Z}$ and $(\mathbb{Z}/m\mathbb{Z})^\times$, and in such cases we may apply the results of Subsection 1.2.1.

Proposition 1.3.3.4. *Let $d \geq 1$ and let p be a prime congruent to 1 modulo d. Then*

$$\sum_{n=0}^{p-1} \exp\left(\frac{2\pi i n^d}{p}\right) = \sum_{\substack{|\chi|=p \\ \chi^d=1}} \sum_{n=0}^{p-1} \chi(n) \exp\left(\frac{2\pi i n}{p}\right).$$

Proof. We rewrite the left hand side:

$$\sum_{n=0}^{p-1} \exp\left(\frac{2\pi i n^d}{p}\right) = \sum_{\substack{n=0 \\ n=y^d \bmod p}}^{p-1} \exp\left(\frac{2\pi i n}{p}\right), \tag{1.11}$$

and then observe that

$$\sum_{\substack{|\chi|=p \\ \chi^d=1}} \chi(n) = \sum_{\chi \in \widehat{(\mathbb{Z}/p\mathbb{Z})}[d]} \chi(n),$$

where we use $G[d]$ to mean the subgroup of elements of order d for an abelian group G. But the map $\phi : \hat{G}[d] \to \widehat{G/G^d}$ defined by

$$\phi(\chi)(g + G^m) = \chi(g)$$

is an isomorphism, so by 1.3,

$$\sum_{\chi \in \hat{G}[d]} \chi(g) = \sum_{\chi \in \widehat{G/G^m}} \chi(g) = \begin{cases} |G/G^d| & g = 0, \\ 0 & g \neq 0, \end{cases}$$

$$= \begin{cases} |G[d]| & g = h^d \text{ for some } h \in G, \\ 0 & g \neq 0, \end{cases}$$

where we used the fact that $\hat{G} \cong G$, so $\hat{G}[d] \cong G[d]$. It follows that

$$\sum_{\substack{|\chi|=p \\ \chi^d=1}} \chi(n) = \#\{y \in \mathbb{Z}/p\mathbb{Z} \mid y^d = n\},$$

so

$$\sum_{\substack{n=0 \\ n=y^d \bmod p}}^{p-1} \exp\left(\frac{2\pi i n}{p}\right) = \sum_{n=0}^{p-1} \sum_{\substack{|\chi|=p \\ \chi^d=1}} \chi(n) \exp\left(\frac{2\pi i n}{p}\right). \tag{1.12}$$

Combining 1.11 and 1.12 proves the proposition. □

We also require a notion of Gauss sums of Dirichlet characters over number fields. In this setting, we have a character

$$\chi : (\mathcal{O}_K/\mathfrak{m})^\times \longrightarrow \mathbb{S}^1,$$

where $\mathfrak{m}$ is an integral ideal such that there exists some $y \in K$ with

$$\mathfrak{d}(y) = \frac{\mathfrak{b}}{\mathfrak{m}}, \quad (\mathfrak{b}, \mathfrak{m}) = 1.$$

Any such Dirichlet character may be extended to all of $\mathcal{O}_K$ by declaring it to be zero on elements x with $(\mathfrak{m}, (x)) \neq 1$.

Definition 1.3.3.5. *With χ a Dirichlet character on $(\mathcal{O}_K/\mathfrak{m})^\times$, $\mathfrak{d}y = \mathfrak{b}\mathfrak{m}^{-1}$ and $(\mathfrak{b}, \mathfrak{m}) = 1$ as above, we define the Gauss sum of χ by:*

$$G_y(\chi, \nu) = \sum_{\mu \in \mathcal{O}_K} \chi(\mu) e^{2\pi i \mathrm{Tr}(\mu\nu y)}.$$

In this setting, y plays the same role as did $1/n$ for Gauss sums of Dirichlet characters over $\mathbb{Q}$.

We have an analogue of Lemma 1.3.3.3:

Lemma 1.3.3.6. *Let χ be a primitive character of $(\mathcal{O}_K/\mathfrak{m})^\times$ and y be as above. Let $\nu \in \mathcal{O}_K$. Then*

$$G_y(\chi, \nu) = \overline{\chi(\nu)} G_y(\chi),$$

where we define $G_y(\chi) = G_y(\chi, 1)$.

Proof. The claim is clear if $((\nu), \mathfrak{m}) = 1$. If instead $((\nu), \mathfrak{m}) = \mathfrak{n} \neq 1$, then since χ is primitive we may find an algebraic integer x such that $\chi(x) \neq 1$ and $x = 1 \bmod \mathfrak{m}\mathfrak{n}^{-1}$. Then $\nu x = \nu \bmod \mathfrak{m}$, and $\nu y(x-1) \in \mathfrak{d}^{-1}$. Then

$$\overline{\chi(x)} G_y(\chi, \nu) = G_y(\chi, \nu),$$

and so $G_y(\chi, \nu) = 0$. □

1.3.4 Dedekind zeta functions and Dirichlet L-functions

We now state the basic facts about zeta functions which we will invoke in Chapter 6. We deal with the Riemann zeta function and its twists by Dirichlet characters, called *L-functions*, first. We are content to state the results we will need as the proofs are well-known.

The Riemann zeta function is one of the central objects which concern us. Its fortunes are closely tied with its companion, Jacobi's theta function. First, we state the fundamental properties of the zeta function.

Proposition 1.3.4.1. *For* $\mathrm{Re}(s) > 1$, *define*

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s}.$$

Then, for $\mathrm{Re}(s)$ *sufficiently large, the zeta function admits an* Euler product*:*

$$\zeta(s) = \prod_{p \text{ prime}} \left(1 - p^{-s}\right)^{-1},$$

and admits a meromorphic extension to the entire complex plane, where it has only one pole, at $s = 1$, and satisfies the functional equation

$$\pi^{-s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s) = \pi^{-(1-s)/2}\Gamma\left(\frac{1-s}{2}\right)\zeta(1-s). \tag{1.13}$$

The pole at $s = 1$ is simple and the residue there is 1. One can see that the zeta function has simple zeros at $s = -2, -4, -6, \ldots$, due to the location of the poles of the gamma functions in the functional equation. The Riemann hypothesis *states that the only other zeros of $\zeta(s)$ have real part $1/2$.*

For us, the most interesting result is the functional equation 1.13. One has

$$\pi^{-s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s) = \frac{1}{2}\int_0^\infty (\theta(iy) - 1) y^{s-1} dy, \tag{1.14}$$

where

$$\theta(z) = \sum_{n=-\infty}^{\infty} e^{2\pi i n^2 z};$$

thus the Riemann zeta function is represented as a Mellin transform of the Jacobi theta function. Using the functional equation 2.2.1.1 for Jacobi's theta function, one may use 1.14 to prove 1.13; conversely, one may use the functional equation 1.13 to prove 2.2.1.1. This correspondence is well-known (and goes all the way back to Riemann), and in Subsection 2.3.3 we use it to prove that a certain function is a *quasimodular form*.

Now we come to Dirichlet L-functions.

Proposition 1.3.4.2. *Let χ be a Dirichlet character modulo m. For* $\mathrm{Re}(s) > 1$, *define*

$$L(\chi, s) = \sum_{n=1}^{\infty} \chi(n) n^{-s}.$$

Then, for Re(s) *sufficiently large, the zeta function admits an* Euler product*:*

$$L(s,\chi) = \prod_{p \text{ prime}} \left(1 - \chi(p)p^{-s}\right)^{-1}.$$

Now suppose that χ *is primitive. Then* χ *admits a* holomorphic *extension to the entire complex plane, and satisfies a functional equation [Neu99, Chapter VII, Theorem 2.8]. If* χ *is the principal character, then* $L(s,\chi)$ *admits a meromorphic extension to* $\mathbb{C}$*, with a simple pole at* $s = 1$ *with residue* $\phi(m)/m$*, where* ϕ *is Euler's totient function. Trivial zeros occur at* $s = 0, -2, -4, -6, \ldots$ *if* χ *is even and non-principal; at* $s = -2, -4, -6, \ldots$ *if* $\chi = 1$*, and at* $s = -1, -3, -5, -7, \ldots$ *if* χ *is odd. The functional equation implies that they are all simple. The* Generalised Riemann Hypothesis *[Win38, Chapter 13] states that all other zeros of* $L(s,\chi)$ *have real part* $1/2$.

As for the Riemann zeta function, the Dirichlet L-functions may be represented as Mellin transforms of *twisted* theta functions.

Lastly, we deal with zeta functions of number fields, although, excepting the Riemann zeta function, we mention them only in passing at various points.

Proposition 1.3.4.3. *Let* K *be a number field with real embeddings* $\sigma_1, \ldots, \sigma_{r_1}$*, complex embeddings* $\sigma_{r_1+1}, \ldots, \sigma_{r_1+2r_2}$ *and set* $n = [K : \mathbb{Q}]$*. Then the* Dedekind zeta function *of* K *is defined for* Re$(s) > 1$ *by*

$$\zeta_K(s) = \sum_{\mathfrak{a}} \mathfrak{N}(\mathfrak{a})^{-s},$$

where $\mathfrak{a}$ *runs over the integral ideals of* $\mathcal{O}_K$*. For* Re(s) *sufficiently large we have*

$$\zeta_K(s) = \prod_{\mathfrak{p}} \left(1 - \mathfrak{n}(\mathfrak{p})\right)^{-s},$$

where $\mathfrak{p}$ *runs over all the prime ideals of* $\mathcal{O}_K$*. The Dedekind zeta function has an analytic continuation of* $\mathbb{C} \setminus \{1\}$*, with a simple pole at* $s = 1$*, and satisfies the functional equation*

$$\zeta_K(1-s) = |d_K|^{s-1/2} \left(\cos\left(\frac{\pi s}{2}\right)\right)^{r_1+r_2} \left(\sin\left(\frac{\pi s}{2}\right)\right)^{r_2} \left(2(2\pi)^{-s}\Gamma(s)\right)^n \zeta_K(s),$$

where d_K *is the discriminant. The* analytic class number formula *[Neu99, Chapter 7, Corollary 5.11] states that the residue of* $\zeta_K(s)$ *at* $s = 1$ *is*

$$\frac{2^{r_1}(2\pi)^{r_2} h_K \mathrm{Reg}_K}{\omega_K |d_K|^{1/2}},$$

where h_K *is the class number of* K*,* Reg_K *is the regulator and* ω_K *is the number of distinct* n*th roots of unity contained in* K*.*

As usual, the functional equation is equivalent to a functional equation for a certain theta function.

1.3.5 The Mordell–Tornheim zeta function

For this subsection we retreat to more concrete territory: the study of the special values of the Riemann zeta function. The results discussed in this subsection will come in useful in the proof of Proposition 2.2.2.2, in which we show that the transformation law for Jacobi's theta function (Theorem 2.2.1.1) is equivalent to a finite collection of identities between special values of zeta functions.

It is well-known [Neu99, Section VII, Theorem 1.10] that the values of the Riemann zeta function at positive even integers are rational multiples of powers of π: one has

$$\zeta(2n) = (-1)^{n-1} B_{2n} \frac{(2\pi)^{2n}}{2(2n)!} \tag{1.15}$$

for all integers $k \geq 1$, where B_{2n} denotes the $2n$th Bernoulli number, defined implicitly by

$$\frac{ze^z}{e^z - 1} = \sum_{n=0}^{\infty} B_n \frac{z^n}{n!}.$$

One may prove 1.15, for small fixed n, using the following integral representation of the zeta function for positive integral n:

$$\zeta(n) = \sum_{k=0}^{\infty} \int_0^1 \cdots \int_0^1 \left(\prod_{i=1}^{n} x_i \right)^k dx_1 \dots dx_n = \int_0^1 \cdots \int_0^1 \frac{1}{1 - \prod_{i=1}^{n} x_i} dx_1 \dots dx_n, \tag{1.16}$$

or alternatively, one may modify Euler's original argument, based on the product representation of the sine. Another route to proving 1.15, or at least 1.17 and 1.18 below, is to use the reduction formula (for k a positive integer)

$$2\left(1 - 2^{-2k-2}\right)\zeta(2k+2) = 2\left(1 - 2^{1-2k}\right)\zeta(2)\zeta(2k) - 4\sum_{n=1}^{k}(2n-1)2^{-2n}\left(1 - 2^{2n-2-2k}\right)\zeta(2n)\zeta(2k+2-2n),$$

proven in 1947 by Estermann [Est47] using elementary methods together with the evaluation of $\zeta(2)$ (see also the paper of Apostol [Apo73] for a proof of a simpler version and a history of related results).

In any case, one may prove, without recourse to Fourier analysis, the formulae

$$\zeta(2) = \frac{\pi^2}{6}$$

$$\zeta(4) = \frac{\pi^4}{90} \tag{1.17}$$

$$\zeta(6) = \frac{\pi^6}{945}. \tag{1.18}$$

The values of the Riemann zeta function at odd integers are much more mysterious, and only $\zeta(3)$ is known to be irrational [Apé79].

A class of zeta-like functions with interesting special values is provided by the *multiple zeta functions*:

Definition 1.3.5.1. *A multiple zeta function (or* Euler sum*) is a function of several complex variables defined by:*

$$\zeta(s_1, \dots, s_k) = \sum_{n_1 > \dots > n_k \geq 1} \frac{1}{n_1^{s_1} \dots n_k^{s_k}}.$$

A multiple zeta value *of length k and weight n is any of the numbers $\zeta(s_1, \dots, s_k)$ such that $s_1, \dots, s_k$ are positive integers which sum to n.*

The multiple zeta values satisfy a bewildering number of identities. In particular, a multiple zeta value can sometimes be expressed in terms of special values of the Riemann zeta function. In 1771, Euler [Eul38, Section 15] used partial fractions and combinatorics to obtain[3]

$$\zeta(4,2) = -\frac{16}{3}\zeta(4)\zeta(2) + \zeta(3)^2 + 8\zeta(6) \tag{1.19}$$

$$\zeta(5,1) = 3\zeta(4)\zeta(2) - \frac{1}{2}\zeta(3)^2 - \frac{9}{2}\zeta(6), \tag{1.20}$$

together with a slew of related identities.

Now we come to the main attraction of this section: the *Mordell–Tornheim zeta function.*

Definition 1.3.5.2. *The Mordell–Tornheim zeta function is a function of several complex variables defined by:*

$$\zeta_{MT}(s_1, \dots, s_k; s) = \sum_{n_1, \dots, n_k \geq 1} \frac{1}{n_1^{s_1} \dots n_k^{s_k}(n_1 + \dots + n_k)^s}.$$

A Mordell–Tornheim zeta value *of depth k and weight n is any of the numbers $\zeta_{MT}(s_1, \dots, s_k; s)$ such that $s_1, \dots, s_k, s$ are positive integers which sum to n.*

[3] Note that Euler's notation is related to the modern notation by

$$\int \frac{1}{z^m}\left(\frac{1}{y^m}\right) = \zeta(m,n) + \zeta(m+n).$$

Tornheim [Tor50] was effectively the first to study what are now called Mordell–Tornheim zeta values, but it was Mordell [Mor58] who, in 1958, gave the remarkable evaluation

$$\zeta_{MT}(2,2;2) = \frac{\pi^6}{2835}. \tag{1.21}$$

His proof uses Fourier analysis. Indeed, he defined a now-deprecated variant of the Bernoulli polynomials $B_r(x)$:

$$\frac{te^{tx}}{e^t-1} = \sum_{r=0}^{\infty} B_r(x)t^r,$$

and cubed both sides of the Fourier expansion of the $B_r(x)$:

$$(2\pi)^r B_r(x) = (-1)^{r/2-1}\sum_{l=1}^{\infty}\frac{2\cos(2\pi l x)}{l^r}.$$

The identity 1.21 emerges upon integrating the resulting mess over $[0,1]$.

In 2010, Bradley and Zhou proved the following remarkable generalisation of Mordell's evaluation:

Theorem 1.3.5.3 (Bradley–Zhou, [BZ10])**.** *Every Mordell–Tornheim zeta value of depth r and weight w can be expressed as a rational linear combination of multiple zeta values of depth r and weight w.*

The remarkable point about their proof is that it is *purely combinatorial* and *constructive*. We will illustrate Theorem 1.3.5.3 by proving Mordell's evaluation 1.21; thus demonstrating that *no Fourier analysis is necessary* for Mordell's result.

Indeed, their result, with indices determined by the special value $(2,2;2)$ replaced with 2s where possible, gives us

$$\zeta_{MT}(2,2,2) = \sum_{j=1}^{2}\left(\prod_{k=1}^{2}\sum_{a_k=0}^{1}\right) M_j T_1\left(\mathop{\mathbf{Cat}}_{\substack{l=1\\ l\neq j}}^{2}\{2-a_l\}, 4+A_j\right),$$

where

$$M_j = (1+A_j)!\prod_{\substack{m=1\\ m\neq j}}^{2}\frac{1}{a_m!}, \quad A_j = \sum_{\substack{m=1\\ m\neq j}}^{2} a_m \quad \text{and} \quad \mathop{\mathbf{Cat}}_{\substack{l=1\\ l\neq j}}^{2}\{2-a_l\} = \begin{cases} 2-a_2 & j=1,\\ 2-a_1 & j=2.\end{cases}$$

We also have $T_1(s,t) = \zeta(t,s)$. Upon simplifying, Theorem 1.3.5.3 yields

$$\zeta(2,2;2) = 2\zeta(4,2) + 4\zeta(5,1).$$

Then, using Euler's formulae 1.19 and 1.20 for the double zeta values, we find that

$$\zeta(2,2;2) = \frac{4}{3}\zeta(4)\zeta(2) - 2\zeta(6),$$

which, using 1.16, simplifies as required.

Chapter 2

Modular forms

To paraphrase the precise definition, a modular form is a function on the upper half plane which transforms in a simple manner under the action of an arithmetic subgroup of $SL(2,\mathbb{R})$. Such a definition does nothing to express the fundamental role played by modular forms in the study of arithmetic. Therefore, instead of beginning with the exact definition, we first illustrate some important examples.

1. Let $n \geq 1$ be an integer, and define $r_4(n)$ to be the number of ways in which n may be expressed as a sum of four squares. Then the function
$$\theta^4(z) = \sum_{n=0}^{\infty} r_4(n) e^{2\pi i n^2 z}$$
is holomorphic for z in the upper half plane $\mathscr{H}$, and it is a modular form, since for all $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL(2,\mathbb{Z})$ with $c = 0 \bmod 4$,
$$\theta^4\left(\frac{az+b}{cz+d}\right) = (cz+d)^2 \theta^4(z).$$

2. Define a function Δ by the following identity:
$$\Delta(z) = \sum_{n=1}^{\infty} \tau(n) e^{2\pi i n z} = e^{2\pi i z} \prod_{n=1}^{\infty} \left(1 - e^{2\pi i n z}\right)^{24}.$$
This function, nowadays known as the *modular discriminant*, was first studied by Ramanujan in 1916 [Ram16]. It is a modular form, since for all $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL(2,\mathbb{Z})$, it satisfies
$$\Delta\left(\frac{az+b}{cz+d}\right) = (cz+d)^{12} \Delta(z).$$

3. Let $\Lambda = \mathbb{Z}\omega_1 \oplus \mathbb{Z}\omega_2$ be a lattice: that is, the set of all integral linear combinations of two fixed complex numbers ω_1 and ω_2 such that the line in the complex plane through ω_1 and ω_2 doesn't pass through the origin. Any lattice may be rotated and scaled so that $\omega_1 = 1$ and ω_2 has positive imaginary part. Furthermore, two lattices $\mathbb{Z} \oplus \tau\mathbb{Z}$ and $\mathbb{Z} \oplus \tau'\mathbb{Z}$ are equal if and only if τ and τ' are related by
$$\frac{a\tau+b}{c\tau+d} = \tau' \text{ for some } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{Z}).$$
Consequently, for $k > 2$, the function
$$G_k(\tau) = \sum_{(m,n)\in\mathbb{Z}^2\setminus\{0,0\}} \frac{1}{(m+n\tau)^k}$$
converges absolutely and uniformly on compact subsets of $\mathscr{H}$ and satisfies
$$G\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^k G(\tau) \text{ for all } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{Z}).$$

Each of the three examples above demarcates a particularly interesting or special class of modular forms. The first is a *theta function*. It is possible to use the theory of modular forms, together with techniques of analysis and geometry, to prove that the numbers $r_4(n)$ are all positive. In fact, it is a general theme that the Fourier coefficients of modular forms often carry interesting number-theoretic information.

The second example is a *cusp form*. Cusp forms are among the most elusive of modular forms, due to the difficulty in studying their Fourier coefficients. A famous result, due to Ramanujan [Har40, Chapter X, pp. 161–185], is that the numbers $\tau(n)$, defined above as the Fourier coefficients of $\Delta(z)$, satisfy the congruence

$$\tau(n) = \sigma_{11}(n) \bmod 691,$$

where $\sigma_{11}(n)$ the sum of the 11th powers of the positive divisors of n. It is still unknown whether $\tau(n) \neq 0$ for all n, although this is conjectured to be true [Leh47; Mur82].

The final example is that of an *Eisenstein series*. These are the modular forms which built by symmetrising simple functions over the action of a discrete arithmetic group. The result (if it converges) is a modular form, but work must be done to unveil the Fourier coefficients. The Fourier coefficients of Eisenstein series are often of arithmetic interest. We shall see in Proposition 2.1.2.4 that the Fourier coefficients of the G_k are essentially sums-of-divisors functions. In Chapter 6 we will investigate a more complicated Eisenstein series, the Fourier coefficients of which turn out to be related to special values of L-functions.

In this chapter, we will investigate modular forms in sufficient detail to be able to explain the usual modern proof of Jacobi's four square theorem:

Theorem (Jacobi, 1829 [Jac29])**.** *For $n \geq 1$, let $r_4(n)$ count the number of ways that n may be represented as a sum of four squares. Then*

$$r_4(n) = 8 \sum_{\substack{d \mid n \\ 4 \nmid d, d>0}} d.$$

Clearly, $r_4(n)$ is always positive since 1 divides n and $4 \nmid 1$.

An earlier version, implicit in the writings of Diophantus, conjectured by Bachet [Dio21, Liber IIII, Quaestio XXXI] and proved by Lagrange [Lag72] in 1770, is the assertion that every integer is representable as a sum of four squares *without* an explicit expression for the number of representations.

In keeping with the emphasis on geometry, we explain how modular forms may be viewed as sections of certain canonical bundles on modular curves, and use the Riemann–Roch theorem to calculate some dimension formulae for spaces of modular forms.

For certain theta functions which will be of interest to us in later chapters, their transformation under the action of the correct arithmetic group is quite complicated. We therefore derive this transformation property for Jacobi's theta function in detail.

2.1 Elementary aspects of modular forms

Modular forms make their presence felt across many areas of modern number theory, so it is not surprising that there are a bewildering number of interrelated places to begin their story. We choose the most direct path: first we will give names to the most important subgroups of $SL(2,\mathbb{Z})$ of finite index, then we will define modular forms as functions on the complex manifold obtained by compactifying the quotient of $\mathscr{H}$ by the action of such subgroups.

2.1.1 Congruence subgroups of $SL(2,\mathbb{Z})$

Let N be a positive integer, and define

$$\Gamma(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2\mathbb{Z}) \mid a = d = 1 = \bmod N, b = c = 0 \bmod N \right\}.$$

This is a normal subgroup of $SL(2,\mathbb{Z})$. It has special prominence amongst the other subgroups will be introduced shortly as it is the kernel of the canonical reduction modulo N:

$$SL(2,\mathbb{Z}) \longrightarrow SL(2,\mathbb{Z}/N\mathbb{Z}).$$

We call $\Gamma(N)$ the *principal congruence subgroup of level* N, and we say that a subgroup of $SL(2,\mathbb{Z})$ is a *congruence subgroup of level* N if it is a subgroup of $\Gamma(N)$ for some N. The most important congruence subgroups are those of the following form:

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{Z}) \text{ such that } c = 0 \bmod N \right\},$$
$$\Gamma_1(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{Z}) \text{ such that } a = d = 1 \bmod N, c = 0 \bmod N \right\}.$$

For each N, the relations between the congruence subgroups is as follows:

$$\Gamma(N) \subseteq \Gamma_1(N) \subseteq \Gamma_0(N) \subseteq SL(2,\mathbb{Z}).$$

These subgroups all have finite index in $SL(2,\mathbb{Z})$, which we now compute.

Lemma 2.1.1.1. *The following statements hold for the indices of congruence subgroups:*

$$[SL(2,\mathbb{Z},\Gamma(N)] = N^3 \prod_{p|N} \left(1 - \frac{1}{p^2}\right), \tag{2.1}$$
$$[\Gamma_1(N) : \Gamma(N)] = N, \tag{2.2}$$
$$[\Gamma_0(N) : \Gamma_1(N)] = N \prod_{p|N} \left(1 - \frac{1}{p}\right), \tag{2.3}$$
$$[SL(2,\mathbb{Z},\Gamma_0(N)] = N \prod_{p|N} \left(1 + \frac{1}{p}\right). \tag{2.4}$$

Proof. To prove 2.1, note that $\Gamma(N)$ is the kernel of the canonical reduction modulo N homomorphism:

$$\Upsilon : SL(2,\mathbb{Z}) \to SL(2,\mathbb{Z}/N\mathbb{Z}).$$

One may check that Υ is surjective [DS05, Exercise 1.2.2], so the index is equal to the order of $SL(2,\mathbb{Z}/N\mathbb{Z})$. By the Chinese Remainder Theorem, if we write $N = \prod_i p_i^{\nu_i}$ as a product of distinct prime powers,

$$SL(2,\mathbb{Z}/N\mathbb{Z}) \cong \prod_i SL(2,\mathbb{Z}/p_i^{\nu_i}\mathbb{Z}),$$

so it suffices to show that $|SL(2,\mathbb{Z}/p^\nu\mathbb{Z})| = p^{3\nu}\left(1 - \frac{1}{p^2}\right)$. This may be proved by induction on ν, using the fact that the kernel of the canonical surjective map

$$\Upsilon_{p^\nu} : SL(2,\mathbb{Z}/p^{\nu+1}\mathbb{Z}) \to SL(2,\mathbb{Z}/p^\nu\mathbb{Z})$$

has order $p^{3\nu}$.

To verify 2.2, we observe that the map

$$\Xi : \Gamma_1(N) \to \mathbb{Z}/N\mathbb{Z}$$

given by $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \mapsto b \bmod N$ is surjective — $\left(\begin{smallmatrix} 1+(b-1)N & b+(b-1)N \\ N & 1+N \end{smallmatrix}\right)$ will do for a preimage of b — and has kernel $\Gamma(N)$.

We turn to 2.3. This time, the map

$$\Phi : \Gamma_0(N) \to (\mathbb{Z}/N\mathbb{Z})^\times$$

given by $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \mapsto d \mod N$ is surjective — if $ad = 1 + lN$, then $\left(\begin{smallmatrix} a & 1 \\ lN & d \end{smallmatrix}\right)$ is a preimage of d — and the kernel is equal to $\Gamma_1(N)$. The order of $(\mathbb{Z}/N\mathbb{Z})^\times$ is given by Euler's totient function $\phi(N)$, which may in turn be expressed as

$$\phi(N) = N \prod_{p|N} \left(1 - \frac{1}{p}\right).$$

Finally we prove 2.4. This is simply a matter of putting together the previous results:

$$[SL(2,\mathbb{Z}) : \Gamma_0(N)] = \frac{[SL(2,\mathbb{Z}) : \Gamma(N)]}{[\Gamma_0(N) : \Gamma_1(N)][\Gamma_1(N) : \Gamma(N)]} = \frac{N^3 \prod_{p|N}\left(1 - \frac{1}{p^2}\right)}{N^2 \prod_{p|N}\left(1 - \frac{1}{p}\right)} = N \prod_{p|N}\left(1 + \frac{1}{p}\right). \qquad \square$$

In general, it is rather difficult to describe congruence subgroups of $SL(2,\mathbb{Z})$ in terms of generators. As an example, which we will make use of in 3.1.2.3, we prove:

Proposition 2.1.1.2. *The congruence subgroup $\Gamma_0(4)$ is generated by $\pm T$ and U, where*

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \textit{and} \quad U = \begin{pmatrix} 1 & 0 \\ 4 & 1 \end{pmatrix}.$$

It follows that $\Gamma_0(4)$ is *equal* to the group generated by the indicated matrices.

Proof. The proof is by descent. For $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(4)$, we have the identity

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} T^n = \begin{pmatrix} * & * \\ c & nc+d \end{pmatrix}.$$

Since c is divisible by 4 and d is odd, if $c \neq 0$, we may choose n so that γT^{-n} has $|d'| < |c'|/2$, where (c', d') is the bottom row.

Similarly, the identity

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} U^n = \begin{pmatrix} * & * \\ 4nd+c & d \end{pmatrix}$$

implies that if $d \neq 0$, we may choose n so that the bottom row of γU^{-n} is (c', d'), where $|c'| < 2|d'|$.

At each step of this process, the quantity $\min\{|c|, 2|d|\}$ decreases, so eventually we must have c' or d' zero. Since we started out with a matrix in $\Gamma_0(4)$, and each of the generators is in $\Gamma_0(4)$, the case $d = 0$ is impossible. So at the final stage, our matrix is of the form $\left(\begin{smallmatrix} \pm 1 & * \\ 0 & \pm 1 \end{smallmatrix}\right)$, which is equal to $\pm T^n$ for some n. The claim follows. □

Let Γ be a congruence subgroup of $SL(2,\mathbb{Z})$ acting on the upper half plane $\mathscr{H}$ by Möbius transformations. Define the *j-factor* (or *factor of automorphy*) $j(\gamma, \tau) = (c\tau + d)$, where $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Note that, for arbitrary matrices γ_1 and γ_2 in $SL(2,\mathbb{Z})$, the j-factor satisfies the cocycle condition,

$$j(\gamma_1\gamma_2, z) = j(\gamma_1, \gamma_2 z) j(\gamma_2, z). \tag{2.5}$$

For a function f on $\mathscr{H}$ and k a positive integer, set

$$f\big|[\gamma]_k(\tau) = j(\gamma, \tau)^{-k} f(\gamma\tau).$$

We now define modular forms on congruence subgroups.

Definition 2.1.1.3. *A function $f : \mathscr{H} \to \mathbb{C}$ is a modular form of weight k with respect to Γ if*

1. *f is holomorphic on $\mathscr{H}$,*
2. *$f\big|[\gamma]_k(\tau) = f(\tau)$ for all $\gamma \in \Gamma$,*
3. *$f\big|[\gamma]_k$ is holomorphic at infinity for all $\gamma \in SL(2,\mathbb{Z})$.*

If, in addition to these requirements, it also transpires that $a_0(f) = 0$ in the Fourier expansion of $f\big|[\gamma]_k$ (described below) for all $\gamma \in SL(2,\mathbb{Z})$, then f is called a cusp form.

The third condition requires some explaining. One easily checks that $f\big|[\gamma]_k$ is $\mathbb{Z}$-periodic, and the $\mathbb{Z}$-periodic map $\tau \mapsto e^{2\pi i\tau}$ effects a biholomorphism between $\mathscr{H}$ and the open unit disk punctured at 0, $\mathbb{D} \setminus \{0\}$. This map identifies the function $f\big|[\gamma]_k(\tau)$ with the Laurent series

$$g(q) = \sum_{n \in \mathbb{Z}} a_n q^n$$

on $\mathbb{D} \setminus \{0\}$. Since $q \to 0$ as $\mathrm{Im}(\tau) \to \infty$, we say that $f\big|[\gamma]_k$ is holomorphic at infinity if g extends holomorphically to all of $\mathbb{D}$. Equivalently, the coefficients in the Laurent expansion of g corresponding to negative powers vanish.

Taking $\gamma = I$ in (3), and bearing in mind the discussion above, we see that f has a Fourier series expansion

$$f(\tau) = \sum_{n=0}^{\infty} a_n(f) e^{2\pi i n\tau/N}, \tag{2.6}$$

where N is the level of Γ. We refer to this as the Fourier expansion at infinity if there is potential for confusion with Fourier expansions at other points.

We also note that the condition (3) need only be checked for the (finitely many) representatives of the cosets of Γ in $SL(2,\mathbb{Z})$.

If f satisfies conditions (1) and (2), then condition (3) is equivalent to the following statement [DS05, Proposition 1.2.4], which is often more easily checked in practice.

3.' *Suppose that in the Fourier expansion at infinity (2.6), there exist positive constants C and k such that for all $n > 0$,*

$$|a_n(f)| \leq Cn^k.$$

One easily checks that the sum of two modular forms (resp. cusp forms) of the same weight is a modular form (resp. cusp form), and the product of a modular form (resp. cusp form) of weight k with a modular form (resp. cusp form) of weight l is a modular form (resp. cusp form) of weight $k+l$. We summarise these observations:

Definition 2.1.1.4. *Let Γ be a congruence subgroup. The set of modular forms of weight k is a complex vector space, denoted*

$$\mathscr{M}_k(\Gamma),$$

and the set of cusp forms of weight k is a vector subspace, denoted

$$\mathscr{S}_k(\Gamma).$$

The set of modular forms for Γ of any weight forms a graded algebra over $\mathbb{C}$:

$$\mathscr{M}(\Gamma) = \bigoplus_{n\geq 0} \mathscr{M}_k(\Gamma),$$

and the graded subalgebra of cusp forms is denoted

$$\mathscr{S}(\Gamma) = \bigoplus_{n\geq 0} \mathscr{S}_k(\Gamma).$$

The set of meromorphic functions on $\mathscr{H}$ satisfying condition (2), and condition (3) with the word "holomorphic" replaced by "meromorphic", is denoted $\mathscr{A}_k(\Gamma)$. This vector space is often called the space of *weakly modular* forms of weight k for Γ.

An important fact is that any space of modular forms of fixed weight is finite dimensional. We will see in Subsection 2.1.3 that the space of modular forms of weight k on $\mathscr{H}$ for a congruence subgroup Γ corresponds to the space of global meromorphic sections (with poles of prescribed order at prescribed points) of certain vector bundles over the modular curve $X(\Gamma)$. By Theorem 1.1.1.17, $X(\Gamma)$ is compact, so the Riemann–Roch theorem (Theorem 2.1.4.3) implies that any such space is finite-dimensional. One may check that conditions (1) and (2) of Definition 2.1.1.3 are enough to ensure that $\mathscr{M}_k(\Gamma)$ may be identified with the space of global sections of a certain bundle over $\Gamma\backslash\mathscr{H}$: the point of condition (3) is to ensure that $\mathscr{M}_k(\Gamma)$ may be identified with the space of global sections of a vector bundle over the *compactification* of $\Gamma\backslash\mathscr{H}$.

2.1.2 Eisenstein series and cusp forms

The space of modular forms of fixed weight for a congruence subgroup Γ may be naturally decomposed into interesting subspaces. One such subspace is the set of *cusp forms*, which we met in Definition 2.1.1.3, and the other is the space of *Eisenstein series*, which is denoted

$$\mathscr{E}_k(\Gamma).$$

The space of Eisenstein series is in some sense complementary to the space of cusp forms: whereas the cusp forms are characterised by vanishing at every cusp, each cusp C gives rise to an Eisenstein series by "averaging" very simple functions over the action by $\Gamma_C\backslash\Gamma$. We will carry out this procedure for $SL(2,\mathbb{Z})$ in some detail, in order to foreshadow the more complicated situation for metaplectic Eisenstein series, described in Chapter 6.

Just for this subsection, set $\Gamma = SL(2,\mathbb{Z})$. The orbit of ∞ under Γ is the entire set $\mathbb{Q} \cup \{\infty\}$, so there is only one cusp for $X(\Gamma)$. An easy calculation shows that the isotropy subgroup of ∞ is

$$\Gamma_\infty = \left\{ \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \in \Gamma \text{ such that } b \in \mathbb{Z}. \right\}$$

For any positive integer k, the kth power of the j-factor satisfies the *cocycle condition*:

$$j(\gamma_1\gamma_2, z)^k = j(\gamma_1, \gamma_2 z)^k j(\gamma_2, z)^k. \tag{2.7}$$

Therefore, we expect that the following function, if it converges, to define a modular form:

$$E_k(z) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} j(\gamma, z)^{-k}.$$

If $k > 2$, then the sum converges absolutely and uniformly on any compact subset of $\mathscr{H}$, so by Morera's theorem, E_k is holomorphic for such k. If k is odd, then the map $\gamma \mapsto -\gamma$ is an automorphism of $SL(2,\mathbb{Z})$ with the property that $j(\gamma, z)^k \mapsto -j(\gamma, z)^k$. So we may assume that k is even and $k \geq 4$.

Proposition 2.1.2.1. *Let $k \geq 3$ be an integer. Then*

$$E_k \in \mathscr{M}_k(\Gamma).$$

Proof. Since $k \geq 3$, we may rearrange the sum at will. Let $\gamma \in \Gamma$.

$$\begin{aligned} E_k(\gamma z) &= \sum_{\kappa \in \Gamma_\infty \backslash \Gamma} j(\kappa, \gamma z)^{-k} \\ &= j(\gamma, z)^k \sum_{\kappa \in \Gamma_\infty \backslash \Gamma} j(\kappa\gamma, z)^{-k} \\ &= j(\gamma, z)^k \sum_{\kappa \in \Gamma_\infty \backslash \Gamma} j(\kappa, z)^{-k}, \end{aligned}$$

where we have used 2.7 to arrive at the second line, and noted that $\Gamma_\infty \backslash \Gamma\gamma = \Gamma_\infty \backslash \Gamma$ to arrive at the final expression. ☐

Suppose that $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right), \left(\begin{smallmatrix} x & y \\ c & d \end{smallmatrix}\right) \in \Gamma$. Then

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x & y \\ c & d \end{pmatrix}^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -y \\ -c & x \end{pmatrix} = \begin{pmatrix} ad - bc & bx - ay \\ cd - dc & dx - cy \end{pmatrix} = \begin{pmatrix} 1 & bx - ay \\ 0 & 1 \end{pmatrix}.$$

Also, when k is even, the map $\gamma \mapsto -\gamma$ is an automorphism of Γ which preserves $j(\gamma, z)^{2k}$. Together with the observation that if $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$, then $(c, d) = 1$, we have proved:

Lemma 2.1.2.2. *Suppose $k \geq 2$ is an integer. Then*

$$E_{2k}(z) = \frac{1}{2} \sum_{\substack{(c,d) \in \mathbb{Z}^2 \\ (c,d) = 1}} \frac{1}{(cz + d)^{2k}}.$$

We now define the Eisenstein series for $SL(2,\mathbb{Z})$ in a form with which we may compute the Fourier expansions.

Definition 2.1.2.3. *Let $k \geq 3$. Then the Eisenstein series for $SL(2,\mathbb{Z})$ of weight k is*

$$G_k(z) = \sum_{\substack{(c,d) \in \mathbb{Z}^2 \\ (c,d) \neq (0,0)}} \frac{1}{(cz + d)^k}.$$

For $k = 2$, we define

$$G_2(z) = \sum_{c \in \mathbb{Z}} \sum_{\substack{d \in \mathbb{Z} \\ d \neq 0 \text{ if } c = 0}} \frac{1}{(cz + d)^2}.$$

By the same arguments as above, for $k \geq 3$, G_k is holomorphic and vanishes for odd k. For even $k \geq 3$, G_k is a multiple of E_k:

$$\begin{aligned} G_k(z) &= \sum_{\substack{(c,d)\in\mathbb{Z}^2 \\ (c,d)\neq(0,0)}} \frac{1}{(cz+d)^k} \\ &= \sum_{m=1}^{\infty} \sum_{\substack{(c,d)\in\mathbb{Z}^2 \\ (c,d)=m}} \frac{1}{(cz+d)^k} \\ &= \sum_{m=1}^{\infty} \frac{1}{m^k} \sum_{\substack{(c,d)\in\mathbb{Z}^2 \\ (c,d)=1}} \frac{1}{(cz+d)^k} \\ &= 2\zeta(k)E_k(z), \end{aligned}$$

where we have used Lemma 2.1.2.2. It follows that $G_k \in \mathscr{M}_k(\Gamma)$ for $k \geq 3$. Since the sum defining G_k isn't absolutely convergent when $k = 2$, we cannot conclude that G_2 is a modular form: indeed, it transforms as a *quasimodular form* under the action of Γ. We will not define them here, but the reader will be pleased to know that there is an extensive literature on quasi-modular forms [CS17, Subsection 5.1.3; Mov12; Zag08].

However, we may compute the Fourier expansion of G_k for all even $k \geq 2$ at once.

Proposition 2.1.2.4. *For all integers $k \geq 1$,*

$$G_{2k}(z) = 2\zeta(2k) + \frac{2(2\pi i)^{2k}}{(k-1)!} \sum_{n=1}^{\infty} \sigma_{2k-1}(n) e^{2\pi i n z},$$

where $\sigma_l(n)$ is the lth-sum-of-divisors function, *defined by*

$$\sigma_l(n) = \sum_{m|n} m^l.$$

Proof. For any integer $k \geq 2$, we may rewrite the Eisenstein series as

$$G_k(z) = 2\zeta(k) + \sum_{\substack{c\in\mathbb{Z} \\ c\neq 0}} \sum_{d\in\mathbb{Z}} \frac{1}{(cz+d)^k}.$$

The inner sum, over d, is clearly invariant under $z \mapsto z+1$, and it converges to a continuous function, so it has a Fourier expansion:

$$\sum_{d\in\mathbb{Z}} \frac{1}{(cz+d)^{2k}} = \sum_{n\in\mathbb{Z}} a_n(y) e^{2\pi i n x}, \text{ where } a_n(y) = \int_0^1 e^{-2\pi i n u} \left(\sum_{d\in\mathbb{Z}} \frac{1}{(cu+icy+d)^{2k}} \right) du.$$

Next, we simplify the integral:

$$\begin{aligned} \int_0^1 e^{-2\pi i n u} \left(\sum_{d\in\mathbb{Z}} \frac{1}{(cu+icy+d)^{2k}} \right) du &= c^{-2k} \int_0^1 e^{-2\pi i n u} \left(\sum_{d\in\mathbb{Z}} \frac{1}{(u+iy+d/c)^{2k}} \right) du \\ &= c^{-2k} \sum_{d \mod c} e^{2\pi i n d/c} \int_{\mathbb{R}} \frac{e^{-2\pi i n u}}{(u+iy)^{2k}} du. \end{aligned}$$

Now we evaluate the integral

$$\int_{\mathbb{R}} \frac{e^{-2\pi i n u}}{(u+iy)^{2k}} du.$$

If $n \leq 0$, then the integrand is holomorphic and decreases rapidly as $|u| \to \infty$ for $u \in \mathscr{H}$. Therefore the integral vanishes by the residue theorem. If $n > 0$, the integrand still decreases rapidly as $|u| \to \infty$ for u in the lower half plane, and there is a residue at $u = -iy$. The residue is easily computed:

$$-\frac{2\pi i}{(k-1)!} \frac{d^{2k-1}}{du^{2k-1}} e^{-2\pi i n u} \bigg|_{u=-iy} = \frac{(2\pi i)^{2k}}{(k-1)!} n^{2k-1} e^{-2\pi n y},$$

and upon noting that $\sum_{d \bmod c} e^{2\pi i n d/c}$ is equal to $|c|$ if $c \mid n$ and 0 otherwise, it follows that

$$\sum_{d\in\mathbb{Z}} \frac{1}{(cz+d)^{2k}} = \frac{(2\pi i)^{2k}}{(k-1)!} \sum_{\substack{n\geq 1\\ c|n}} \frac{|c|n^{2k-1}}{c^{2k}} e^{2\pi i n z}$$

Summing over nonzero c and noting that as c runs over all the divisors of n, so does n/c, we have the complete Fourier expansion for $G_{2k}(z)$:

$$G_{2k}(z) = 2\zeta(2k) + \frac{2(2\pi i)^{2k}}{(k-1)!} \sum_{n=1}^{\infty} \sigma_{2k-1}(n) e^{2\pi i n z}. \qquad \square$$

Cusp forms are much more difficult to get hold of. They are distinguished by the fact that their Fourier coefficients grow noticeably more slowly than those for Eisenstein series.

Proposition 2.1.2.5. *Let $f \in \mathscr{S}_k(\Gamma)$ for some congruence subgroup Γ, with Fourier expansion*

$$f(z) = \sum_{n=1}^{\infty} a_n e^{2\pi i n z}.$$

Then $a_n = O(n^{k/2})$.

Proof. First, note that $F(z) = y^{\frac{k}{2}}|f(z)|$ is bounded on $\mathscr{H}$, since F is invariant under the action of Γ and decays exponentially at the cusps. It follow that $f(z) = O(y^{-\frac{k}{2}})$ for all $z \in \mathscr{H}$, where the implied constant depends on f. By Cauchy's integral formula,

$$a_n = \frac{1}{2\pi i} \int \frac{\tilde{f}(q)}{q^{n+1}} dq,$$

where $\tilde{f}(e^{2\pi i z}) = f(z)$, and the integral is taken around a small circle of radius r centered at the origin. By preservation of inequalities for line integrals, for any $r < 1$, we have

$$|a_n| \leq r^{-n} \sup_{|e^{2\pi i z}|=r} f(z) = e^{2\pi n y} O\left(y^{-\frac{k}{2}}\right).$$

The last term is minimised when $y = 1/n$, which proves the claim. $\square$

Remark 2.1.2.6. *The best possible bound for Proposition 2.1.2.5, $a_n = O(n^{\frac{k-1}{2}})$, was known as the Ramanujan conjecture*[4]*, until it was proven in 1971 by Deligne [Del71]. The appropriate generalisation to automorphic forms on $GL(n)$ is still unproven, but is known to follow from a special case of Langlands' conjectures [Lan70].*

2.1.3 Modular forms as sections of line bundles

The reader may well wonder why, in Definition 2.1.1.3, we stipulate that a modular form must transform with the particular factor of automorphy $(cz + d)^k$. The simplest kind of function which transforms in such a way under the action of the modular group is surely one which satisfies $f(\gamma z) = f(z)$, but as the reader may check, all such functions f must be constant by Liouville's Theorem[5].

But a geometer will recognise that the factor of automorphy appears for good reason: indeed, the factor of automorphy is precisely the sort of thing that arises if we pull back a section of a line bundle to a function via local coordinates. The point of this subsection is to explain how some modular forms may be viewed as sections of line bundles.

Unfortunately, modular forms do not correspond to *holomorphic* sections of line bundles. Instead, they correspond to meromorphic sections with poles of prescribed order permitted only at certain places. We explain how to define orders of poles of meromorphic sections, and then we formalise the idea of restricting the location and severity of poles using the notion of a *divisor*.

[4]Ramanujan only made the conjecture [Ram16] for $\mathscr{S}_{12}(SL(2,\mathbb{Z}))$, where it is equivalent to the bound $\tau_n = O\left(n^{\frac{11}{2}}\right)$ for the Fourier coefficients τ_n of the modular discriminant Δ (cf. the introduction to this chapter).

[5]If we drop the condition that f be holomorphic, then such functions do exist, and are called modular functions. They are alternatively realised as quotients of modular forms of the same weight.

Definition 2.1.3.1. *Suppose that X is a Riemann surface and that E is a holomorphic vector bundle over X. Let s be a meromorphic section of E; that is, an element of $\mathscr{O}(E) \otimes_{\mathscr{O}(X)} \mathscr{K}(X)$. For every $x \in X$, there is a neighbourhood U of x over which s is represented in coordinates as $h(x)\tilde{s}(x)$, where h is a meromorphic function on X and $\tilde{s}$ is a holomorphic function from x into $\mathbb{C}^k$. At points where h has a pole of order n, we say that s has a pole of order n, and we also define $ord_x(s) = ord_x(h)$. These definitions may be easily checked to be independent of the choices of U and h.*

We note that if X is compact, there are only finitely many points $x \in X$ for which there does not exist a neighbourhood over which s is actually a holomorphic section.

Definition 2.1.3.2. *A divisor is an expression of the form*

$$\mu = \sum_{x \in X} \mu_x \cdot x,$$

where $\mu_x \in \mathbb{Z}$ for all x and $\mu_x = 0$ for all but finitely many x. The degree $\deg(\mu)$ of a divisor μ is defined to be the quantity $\sum_{x \in X} \mu_x$. The set of all divisors form a $\mathbb{Z}$-module $\mathrm{Div}(X)$ under pointwise addition, and the map $\deg : \mathrm{Div}(X) \to \mathbb{Z}$ is a group homomorphism. If E is a holomorphic vector bundle over a compact Riemann surface X and s is a meromorphic section, the divisor of s, $\mathrm{div}(s)$, is defined to be

$$\mathrm{div}(s) = \sum_{x \in X} ord_x(s) \cdot x \in \mathrm{Div}(X),$$

where $ord_x(s)$ is from Definition 2.1.3.1. For μ and ν in $\mathrm{Div}(X)$, the relation

$$\mu \geq \nu \text{ if } \deg(\mu - \nu) \geq 0$$

is a well-defined partial order.

For any divisor $\mu \in \mathrm{Div}(X)$, we define

$$L(D) = \mathscr{K}(X(\Gamma))[D],$$

and set $\ell(D) = \dim_{\mathbb{C}} L(D)$.

Definition 2.1.3.3. *Suppose that X is a compact Riemann surface and E is a holomorphic vector bundle. Let $\mu \in \mathrm{Div}(X)$. Then $\Gamma(X, E)[\mu]$ is defined to be the complex vector space of meromorphic sections s of E such that $\mathrm{div}(s) \geq \mu$ for $s \neq 0$.*

Theorem 2.1.3.4. *Let $k \geq 0$ be an integer, and Γ be a congruence subgroup of $SL(2, \mathbb{Z})$. Then the natural map $\pi : \mathscr{H} \to X(\Gamma)$ induces an isomorphism*

$$\pi^* : \otimes^k \Omega^1 X(\Gamma) \longrightarrow \mathscr{A}_{2k}(\Gamma).$$

Proof. Let ω be an element of $\otimes^k \Omega^1 X(\Gamma)$. We will show first that the pullback of ω by π can be identified with an element of $\mathscr{A}_{2k}(\Gamma)$. Indeed, $\pi^*(\omega)$ is invariant under the action on Γ: let $\gamma \in \Gamma$, $\tau \in \mathscr{H}$ and $X \in \otimes^k T\mathscr{H}$ and calculate:

$$\pi^*\omega\big|_{\gamma\tau} X\big|_{\gamma\tau} = \omega(d\pi\big|_{\gamma\tau} X\big|_{\gamma\tau}) = \omega(d\pi(X)\big|_{\pi(\gamma\tau)}) = \omega(d\pi(X)\big|_{\pi(\tau)}) = \pi^*\omega\big|_{\tau} X\big|_{\tau}.$$

Since ω is an element of $\otimes^k \Omega^1 \mathscr{H}$, it is of the form $f(\tau) \otimes^k d\tau$ for some meromorphic function f on $\mathscr{H}$. Viewing the action of γ as a biholomorphism of $\mathscr{H}$ and recognising that $\gamma^*\pi^*\omega\big|_{\tau} X\big|_{\tau} = \pi^*\omega\big|_{\gamma\tau} X\big|_{\gamma\tau}$, we have

$$\begin{aligned} f(\tau) \otimes^k d\tau &= \pi^*\omega \\ &= \gamma^*\pi^*\omega \\ &= \gamma^*\left(f(\tau) \otimes^k d\tau\right) \\ &= f(\gamma\tau)\left(\frac{d\gamma(\tau)}{d\tau}\right)^k \otimes^k d\tau \\ &= j(\gamma, \tau)^{-2k} f(\gamma\tau) \otimes^k d\tau. \end{aligned}$$

So f satisfies condition (2) of Definition 2.1.1.3. In order to show that it is in $\mathscr{A}_{2k}(\Gamma)$, we need to verify that it is meromorphic at the cusps. But this is immediate from the calculation above:

$$j(\gamma,\tau)^{-2k} f(\gamma\tau) \otimes^k d\tau\big|_{\tau=\infty} = \gamma^*\pi^*\omega\big|_{\tau=\infty} \tag{2.8}$$

and γ is a biholomorphism whilst ω is meromorphic at $\Gamma\infty$ by assumption.

We now need to verify that if $f \in \mathscr{A}_{2k}(\Gamma)$, then $f(\tau) \otimes^k d\tau \in \otimes^k\Omega^1(\Gamma\backslash X)$ extends to a section of $\otimes^k\Omega^1 X(\Gamma)$. Clearly, $f(\tau)\otimes^k d\tau$ is a section of the indicated bundle; we need to check that it is meromorphic at the images of the cusps. But this is also immediate from 2.8. □

Theorem 2.1.3.4 doesn't quite hold as stated with the word "meromorphic" replaced by holomorphic. But we can easily work out the correct statement by quantifying the non-holomorphicness of the indicated sections using divisors.

Theorem 2.1.3.5. *Let $k \geq 0$ be an integer, and Γ be a congruence subgroup of $SL(2,\mathbb{Z})$. Let μ and ν denote the divisors*

$$\mu = -k\sum_i (1-e_i^{-1})E_i - k\sum_j C_j,$$
$$\nu = -k\sum_i (1-e_i^{-1})E_i - (k-1)\sum_j C_j,$$

where the E_i are the distinct images of the elliptic points in $X(\Gamma)$ and C is the number of inequivalent cusps. Then the natural map $\pi : \mathscr{H} \to X(\Gamma)$ induces isomorphisms

$$\pi^* : \otimes^k\Omega^1 X(\Gamma)[\mu] \longrightarrow \mathscr{M}_{2k}(\Gamma),$$
$$\pi^* : \otimes^k\Omega^1 X(\Gamma)[\nu] \longrightarrow \mathscr{S}_{2k}(\Gamma).$$

Proof. If $P = \pi(Q) \in X(\Gamma)$ is neither a cusp nor an elliptic point, then the quotient map $\pi : \mathscr{H} \to \Gamma\backslash\mathscr{H}$ is a local biholomorphism, so $\mathrm{ord}_Q(f) = \mathrm{ord}_P(f(\tau)\otimes^k d\tau)$.

If $E = \pi(Q)$ is the image of an elliptic point, then we saw in the alternative proof of Theorem 1.1.1.17 that there is a chart around E which is locally an e-fold covering of disks, where e is the order of the reduced isotropy subgroup of Γ at E. We calculate the order of our one-form at E_i:

$$f(\tau^e)\otimes^k d(\tau^e) = f(\tau^e)\otimes^k \left(e\tau^{e-1}d\tau\right) = e\tau^{e-1} f(\tau^e)\otimes^k d\tau,$$

so $\mathrm{ord}_Q(f) = e\cdot \mathrm{ord}_E(f(\tau)\otimes^k d\tau) + k(e-1)$.

If $C = \pi(Q)$ is the image of a cusp, then near C we have a local coordinate defined implicitly by $\phi : \mathscr{H} \to \mathbb{D}\setminus\{0\}$, where $\phi(\tau) = e^{2\pi i\tau/h}$. We calculate the order of the one-form:

$$f(\phi(\tau))\otimes^k d(\phi(\tau)) = f(\phi(\tau))\otimes^k \left(\frac{2\pi i}{h}e^{2\pi i\tau/h}d\tau\right) = \left(\frac{2\pi i}{h}\right)^k \phi(\tau)^k f(\phi(\tau))\otimes^k d\tau,$$

so $\mathrm{ord}_Q(f) = \mathrm{ord}_C(f(\tau)\otimes^k d\tau) + k$.

If we are interested in cusp forms, we need only take into account the condition that a cusp form vanish to order 1 at all cusps. Therefore, if f is a cusp form of weight $2k$, we have $1 = \mathrm{ord}_Q(f) = \mathrm{ord}_C(f(\tau)\otimes^k d\tau) + k$. □

2.1.4 Dimension formulae and the Riemann–Roch Theorem

In the last subsection, we saw that, for a congruence subgroup Γ we may identity the space of modular forms of even weight with the space of meromorphic sections of a certain vector bundle with restrictions on the locations and orders of the possible poles.

Since $X(\Gamma)$ is a complex manifold of dimension one, $\Omega^1(X(\Gamma))$ is a vector bundle of rank one: a line bundle. If we select any nonzero $s_0 \in \Omega^1 X(\Gamma)$, all $s \in \otimes^k\Omega^1 X(\Gamma)$ are of the form $s = fs_0^k$ for some $f \in \mathscr{K}(X(\Gamma))$, where s_0^k is short for $\otimes^k s_0$. Consequently,

$$\mathscr{M}_{2k}(\Gamma) \cong \otimes^k\Omega^1 X(\Gamma)[\mu] = \{fs_0^k \mid \mathrm{div}(fs_0^k) \geq \mu\},$$

where μ is the divisor from Theorem 2.1.3.5, and the last set is isomorphic as a complex vector space to

$$\{f \in \mathscr{K}(X(\Gamma)) \mid \mathrm{div}(f) \geq \mu - \mathrm{div}(s_0^k)\}.$$

As it stands, μ is a rational divisor: that is, $\mu \in \mathrm{Div}_{\mathbb{Q}}(X) = \mathrm{Div}(X) \otimes_{\mathbb{Z}} \mathbb{Q}$. However, since f is a meromorphic function on $X(\Gamma)$ and s_0 is a holomorphic section, both $\mathrm{div}(f)$ and $\mathrm{div}(s_0^k)$ are integral divisors, so

$$\mathrm{div}(f) \geq \mu - \mathrm{div}(s_0^k) \text{ if and only if } \mathrm{div}(f) \geq \lfloor \mu \rfloor - \mathrm{div}(s_0^k),$$

where

$$\lfloor \mu \rfloor = \sum_{x \in X(\Gamma)} \lfloor \mu_x \rfloor \cdot x.$$

In the end, using the notation of 2.1.3.2, we have

$$\mathscr{M}_{2k}(\Gamma) = L(\mathrm{div}(s_0^k) - \lfloor \mu \rfloor),$$

and the analogous statement applies for cusp forms:

$$\mathscr{S}_{2k}(\Gamma) = L(\mathrm{div}(s_0^k) - \lfloor \nu \rfloor),$$

where ν is the divisor defined in Theorem 2.1.3.5.

We will now state the Riemann–Roch theorem, which will allow us to compute the dimensions of $\mathscr{M}_{2k}$ and $\mathscr{S}_{2k}$ based on the information above. There are many good sources for information on the Riemann–Roch theorem [Gun66, Chapter 7; Har77, Chapter IV].

Definition 2.1.4.1. *A* canonical divisor K *on a compact Riemann surface is a divisor of the form* $\mathrm{div}(\lambda)$*, for some one-form* $\lambda \in \Omega^1(X)$*.*

Definition 2.1.4.2. *The* genus g *of a compact Riemann surface is defined to be*

$$g = \dim H^0(X, \mathscr{O}),$$

where $H^0(X, \mathscr{O})$ *is the (finite-dimensional) space of global holomorphic functions on* X*.*

There are many equivalent definitions of the genus for a surface, but from the point of view of the Riemann–Roch theorem, this one is the most natural. One may prove that for a triangulation of X with V vertices, E edges and F faces, one has

$$2g - 2 = V - E + F. \tag{2.9}$$

Theorem 2.1.4.3 (Riemann–Roch)**.** *Let* X *be a compact Riemann surface of genus* g*. Let* K *be a canonical divisor. Then for any divisor* $D \in \mathrm{Div}(X)$*,*

$$\ell(D) - \ell(K - D) = \deg(D) - g + 1.$$

Some immediate consequences are: if we set $D = 0$, then $1 + \ell(K) = 0 - g + 1$, so $\ell(K) = g$. If $D = K$, then $g - 1 = \deg(K) - g + 1$, so $\deg(K) = 2g - 2$. If $\deg(D) > 2g - 2$, then for all $f \in \mathscr{K}(X)$, $\deg(\mathrm{div}(f) - K + D) < 0$ (where we have used that $\deg\ \mathrm{div}(f) = 0$ and $\deg(K) = 2g - 2$), so $\ell(K - D) = 0$. This leads to:

Corollary 2.1.4.4. *If* $\deg(D) > 2g - 2$*,*

$$\ell(D) = \deg(D) - g + 1.$$

Now we turn our attention to the problem of calculating the dimension of $\mathscr{M}_{2k}(\Gamma)$. We verify that for $X = X(\Gamma)$ and

$$D = \sum_i \lfloor k(1 - e_i^{-1}) \rfloor E_i + k \sum_j C_j + \mathrm{div}(s_0^k),$$

the hypothesis of Corollary 2.1.4.4 is satisfied. Note first that $\mathrm{div}(s_0)$ is a canonical divisor on $X(\Gamma)$, so by the discussion preceding Corollary 2.1.4.4, $\deg\ \mathrm{div}(s_0) = 2g - 2$. It follows that $\deg\ \mathrm{div}(s_0^k) = 2k(g - 1)$, so for $k \geq 1$,

$$\deg(D) = \sum_i \lfloor k(1 - e_i^{-1}) \rfloor + kC + 2k(g - 1) > 2g - 2, \tag{2.10}$$

where C is the (obviously positive) number of inequivalent cusps.

We repeat the calculation for the cusp forms. Set

$$D = \sum_i \lfloor k(1-e_i^{-1}) \rfloor E_i + (k-1)\sum_j C_j + \mathrm{div}(s_0^k). \tag{2.11}$$

If $k \geq 2$, then the degree is sufficiently large for us to apply Corollary 2.1.4.4:

$$\deg(D) = \sum_i \lfloor k(1-e_i^{-1}) \rfloor E_i + (k-1)C + 2k(g-1) > 2g-2. \tag{2.12}$$

If $k = 1$, then the divisor D at 2.11 reduces to $\mathrm{div}(s_0^k)$, and is therefore canonical, so $\deg(D) = 2g-2$.

We assumed that $k \geq 1$ to obtain 2.10 and 2.12, but the other cases are easily dealt with. The isomorphism

$$\mathscr{M}_0(\Gamma) \cong \mathscr{O}(X(\Gamma))$$

implies that $\mathscr{M}_0(\Gamma) = \mathbb{C}$, since the set of holomorphic functions on a compact complex manifold consists of only the constant functions, whilst no modular form of negative weight can be holomorphic. We collect together these observations and apply Corollary 2.1.4.4 to obtain:

Proposition 2.1.4.5. *For $k < 0$, the spaces $\mathscr{M}_k(\Gamma)$ are zero-dimensional; $\mathscr{M}_0(\Gamma)$ consists of the constant functions, and for $k \geq 1$,*

$$\dim \mathscr{M}_{2k}(\Gamma) = \sum_i \lfloor k(1-e_i^{-1}) \rfloor + kC + (2k-1)(g-1),$$
$$\dim \mathscr{S}_{2k}(\Gamma) = \begin{cases} k(1-e_i^{-1})\rfloor + (k-1)C + (2k-1)(g-1) & k \geq 2, \\ g & k = 1, \end{cases}$$

where the sum runs over the distinct elliptic points, e_i is the order of the reduced isotropy subgroup (Definition 1.1.1.3) of the ith elliptic point, and C is the number of inequivalent cusps for Γ.

There are analogous formulae for the dimension of spaces of modular forms and cusp forms of integral (not necessarily even) weight k. These formulae are achieved by considering the meromorphic one-form $f(\tau^2) \otimes^k d\tau$ and using the Riemann–Roch theorem. However, the details are cumbersome — one must differentiate between *regular* and *irregular* cusps for Γ — so we refer the reader to the usual literature [DS05, Chapter 3 Section 6; Shi71, Chapter 2].

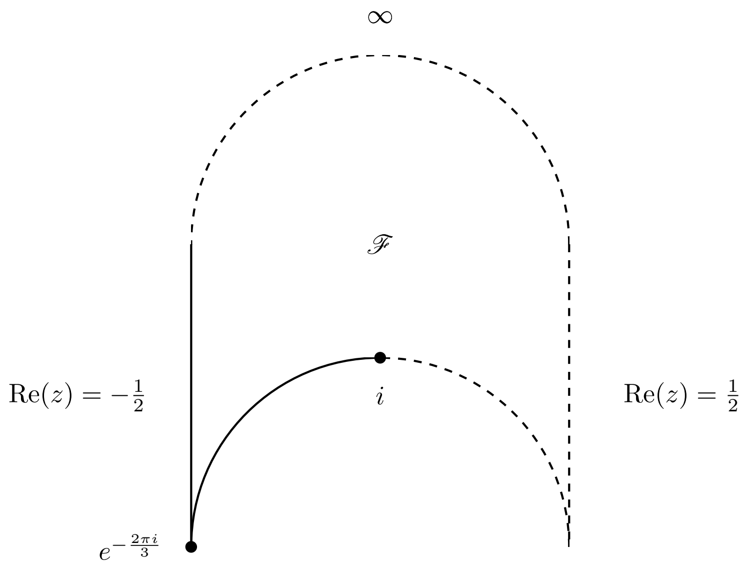
The rest of this subsection is devoted to working out explicit formulae for the genus of $X(\Gamma)$, so that we may use Proposition 2.1.4.5 to calculate the dimensions of spaces of modular forms for, say, $\Gamma_0(4)$.

Proposition 2.1.4.6. *The Riemann surface $X(SL(2,\mathbb{Z}))$ has genus zero.*

Proof. Recall the fundamental domain for $SL(2,\mathbb{Z})$ from Proposition 1.1.1.16. Algebraically, it is described as

$$\mathscr{F} = \left\{ x \in \mathscr{H} \mid \mathrm{Re}(x) < \frac{1}{2}, |x| > 1 \right\} \cup \left\{ -\frac{1}{2} + it \in \mathscr{H} \mid t > \frac{\sqrt{3}}{2} \right\} \cup \left\{ e^{i\theta} \mid \frac{\pi}{2} \leq \theta \leq \frac{2\pi}{3} \right\},$$

where we have taken care to determine which segments of edges should be included. The fundamental domain is compactified by identifying the vertical left-hand boundary with the vertical right-hand boundary, gluing together the segments of the semicircles in the first and second quadrants (in such a way that a path following the circle from $e^{\frac{2\pi i}{3}}$ to i in the *inside* of the fundamental domain is identified with a path following the circle from $e^{\frac{\pi i}{3}}$ to i on the *outside* of the fundamental domain), and adding a point at $i\infty$.



Therefore, keeping in mind the boundary identifications, we may triangulate the fundamental domain as follows. Take the point at infinity, i, $e^{\frac{2\pi i}{3}}$ and some other point in the interior of the fundamental domain as vertices. Upon introducing edges between each of these four vertices, we have five edges in total: one for both the vertical sides, one for the two arcs of the semicircular boundary, and three through the interior of $\mathscr{F}$ to connect the vertices on the boundary to the inner vertex.

Upon applying 2.9, we have

$$2 - 2g = V - E + F = 4 - 5 + 3 = 2,$$

so it follows that $g = 0$. □

We now determine the elliptic points of $SL(2,\mathbb{Z})$. Suppose $\tau \in \mathscr{H}$ is fixed by some matrix $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, $\gamma \neq \pm I$. Then $a\tau + b = c\tau^2 + d\tau$. With the additional constraint that $\tau \in \mathscr{H}$, this implies that $c \neq 0$ and $|a + d| < 2$. Therefore $a = \pm 1 - d$ or $a = -d$. In the first case, $\gamma^2 + I = 0$ and in the second case, $\gamma^2 \pm \gamma + I = 0$. It follows that $\gamma^4 = I$ or $\gamma^6 = I$, so the order of γ is one of $1, 2, 3, 4$ or 6.

Lemma 2.1.4.7. *Let $\gamma \in SL(2,\mathbb{Z})$. If γ has order 3 then γ is conjugate to $\left(\begin{smallmatrix} 0 & 1 \\ -1 & -1 \end{smallmatrix}\right)^{\pm 1}$. If γ has order 4 then γ is conjugate to $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)^{\pm 1}$. If γ has order 6 then γ is conjugate to $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right)^{\pm 1}$.*

The proof is a tedious computation [DS05, Proposition 2.3.3].

Now we can determine the elliptic points for $SL(2,\mathbb{Z})$.

Lemma 2.1.4.8. *The elliptic points for $SL(2,\mathbb{Z})$ are represented by the orbits of i and $\rho = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$. The isotropy subgroup of i is $\left\langle \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right) \right\rangle$ and the isotropy subgroup of ρ is $\left\langle \left(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right) \right\rangle$.*

Proof. Note that i and ρ are the fixed points of the matrices in Lemma 2.1.4.7. The statement about the isotropy subgroup of i is easily checked since $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)i = i$ implies that $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} d & -c \\ c & d \end{smallmatrix}\right)$, and $d^2 + c^2 = 1$ leaves only finitely many options. The same procedure for ρ shows that the only matrices in $SL(2,\mathbb{Z})$ fixing ρ are of the form $\left(\begin{smallmatrix} a & b \\ -b & a-b \end{smallmatrix}\right)$, so $a^2 - ab + b^2 = 1$, which is equivalent to $(2a - b)^2 + 3b^2 = 4$. □

Given a nonconstant holomorphic map $f : X \to Y$ between two compact connected Riemann surfaces, it is natural to wonder if there is some relation between the genera of X and Y given in terms of data associated to f. In fact such a relation does exist: it was first discovered by Riemann and used by him without proof. Proofs were given later by Hurwitz [Hur91; Hur93] and Zeuthen [Zeu71].

Any such f is surjective by the open mapping theorem. By assumption, f is non-constant, so $f^{-1}(y)$ is discrete for all $y \in Y$. Since X is compact, $f^{-1}(y)$ must be finite.

If $f(x) = y$, then in local coordinates centred at x and y, f is represented by a holomorphic function of the form $\tilde{f}(z) = z^n + \text{H.O.T.}(z)$ for some integer $n \geq 1$. Set $n = \nu_x(f)$: this number is called the ramification degree. By the preceding paragraph, the quantity $\sum_{x\in f^{-1}(y)} \nu(x)$ is finite. Applying techniques of elementary complex analysis in a chart around an arbitrary $y \in Y$ reveals that $\sum_{x\in f^{-1}(y)} \nu_x(f)$ is locally constant, and must therefore be constant since Y is connected. We call this constant the degree of f, and write it as $\deg(f)$.

It is also easy to see that the set S of points of X for which $\nu_x \geq 2$ is finite, since S is must be closed if f is nonconstant, and X is compact.

Theorem 2.1.4.9 (Riemann–Hurwitz)**.** *Let $f : X \to Y$ be a nonconstant holomorphic map of compact connected Riemann surfaces. Let $g(X)$ and $g(Y)$ stand for the genera of X and Y respectively. Then with the notation introduced above,*

$$2g(X) - 2 = \deg(f)(2g(Y) - 2) + \sum_{x\in X}(\nu_x(f) - 1).$$

Proof. We sketch a topological proof, although it is possible to deduce the result from the Riemann–Roch theorem.

First, triangulate Y so that each ramification point is a vertex. Pull back the triangulation to X, and compare the counts of vertices, faces and edges:

$$V_X = \deg(f)V_Y - \sum_{x\in X}(\nu_x(f) - 1),$$
$$E_X = \deg(f)E_Y,$$
$$F_X = \deg(f)F_Y.$$

The Riemann–Hurwitz formula emerges upon using 2.9:

$$2 - 2g_X = V_X - E_X + F_X = \deg(f)(2 - 2g_Y) - \sum_{x\in X}(\nu_x(f) - 1).$$ □

We can now express the genus g of a modular curve $X(\Gamma)$ in terms of its elliptic points and cusps.

Corollary 2.1.4.10. *Let Γ be a congruence subgroup of $SL(2,\mathbb{Z})$, let g stand for the genus of $X(\Gamma)$, and $\deg(f)$ stand for the degree of the natural projection $f : X(\Gamma) \to X(SL(2,\mathbb{Z}))$. Let E_2 and E_3 stand for the numbers of elliptic points of multiplicity 2 and 3 for Γ and let C stand for the total number of cusps of Γ. Then*

$$g = 1 + \frac{1}{12}[PSL(2,\mathbb{Z}) : \overline{\Gamma}] - \frac{1}{4}E_2 - \frac{1}{3}E_3 - \frac{1}{2}C.$$

Proof. We proceed by computing $\sum_{x\in f^{-1}(y)}(\nu_x(f) - 1)$, as y runs over the non-elliptic points, the elliptic points and the cusps.

If y is not an elliptic point, and $x \in f^{-1}(y)$, then f doesn't ramify at x, so $\nu_x(f) = 1$.

If $x \in f^{-1}([i])$, then x is either non-elliptic or it is elliptic of order 2. If x is non-elliptic, $\nu_x(f) = 2$, and if x is elliptic, $\nu_x(f) = 1$. By the definition of the degree of f,

$$\deg(f) = E_2 + 2\#\{x \in f^{-1}([i]) \mid x \text{ is not elliptic}\},$$

so $\#\{x \in f^{-1}([i]) \mid x \text{ is not elliptic}\} = \frac{1}{2}(\deg(f) - E_2)$.

Set $\rho = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$. If $x \in f^{-1}([\rho])$, then $\nu_x(f) = 1$ if x is elliptic and 3 if not. Using the definition of degree again,

$$\#\{x \in f^{-1}([i]) \mid x \text{ is not elliptic}\} = \frac{1}{3}(\deg(f) - E_3).$$

It follows that

$$\sum_{x\in f^{-1}([\rho])}(\nu_x(f) - 1) = \frac{2}{3}(\deg(f) - E_3).$$

Lastly, we deal with the cusps. But $SL(2,\mathbb{Z})$ has only one cusp, and the fibre of that cusp consists of all the cusps of Γ, so

$$\sum_{x\in f^{-1}([\infty])}(\nu_x(f) - 1) = \sum_{x\in f^{-1}([\infty])}\nu_x(f) - C = \deg(f) - C.$$ □

2.2 Jacobi's theta function

An important source of modular forms is the class of functions referred to as theta functions. This is an umbrella term, encompassing all functions whose q-expansion coefficients are related in some way to solutions counts of systems of quadrics over integers. The earliest appearance of theta functions is in Bernoulli's *Ars Conjectandi* [Ber13], a treatise on probability, but the first systematic investigation into their mysterious properties was Jacobi's *Fundamenta Nova* [Jac29]. In this section, we shall prove the fundamental transformation laws for some very basic theta functions, and investigate their behaviour under modular substitutions.

2.2.1 Jacobi's transformation law

In this section we present a *complete* proof of Jacobi's transformation law for the theta function.

Theorem 2.2.1.1. *For $z \in \mathscr{H}$, set*

$$\theta(z) = \sum_{n=-\infty}^{\infty} e^{2\pi i n^2 z},$$

where the branch cut for the square root is taken along the negative imaginary axis. Then

$$\theta\left(-\frac{1}{4z}\right) = \sqrt{-2iz}\,\theta(z). \tag{2.13}$$

Proof. By uniqueness of analytic continuation, it suffices to prove the following transformation law

$$\sum_{n=-\infty}^{\infty} e^{-\pi n^2/x} = \sqrt{x} \sum_{n=-\infty}^{\infty} e^{-\pi n^2 x}, \tag{2.14}$$

for positive real numbers x. We define a real-valued function Schwartz function of n by

$$f_x(n) = e^{-\pi n^2 z}$$

and calculate its Fourier transform, with respect to n, for a real variable y:

$$\begin{aligned}\hat{f}_x(y) &= \int_{\mathbb{R}} e^{-\pi n^2 x - 2\pi i n y} dy \\ &= e^{-\pi y^2/x} \int_{\mathbb{R}} e^{-\pi x (n + iy/x)^2} dn \\ &= x^{-1/2} e^{-\pi y^2/x} \int_{\mathbb{R}} e^{-\pi n^2} dn,\end{aligned}$$

where in order to arrive at the final line we have used Cauchy's integral theorem together with standard estimates on the integrand to shift the line of integration. But this last integral is well-known: it has the value 1. The formula 2.14 follows upon applying Poisson summation (Theorem 1.2.2.4). □

2.2.2 Jacobi's transformation law via special values of zeta functions

In this subsection, we will use the results of Subsection 1.3.5 on the special values of the Riemann and Mordell–Tornheim zeta functions to prove Jacobi's transformation formula (Theorem 2.2.1.1) *without Fourier analysis.*

We define the Laplace transform $\mathscr{L}f(s)$ of a continuous function $f(t)$ by

$$\mathscr{L}f(s) = \int_0^{\infty} f(t) e^{-st} dt,$$

when the integral converges, and we define the Heaviside step function by

$$H(t) = \begin{cases} 1 & t \geq 0 \\ 0 & t < 0. \end{cases}$$

Set

$$A = t - \frac{4t^{3/2}}{3\sqrt{\pi}} \qquad B = \sum_{n=1}^{\infty}(t-\pi n^2)H(t-\pi n^2)$$

$$C = \sum_{n=1}^{\infty}\frac{\sin(2n\sqrt{\pi t})}{2\pi^2 n^3} \qquad D = \sum_{n=1}^{\infty}\frac{\sqrt{t}\cos(2n\sqrt{\pi t})}{\pi^{3/2}n^2},$$

and define

$$F(t) = A + 2(B - C + D).$$

Clearly each series above is absolutely convergent, so F is continuous.

Proposition 2.2.2.1. *The Laplace transform of F is*

$$u^{-2}\left(\sum_{n\in\mathbb{Z}} e^{-\pi n^2 u} - \frac{1}{\sqrt{u}}\sum_{n\in\mathbb{Z}} e^{-\pi n^2/u}\right).$$

Proof. Since each of the sums converge absolutely, we may commute the Laplace transform with the limits where necessary, and the result follows from the easily proved formulae:

$$\int_0^\infty \left(t - \frac{4t^{3/2}}{3\sqrt{\pi}}\right)e^{-st}dt = s^{-2} - s^{-5/2},$$

$$\int_0^\infty (t-\pi n^2)H(t-\pi n^2)\, e^{-st}dt = s^{-2}e^{-\pi n^2 s},$$

$$\frac{1}{2\pi^2 n^3}\int_0^\infty \sin(2n\sqrt{\pi t})e^{-st}dt = \frac{1}{2\pi n^2}s^{-3/2}e^{-\pi n^2/s},$$

$$\frac{1}{\pi^{3/2}n^2}\int_0^\infty \sqrt{t}\cos(2n\sqrt{\pi t})e^{-st}dt = \frac{1}{2\pi n^2}(s^{-3/2} - 2\pi n^2 s^{-5/2})e^{-\pi n^2/s}.$$

□

Therefore, the functional equation 2.13 for Jacobi's theta function will be proved once we know that F is zero. Our strategy is to consider the L^2 norm of F:

$$\|F\|_2^2 = \int_0^\infty F^2(t)dt = \sum_{k=0}^{\infty}\int_{k^2\pi}^{(k+1)^2\pi} F^2(t)dt = 2\pi\sum_{k=0}^{\infty}\int_k^{k+1} sF^2(\pi s^2)ds.$$

Since F is continuous, the functional equation for the theta function follows from the next result.

Proposition 2.2.2.2. *With F as defined above,*

$$\int_k^{k+1} sF^2(\pi s^2)ds = 0$$

for all integers $k > 0$.

Proof. It will transpire that the integral in question is elementary, and we will express the resulting numbers in terms of special values of the Riemann zeta function and the Mordell–Tornheim zeta function. First, we write F^2 in terms of A, B, C and D:

$$F^2 = (A + 2(B - C + D))^2 = A^2 + 4(AB - AC + AD - 2BC + 2BD - 2CD + B^2 + C^2 + D^2). \tag{2.15}$$

Note that for $0 \leq k \leq s \leq k+1$, we have

$$B(\pi s^2) = \pi\sum_{n=1}^{k}(s^2 - n^2) = \pi\left(ks^2 - \frac{1}{6}k(k+1)(2k+1)\right).$$

Now we integrate separately each of the terms on the right hand side of 2.15 over $[k, k+1]$ for non-negative integers k (using the notation of Subsection 1.3.5 where necessary).

$$\begin{aligned}\int_k^{k+1} s(A(\pi s^2))^2 ds &= \pi^2 \int_k^{k+1} \left(s^5 - \frac{8}{3}s^6 + \frac{16}{9}s^7\right) ds \\ &= \pi^2 \left(\frac{16k^7}{9} + \frac{32k^6}{9} + \frac{49k^5}{9} + \frac{85k^4}{18} + \frac{22k^3}{9} + \frac{13k^2}{18} + \frac{k}{9} + \frac{1}{126}\right).\end{aligned}$$

$$\begin{aligned}\int_k^{k+1} sC(\pi s^2)D(\pi s^2)ds &= \frac{1}{2\pi^3}\int_k^{k+1} s^2 \left(\sum_{n=1}^{\infty}\frac{\sin(2\pi ns)}{n^3}\right)\left(\sum_{m=1}^{\infty}\frac{\cos(2\pi ms)}{m^2}\right) ds \\ &= \frac{1}{2\pi^3}\sum_{m,n=1}^{\infty}\frac{1}{n^3m^2}\int_k^{k+1} s^2 \sin(2\pi ns)\cos(2\pi ms)ds \\ &= \frac{1}{2\pi^3}\left(\sum_{\substack{m,n=1\\ m\neq n}}^{\infty}\frac{1}{n^3m^2}\left(\frac{2k+1}{4\pi}\left(\frac{1}{m-n}-\frac{1}{m+n}\right)\right) - \sum_{m=n=1}^{\infty}\frac{1}{n^3m^2}\left(\frac{2k+1}{8\pi n}\right)\right) \\ &= \frac{2k+1}{8\pi^4}\sum_{\substack{m,n=1\\ m\neq n}}^{\infty}\frac{1}{n^3m^2}\frac{2n}{m^2-n^2} - \frac{2k+1}{16\pi^4}\sum_{n=1}^{\infty}\frac{1}{n^6} \\ &= -\frac{2k+1}{16\pi^4}\zeta(6).\end{aligned}$$

$$\begin{aligned}\int_k^{k+1} sD^2(\pi s^2)ds &= \frac{1}{\pi^2}\int_k^{k+1} s^3 \left(\sum_{n=1}^{\infty}\frac{\cos(2\pi ns)}{n^2}\right)\left(\sum_{m=1}^{\infty}\frac{\cos(2\pi ms)}{m^2}\right) ds \\ &= \frac{1}{\pi^2}\sum_{m,n=1}^{\infty}\frac{1}{n^2m^2}\int_k^{k+1} s^3\sin(2\pi ms)\sin(2\pi ns)ds \\ &= \frac{2k+1}{\pi^2}\left(\sum_{\substack{m,n=1\\ m\neq n}}^{\infty}\frac{1}{n^2m^2}\left(\frac{3}{8\pi^2}\left(\frac{1}{(m-n)^2}+\frac{1}{(m+n)^2}\right)\right) + \sum_{m=n=1}^{\infty}\frac{1}{n^2m^2}\left(\frac{2k^2+2k+1}{8}+\frac{3}{32\pi^2n^2}\right)\right) \\ &= \frac{3(2k+1)}{8\pi^4}\sum_{\substack{m,n=1\\ m\neq n}}^{\infty}\frac{1}{n^2m^2}\left(\frac{1}{(m-n)^2}+\frac{1}{(m+n)^2}\right) + \frac{2k+1}{\pi^2}\left(\frac{2k^2+2k+1}{8}\zeta(4)+\frac{3\zeta(6)}{32\pi^2}\right) \\ &= \frac{3(2k+1)}{8\pi^4}\left(2\sum_{\substack{m,n=1\\ m>n}}^{\infty}\frac{1}{n^2m^2(m-n)^2} + \sum_{m,n=1}^{\infty}\frac{1}{n^2m^2(m+n)^2} - \frac{1}{4}\zeta(6)\right) + \frac{2k+1}{\pi^2}\left(\frac{2k^2+2k+1}{8}\zeta(4)+\frac{3\zeta(6)}{32\pi^2}\right) \\ &= \frac{9(2k+1)}{8\pi^4}\zeta_{MT}(2,2;2) + \frac{(2k+1)(2k^2+2k+1)}{8\pi^2}\zeta(4) \\ &= \frac{\zeta(4)}{2\pi^2}k^3 + \frac{3\zeta(4)}{4\pi^2}k^2 + \left(\frac{\zeta(4)}{2\pi^2}+\frac{9\zeta_{MT}(2,2;2)}{4\pi^4}\right)k + \frac{\zeta(4)}{8\pi^2} + \frac{9\zeta_{MT}(2,2;2)}{8\pi^4}.\end{aligned}$$

$$\begin{aligned}
\int_k^{k+1} sC^2(\pi s^2)ds &= \frac{1}{4\pi^4}\int_k^{k+1} s\left(\sum_{n=1}^{\infty}\frac{\sin(2\pi ns)}{n^3}\right)\left(\sum_{m=1}^{\infty}\frac{\sin(2\pi ms)}{m^3}\right)ds \\
&= \frac{1}{4\pi^4}\sum_{m,n=1}^{\infty}\frac{1}{n^3m^3}\int_k^{k+1} s\sin(2\pi ns)\sin(2\pi ms)ds \\
&= \frac{2k+1}{16\pi^4}\sum_{m=n=1}^{\infty}\frac{1}{n^3m^3} \\
&= \frac{2k+1}{16\pi^4}\zeta(6).
\end{aligned}$$

$$\begin{aligned}
\int_k^{k+1} sB^2(\pi s^2)ds &= \pi^2\int_k^{k+1} s\left(ks^2-\frac{k(k+1)(2k+1)}{6}\right)^2 ds \\
&= \frac{\pi^2}{72}\left(32k^7+64k^6+98k^5+85k^4+38k^3+7k^2\right).
\end{aligned}$$

$$\begin{aligned}
\int_k^{k+1} sB(\pi s^2)C(\pi s^2)ds &= \frac{1}{2\pi}\int_k^{k+1} s\left(ks^2-\frac{k(k+1)(2k+1)}{6}\right)\sum_{n=1}^{\infty}\frac{\sin(2n\pi s)}{n^3}ds \\
&= \frac{k}{2\pi}\sum_{n=1}^{\infty}\frac{1}{n^3}\int_k^{k+1} s^3\sin(2n\pi s)ds-\frac{k(k+1)(2k+1)}{12\pi}\sum_{n=1}^{\infty}\frac{1}{n^3}\int_k^{k+1} s\sin(2n\pi s)ds \\
&= \frac{k}{2\pi}\sum_{n=1}^{\infty}\left(\frac{3}{4\pi^3n^6}-\frac{3k^2+3k+1}{2\pi n^4}\right)+\frac{k(k+1)(2k+1)}{24\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^4} \\
&= \frac{3k}{8\pi^4}\zeta(6)-\frac{k(3k^2+3k+1)}{4\pi^2}\zeta(4)+\frac{k(k+1)(2k+1)}{24\pi^2}\zeta(4) \\
&= -\frac{2\zeta(4)}{3\pi^2}k^3-\frac{5\zeta(4)}{8\pi^2}k^2+\left(-\frac{5\zeta(4)}{24\pi^2}+\frac{3\zeta(6)}{8\pi^4}\right)k.
\end{aligned}$$

$$\begin{aligned}
\int_k^{k+1} sB(\pi s^2)D(\pi s^2)ds &= \int_k^{k+1} s^2\left(ks^2-\frac{k(k+1)(2k+1)}{6}\right)\sum_{n=1}^{\infty}\frac{\cos(2n\pi s)}{n^2}ds \\
&= k\sum_{n=1}^{\infty}\frac{1}{n^2}\int_k^{k+1} s^4\cos(2n\pi s)ds-\frac{k(k+1)(2k+1)}{6}\sum_{n=1}^{\infty}\frac{1}{n^2}\int_k^{k+1} s^2\cos(2\pi ns)ds \\
&= k\left(\frac{3k^2+3k+1}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^4}-\frac{3}{2\pi^4}\sum_{n=1}^{\infty}\frac{1}{n^6}\right)-\frac{k(k+1)(2k+1)}{12\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^4} \\
&= \frac{k(3k^2+3k+1)}{\pi^2}\zeta(4)-\frac{3k}{2\pi^4}\zeta(6)-\frac{k(k+1)(2k+1)}{12\pi^2}\zeta(4) \\
&= \frac{17\zeta(4)}{6\pi^2}k^3+\frac{11\zeta(4)}{4\pi^2}k^2+\left(\frac{11\zeta(4)}{12\pi^2}-\frac{3\zeta(6)}{2\pi^4}\right)k.
\end{aligned}$$

$$\int_k^{k+1} sA(\pi s^2)D(\pi s^2)ds = \int_k^{k+1} s\left(\pi s^2 - \frac{4\pi s^3}{3}\right)\left(\sum_{n=1}^{\infty}\frac{s\cos(2n\pi s)}{\pi n^2}\right)ds$$
$$= \sum_{n=1}^{\infty}\frac{1}{n^2}\int_k^{k+1} s^4\cos(2n\pi s)ds - \frac{4}{3}\sum_{n=1}^{\infty}\frac{1}{n^2}\int_k^{k+1} s^5\cos(2\pi ns)ds$$
$$= \frac{3k^2+3k+1}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^4} - \frac{3}{2\pi^4}\sum_{n=1}^{\infty}\frac{1}{n^6} - \frac{5(4k^3+6k^2+4k+1)}{3\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^4} + \frac{5(2k+1)}{\pi^4}\sum_{n=1}^{\infty}\frac{1}{n^6}$$
$$= \frac{3k^2+3k+1}{\pi^2}\zeta(4) - \frac{3}{2\pi^4}\zeta(6) - \frac{5(4k^3+6k^2+4k+1)}{3\pi^2}\zeta(4) + \frac{5(2k+1)}{\pi^4}\zeta(6)$$
$$= -\frac{20\zeta(4)}{3\pi^2}k^3 - \frac{7\zeta(4)}{\pi^2}k^2 + \left(-\frac{11\zeta(4)}{3\pi^2} + \frac{10\zeta(6)}{\pi^4}\right)k - \frac{2\zeta(4)}{3\pi^2} + \frac{7\zeta(6)}{2\pi^4}.$$

$$\int_k^{k+1} sA(\pi s^2)C(\pi s^2)ds = \int_k^{k+1} s\left(\pi s^2 - \frac{4\pi s^3}{3}\right)\left(\sum_{n=1}^{\infty}\frac{\sin(2n\pi s)}{2\pi^2 n^3}\right)ds$$
$$= \frac{1}{2\pi}\sum_{n=1}^{\infty}\frac{1}{n^3}\int_k^{k+1} s^3\sin(2\pi ns)ds - \frac{2}{3\pi}\sum_{n=1}^{\infty}\frac{1}{n^3}\int_k^{k+1} s^4\sin(2\pi ns)ds$$
$$\frac{1}{2\pi}\left(\frac{3}{4\pi^3}\sum_{n=1}^{\infty}\frac{1}{n^6} - \frac{3k^2+3k+1}{2\pi}\sum_{n=1}^{\infty}\frac{1}{n^4}\right) + \frac{2}{3\pi}\left(\frac{4k^3+6k^2+4k+1}{2\pi}\sum_{n=1}^{\infty}\frac{1}{n^4} - \frac{3(2k+1)}{2\pi^3}\sum_{n=1}^{\infty}\frac{1}{n^6}\right)$$
$$= \frac{3}{8\pi^4}\zeta(6) - \frac{3k^2+3k+1}{4\pi^2}\zeta(4) + \frac{2(4k^3+6k^2+4k+1)}{6\pi^2}\zeta(4) - \frac{2k+1}{\pi^4}\zeta(6)$$
$$= \frac{4\zeta(4)}{3\pi^2}k^3 + \frac{5\zeta(4)}{4\pi^2}k^2 + \left(\frac{7\zeta(4)}{12\pi^2} - \frac{2\zeta(6)}{\pi^4}\right)k + \frac{\zeta(4)}{12\pi^2} - \frac{5\zeta(6)}{8\pi^4}.$$

$$\int_k^{k+1} sA(\pi s^2)B(\pi s^2)ds = \pi^2 k\int_k^{k+1} s^3\left(1 - \frac{4\pi s}{3}\right)\left(s^2 - \frac{(k+1)(2k+1)}{6}\right)ds$$
$$= -\left(\frac{8\pi^2}{9}k^7 + \frac{16\pi^2}{9}k^6 + \frac{49\pi^2}{18}k^5 + \frac{85\pi^2}{36}k^4 + \frac{17\pi^2}{15}k^3 + \frac{97\pi^2}{360}k^2 + \frac{53\pi^2}{2520}k\right)$$

Collecting the calculations, it follows that

$$F = \left(\frac{\pi^2}{45} - \frac{\zeta(4)}{2\pi^2}\right)k^3 + \left(\frac{\pi^2}{30} - \frac{3\zeta(4)}{\pi^2}\right)k^2 + \left(\frac{17\pi^2}{630} - \frac{6\zeta(4)}{\pi^2} + \frac{69\zeta(6)}{2\pi^4} + \frac{9\zeta_{MT}(2,2;2)}{\pi^4}\right)k$$
$$+ \frac{\pi^2}{126} - \frac{5\zeta(4)}{2\pi^2} + \frac{69\zeta(6)}{4\pi^4} + \frac{9\zeta_{MT}(2,2;2)}{2\pi^4},$$

where the coefficients of k^7, k^6, k^5 and k^4 have already vanished. We observe that the vanishing of the coefficients of k^3 and k^2 is equivalent to the evaluation $\zeta(4) = \frac{\pi^4}{90}$ (see 1.17 from Subsection 1.3.5), and the vanishing of the coefficient of k and the constant coefficient follows from the evaluation of $\zeta_{MT}(2,2;2)$ (see 1.21) together with 1.17 and 1.18. □

The author's attempts to generalise the relationship above to a result linking functional equations of L-functions of modular forms of higher weight to identities between special values of those L-functions have not met with success. If the Dirichlet coefficients of the L-function in question do not vanish off the perfect squares, then the integrals over $[k, k+1]$ become much more difficult to evaluate and are not independent of k. Consequently, the expression occurring in lieu of F is a finite sum of special values of functions which are not related to the original L-function in any obvious way.

2.2.3 Modularity of theta functions

We proved in Theorem 2.2.1.1 that Jacobi's theta function satisfies a functional equation

$$\theta(-1/4z) = \sqrt{-2iz}\theta(z),$$

but we have remarked many times that theta functions are examples of modular forms. Therefore, we need to know how $\theta(z)$ transforms under the action of an arithmetic subgroup of $SL(2,\mathbb{R})$. Our proof of the next result follows that of Eichler [Eic77] (who in turn claims to follow Hermite). A longer proof is given by Koblitz [Kob84, Chapter III, Section 4], who treats modular forms of *half-integral weight* (of which Jacobi's theta function is the simplest example) in detail. A systematic study of modular forms of half-integral weight was begun by Shimura [Shi72].

Proposition 2.2.3.1. *For all $z \in \mathscr{H}$,*

$$\theta(\gamma z) = j(\gamma, z)\theta(z),$$

for all $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(4)$, where

$$\Gamma_0(4) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{Z}) \text{ such that } c = 0 \bmod 4 \right\}$$

and

$$j(\gamma, z) = \begin{cases} \left(\dfrac{c}{d}\right)\epsilon_d^{-1}(cz+d)^{1/2} \text{ where } \left(\dfrac{c}{d}\right) \text{ and } \epsilon_d \text{ are as usual and } |\arg{(cz+d)^{1/2}}| < \dfrac{\pi}{2}, & d > 0, \\ j(-\gamma, z) & d < 0. \end{cases}$$

Recall that we defined the Legendre symbol in Subsection 1.3.2 for odd prime denominators; here, we extend it to all odd denominators by multiplicativity. The epsilon factor, *which will feature prominently in Chapter 4, is defined by*

$$\epsilon_d = \begin{cases} 1 & d = 1 \bmod 4, \\ i & d = 3 \bmod 4. \end{cases}$$

The reader should note that during the proof of Proposition 2.2.3.1, we will require a few results which are proved in Chapter 4.

Proof. We proceed indirectly in order to avoid having to evaluate difficult Gauss sums. Instead of investigating the behaviour of θ under a matrix in $\Gamma_0(4)$, we will begin by considering the action on θ of a matrix of the form $\left(\begin{smallmatrix} b & -a \\ d & -c \end{smallmatrix}\right)$, where $b = 0 \bmod 2$ and $c = 0 \bmod 2$. We also define

$$\vartheta(z) = \theta(z/2) = \sum_{n\in\mathbb{Z}} e^{\pi i n z}.$$

Then we have

$$\begin{aligned} \vartheta\left(\frac{bz-a}{dz-c}\right) &= \vartheta\left(\frac{b}{d} - \frac{1}{d(dz-c)}\right) \\ &= \sum_{n=0}^{|d|-1} e^{\pi i n^2 b/d} \sum_{k=-\infty}^{\infty} e^{-\pi i (n/d+k)^2} d/(dz-c) \\ &= (id)^{-1/2}(dz-c)^{1/2} \sum_{n=0}^{|d|-1} e^{\pi i n^2 b/d} \sum_{k=-\infty} e^{\pi i (2nk/d + k^2(z-c/d))}, \end{aligned} \tag{2.16}$$

where we have used Poisson summation to arrive at the third line. Upon interchanging the order of summation, the sum over n resolves to an expression known as a Gauss sum:

$$\sum_{n=0}^{|d|-1} \exp\left(\frac{\pi i(n^2 b + 2nk)}{d}\right) = \sum_{n=0}^{|d|-1} \exp\left(\frac{2\pi i(b'n^2 + kn)}{d}\right),$$

where we have rewritten the sum to take advantage of the fact that $b = 2b'$ is even. Since $(b, d) = 1$, there exists some b'' so that $b'b'' = 1 \bmod d$, so we may make the substitution $n \mapsto b''n$ in the sum:

$$\sum_{n=0}^{|d|-1} \exp\left(\frac{\pi i(n^2 b + 2nk)}{d}\right) = \sum_{n=0}^{|d|-1} \exp\left(\frac{2\pi i b''(n^2 + kn)}{d}\right)$$

Then, since d is odd, there is some τ so that $2\tau = 1 \bmod d$, and we let $n \mapsto n + k\tau$:

$$\sum_{n=0}^{|d|-1} \exp\left(\frac{\pi i(n^2 b + 2nk)}{d}\right) = \exp\left(-\frac{2\pi i b''\tau^2 k^2}{d}\right) \sum_{n=0}^{|d|-1} \exp\left(\frac{2\pi i b'' n^2}{d}\right).$$

The sum remaining on the right hand side of the preceding display may be evaluated Proposition 4.3.2.1:

$$\sum_{n=0}^{|d|-1} \exp\left(\frac{2\pi i b'' n^2}{d}\right) = \epsilon_{|d|}\left(\frac{\operatorname{sgn}(d) b''}{|d|}\right)\sqrt{|d|}.$$

Observe that since $-bc + ad = 1$, $-2b'c = 1 \bmod d$. Therefore $\tau = -b'c \bmod d$, so

$$\exp\left(-\frac{2\pi i b''\tau^2 k^2}{d}\right) = \exp\left(-\frac{2\pi i b''(b')^2 c^2 k^2}{d}\right) = \exp\left(-\frac{2\pi i b' c^2 k^2}{d}\right) = \exp\left(-\frac{\pi i c k^2}{d}\right).$$

Substituting our results into 2.16, we find that

$$\vartheta\left(\frac{bz - a}{dz - c}\right) = (id)^{-1/2}(dz - c)^{1/2}\epsilon_{|d|}\left(\frac{\operatorname{sgn}(d) b''}{|d|}\right)\sqrt{|d|}\vartheta(z),$$

and upon making the substitution $z \mapsto -1/z$ and applying analytic continuation to 2.14,

$$\begin{aligned}\vartheta\left(\frac{az + b}{cz + d}\right) &= (id)^{-1/2}(-d/z - c)^{1/2}\epsilon_{|d|}\left(\frac{\operatorname{sgn}(d) b''}{|d|}\right)\sqrt{|d|}\vartheta\left(-\frac{1}{z}\right) \\ &= d^{-1/2}|d|^{1/2}(cz + d)^{1/2}\epsilon_{|d|}\left(\frac{\operatorname{sgn}(d) b''}{|d|}\right)\vartheta(z).\end{aligned}$$

Lastly, note that $b'b'' = 1 \bmod d$ implies $\left(\frac{b''}{d}\right) = \left(\frac{b'}{d}\right)$, and $-2b'c = 1 \bmod d$ implies that $\left(\frac{b'}{d}\right) = \left(\frac{-2c}{d}\right)$. Then

$$\begin{aligned}\vartheta\left(\frac{az + b}{cz + d}\right) &= \begin{cases} (cz + d)^{1/2}\epsilon_d\left(\frac{2c}{d}\right)\vartheta(z) & d > 0 \\ i(cz + d)^{1/2}\epsilon_{-d}\left(\frac{2c}{-d}\right)\vartheta(z) & d < 0 \end{cases} \\ &= j\left(\begin{pmatrix} a & b \\ 2c & d \end{pmatrix}, z/2\right)\vartheta(z),\end{aligned}$$

and upon replacing $\vartheta(z)$ with $\theta(z/2)$, the claim is proved. □

When Jacobi's theta function is raised to an even power, the Legendre symbol no longer appears in the description of the behaviour of such functions under the action of elements of $SL(2, \mathbb{Z})$, and it is much more straightforward to find congruence subgroups under which such functions are modular forms. The case which interests us most, particularly in Chapter 3, is the fourth power of θ, on account of the following remarkable fact:

Observation 2.2.3.2. *The nth Fourier coefficient of θ^4, $r_4(n)$, is equal to the number of ways in which n may be represented as a sum of four squares of integers. More precisely,*

$$r_4(n) = \#\{(x, y, z, w) \in \mathbb{Z}^4 \mid x^2 + y^2 + z^2 + w^2 = n\}.$$

The proof of the Observation is clear from the definition

$$\theta^4(z) = \left(\sum_{n \in \mathbb{Z}} e^{2\pi i n^2 z}\right)^4. \tag{2.17}$$

The next result, together with Observation 2.2.3.2, provides to vital link between the theory of modular forms and the study of representations of squares.

Proposition 2.2.3.3. *Let θ denote Jacobi's theta function. Then θ^4 is a modular form of weight 2 for the congruence subgroup $\Gamma_0(4)$.*

Proof. By Jacobi's transformation law (Theorem 2.2.1.1), θ^4 is invariant under the action of $\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$, and it satisfies

$$\theta^4\left(-\frac{1}{4z}\right) = -4z^2\theta(z). \tag{2.18}$$

The element of $GL(2,\mathbb{Q})$ effecting the transformation in 2.18 is $\left(\begin{smallmatrix}0&-1\\4&0\end{smallmatrix}\right)$, which doesn't have determinant equal to unity. However, the product

$$\begin{pmatrix}0&-1\\4&0\end{pmatrix}^{-1}\begin{pmatrix}1&1\\0&1\end{pmatrix}\begin{pmatrix}0&-1\\4&0\end{pmatrix} = \begin{pmatrix}1&0\\4&1\end{pmatrix}$$

is comprised entirely of matrices such that the transformation of θ^4 under their action is known; and the product lies in the congruence subgroup $\Gamma_0(4)$. It follows that

$$\theta^4(\gamma z) = j(\gamma,z)^2\theta^4(z) \text{ for } \gamma \in \left\langle \pm\begin{pmatrix}1&1\\0&1\end{pmatrix}, \pm\begin{pmatrix}1&0\\4&1\end{pmatrix}\right\rangle.$$

But by Proposition 2.1.1.2, $\Gamma_0(4)$ is exactly the group generated by $\pm\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$ and $\pm\left(\begin{smallmatrix}1&0\\4&1\end{smallmatrix}\right)$. The last condition which must be satisfied in order for θ^4 to be a modular form is that it be holomorphic at the cusps of $\Gamma_0(4)$. According to the commentary following Definition 2.1.1.3, we need only check that there exist positive constants C and k such that the absolute value of the nth Fourier coefficient of θ^4 is bounded by Cn^k. We observed above that the nth Fourier coefficient of θ^4, denoted $r_4(n)$, is equal to the number of ways in which n may be written as a sum of four squares of integers. Ignoring changes of sign and orderings, an integer n can't be written as a sum of four squares in more than n^2 ways, since there are no more than $\sqrt{n}$ choices for each square appearing in the sum. It follows that $r_4(n)$ is bounded by a multiple of n^2, as desired. □

2.3 Modular forms and Jacobi's theorem on sums of four squares

We have seen that the fourth power of Jacobi's theta function satisfies a functional equation:

$$\theta^4\left(-\frac{1}{4z}\right) = -4z^2\theta^4(z),$$

with the consequence that θ^4 is a modular form of weight 2 for the congruence subgroup $\Gamma_0(4)$; indeed, for any $\gamma \in \Gamma_0(4)$,

$$\theta^4(\gamma z) = j(\gamma,z)^2\theta^4(z).$$

We have investigated the spaces of modular forms on congruence subgroups in Section 2.1 in enough detail to conclude that $\mathscr{M}_2(\Gamma_0(4))$ is finite-dimensional, with the exact dimension given by the Riemann–Roch theorem. In Subsection 2.3.1, we deduce that

$$\dim \mathscr{M}_2(\Gamma_0(4)) = 2.$$

Therefore, if we could only find a basis of $\mathscr{M}_2(\Gamma_0(4))$ with explicitly given Fourier coefficients, we should be able to deduce Jacobi's theorem (see the introduction to this chapter). There enough evidence already that this is possible: note that $\dim \mathscr{S}_2(\Gamma_0(4)) = 0$, so we expect that $\mathscr{M}_2(\Gamma_0(4))$ consists entirely of sums of Eisenstein series, which, in principle, have explicitly determinable Fourier coefficients.

In practice, things are more complicated: the Eisenstein series G_k for $SL(2,\mathbb{Z})$ doesn't converge uniformly for $k = 2$, so the argument which show that G_k is a modular form for $k > 2$ doesn't apply here. In Subsection 2.3.2, we carry out the tedious calculation that determines exactly how G_2 transforms under the action of $SL(2,\mathbb{Z})$ and from it we fashion two linearly independent Eisenstein series for $\mathscr{M}_2(\Gamma_0(4))$. Jacobi's theorem emerges upon a comparison of Fourier coefficients.

2.3.1 The dimension of $\mathscr{M}_2(\Gamma_0(4))$

Recall the formulae from Proposition 2.1.4.5, specialised to the case $k = 2$:

$$\dim \mathscr{M}_2(\Gamma) = \sum_i \lfloor (1 - e_i^{-1}) \rfloor + C + g - 1, \tag{2.19}$$

$$\dim \mathscr{S}_2(\Gamma) = g, \tag{2.20}$$

where the sum runs over the elliptic points of Γ.

We will now determine the invariants of the action of $\Gamma_0(4)$ required to compute the dimension of $\mathscr{M}_2(\Gamma_0(4))$.

Proposition 2.3.1.1. *The genus of $X(\Gamma_0(4))$ is 0. The action of $\Gamma_0(4)$ on $\mathscr{H}$ produces no elliptic points, and three inequivalent cusps.*

Proof. First, we show that $\Gamma_0(4)$ acts freely on $\mathscr{H}$. By Lemma 2.1.4.8, the distinct points of $\mathscr{H}$ fixed by $SL(2,\mathbb{Z})$, up to the action of $SL(2,\mathbb{Z})$, are

$$i = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} i \text{ and } -\frac{1}{2} + \frac{\sqrt{3}}{2} i = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} \left(-\frac{1}{2} + \frac{\sqrt{3}}{2} i \right).$$

Therefore, the matrices in $SL(2,\mathbb{Z})$ which fix i and $-\frac{1}{2} + \frac{\sqrt{3}}{2} i$ are all of the form

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ac+bd & -(a^2+b^2) \\ c^2+d^2 & -(ac+bd) \end{pmatrix}$$

and

$$B = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ac+ad+bd & -(a^2+ab+b^2) \\ c^2+cd+d^2 & -(ac+bc+bd) \end{pmatrix}.$$

But if $c^2 + d^2$ or $c^2 + cd + d^2$ is congruent to 0 modulo 4, c and d must both be even, which contradicts the requirement that $a - bc = 1$.

Now we turn to the problem of determining the distinct classes of cusps. We claim that there are three cusps, represented by 0, $\frac{1}{2}$ and ∞. To prove this, we will find the orbit of each of these points, and then it will be clear than the orbits are disjoint and that their union is $\mathbb{Q} \cup \{\infty\}$.

1. The orbit of ∞:

   $$\text{Claim:} \quad \Gamma_0(4) \cdot \infty = \left\{ \frac{p}{q} \text{ such that } (p,q) = 1, 4 \mid q \right\}.$$

   Since $(p,q) = 1$, there exist $a, b \in \mathbb{Z}$ such that $ap - bq = 1$. Then the matrix $\left(\begin{smallmatrix} p & b \\ q & a \end{smallmatrix}\right) \in \Gamma_0(4)$, and

   $$\begin{pmatrix} p & b \\ q & a \end{pmatrix} \cdot \infty = \frac{p}{q}.$$

2. The orbit of 0:

   $$\text{Claim:} \quad \Gamma_0(4) \cdot 0 = \left\{ \frac{p}{q} \text{ such that } (p,q) = 1, 2 \nmid q \right\}$$

   Since q is odd, $(4p, q) = 1$, so there exist $a, b \in \mathbb{Z}$ such that $ap - 4bq = 1$. Then the matrix $\left(\begin{smallmatrix} a & p \\ 4b & q \end{smallmatrix}\right) \in \Gamma_0(4)$, and

   $$\begin{pmatrix} a & p \\ 4b & q \end{pmatrix} \cdot \infty = \frac{p}{q}.$$

3. The orbit of $\frac{1}{2}$:

   $$\text{Claim:} \quad \Gamma_0(4) \cdot \frac{1}{2} = \left\{ \frac{p}{q} \text{ such that } (p,q) = 1, 2 \mid q, 4 \nmid q \right\}.$$

   Since $(4p, q) = 2$, there exist $a, b \in \mathbb{Z}$ such that $4ap - bq = 2$. Set $q = 2r$ and note that b is odd, since $4 \mid 2$ otherwise. We have a candidate matrix in $\Gamma_0(4)$:

   $$\begin{pmatrix} b & -\frac{1}{2}(b+p) \\ 4a & -(2a+r) \end{pmatrix}, \quad \det \begin{pmatrix} b & -\frac{1}{2}(b+p) \\ 4a & -(2a+r) \end{pmatrix} = -2ab - br + 2ab + 2ap = (4ap - 2br)/2 = 1,$$

and

$$\begin{pmatrix} b & -\frac{1}{2}(b+p) \\ 4a & -(2a+r) \end{pmatrix} \cdot \frac{1}{2} = \frac{-p}{-2r} = \frac{p}{q}.$$

Finally, we compute the genus. We have already computed that the genus of $X(SL(2,\mathbb{Z}))$ is zero in 2.1.4.6, and one may check that the projection $f : X(\Gamma) \to X(SL(2,\mathbb{Z})$ is a nonconstant holomorphic map for any congruence subgroup Γ. If we set $\Gamma = \Gamma_0(4)$ and apply Corollary 2.1.4.10, using g to denote the genus of $X(\Gamma_0(4))$, we have

$$g = 1 + \frac{1}{12}[PSL(2,\mathbb{Z}) : \overline{\Gamma_0(4)}] - \frac{3}{2},$$

since we have just proved that $\Gamma_0(4)$ has no elliptic points and exactly three cusps. Setting $N = 4$ in our calculation of the index of $\Gamma_0(N)$ in $SL(2,\mathbb{Z})$ in Lemma 2.1.1.1, we obtain

$$[PSL(2,\mathbb{Z}) : \overline{\Gamma_0(4)}] = [SL(2,\mathbb{Z}) : \Gamma_0(4)] = 6,$$

and it follows that the genus is zero. □

Upon substituting the results of Proposition 2.3.1.1 into 2.19 and 2.20, we obtain the dimension results.

Corollary 2.3.1.2. *We have* $\mathscr{M}_2(\Gamma_0(4)) = 2$ *and* $\mathscr{S}_2(\Gamma_0(4)) = 0$.

2.3.2 Eisenstein series for $\mathscr{M}_2(\Gamma_0(4))$

We must now find a basis for $\mathscr{M}_2(\Gamma_0(4))$. We know from the previous subsection that there are no cusp forms to worry about, which is fortunate as questions of positivity of the Fourier coefficients of cusp forms is often extremely involved (especially for spaces of low weights [Ser81]). However, the road to actually finding the two Eisenstein series and computing their Fourier coefficients is not as straight as one might expect. In this subsection, we present the "usual" proof, making use of the observation that G_2 transforms under the action of $SL(2,\mathbb{Z})$ as a quasi-modular form, rather than a modular form, then averaging out in two different ways to obtain two linearly independent modular forms for $\mathscr{M}_2(\Gamma_0(4))$. In Chapter 3, we present an alternative proof, in which we will prove that G_2 is a quasi-modular form using the properties of a certain differential equation.

We begin by determining how the weight 2 Eisenstein series

$$G_2(\tau) = \sum_{c\in\mathbb{Z}} \sum_{\substack{d\in\mathbb{Z} \\ d\neq 0 \text{ if } c\neq 0}} (c\tau + d)^{-2} \tag{2.21}$$

transforms under the action of $SL(2,\mathbb{Z})$. A tedious calculation reveals that G_2 transforms as a quasi-modular form.

Proposition 2.3.2.1. *For all* $\gamma \in SL(2,\mathbb{Z})$,

$$G_2\big|[\gamma]_2 = G_2(\tau) - \frac{2\pi i c}{c\tau + d} \tag{2.22}$$

Since G_2 has a Fourier expansion, the claim is true for all matrices of the form T^n, where $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$.We proceed by first proving the proposition for $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$.

Lemma 2.3.2.2. *For all* $\tau \in \mathscr{H}$,

$$\tau^{-2} G_2\left(-\frac{1}{\tau}\right) = G_2(\tau) - \frac{2\pi i}{\tau}.$$

Proof. We begin with the left hand side of 2.22.

$$\begin{aligned} \tau^{-2} G_2\left(-\frac{1}{\tau}\right) &= \sum_{c\in\mathbb{Z}} \sum_{\substack{d\in\mathbb{Z} \\ d\neq 0 \text{ if } c\neq 0}} \frac{1}{(-c + d\tau)^2} \\ &= \sum_{d\in\mathbb{Z}} \sum_{\substack{c\in\mathbb{Z} \\ c\neq 0 \text{ if } d\neq 0}} \frac{1}{(c\tau - d)^2} \\ &= 2\zeta(2) + \sum_{d\in\mathbb{Z}} \sum_{\substack{c\in\mathbb{Z} \\ c\neq 0}} \frac{1}{(c\tau + d)^2} \end{aligned}$$

On the other hand,

$$G_2(\tau) = 2\zeta(2) + \sum_{\substack{c\in\mathbb{Z}\\ c\neq 0}}\sum_{d\in\mathbb{Z}} \frac{1}{(c\tau+d)^2}.$$

Flourishing a rabbit, we subtract the telescoping series

$$\sum_{\substack{c\in\mathbb{Z}\\ c\neq 0}}\sum_{d\in\mathbb{Z}} \frac{1}{(c\tau+d)(c\tau+d+1)} = \sum_{\substack{c\in\mathbb{Z}\\ c\neq 0}}\sum_{d\in\mathbb{Z}} \left(\frac{1}{c\tau+d} - \frac{1}{c\tau+d+1}\right) = 0$$

from $G_2(\tau)$:

$$G_2(\tau) - \sum_{\substack{c\in\mathbb{Z}\\ c\neq 0}}\sum_{d\in\mathbb{Z}} \frac{1}{(c\tau+d)(c\tau+d+1)} = 2\zeta(2) + \sum_{\substack{c\in\mathbb{Z}\\ c\neq 0}}\sum_{d\in\mathbb{Z}} \frac{1}{(c\tau+d)^2(c\tau+d+1)}.$$

This last sum is absolutely convergent, so we may interchange the order of summation. Upon subtracting the expression at 2.21, we have

$$\begin{aligned}
G_2(\tau) - \tau^{-2}G_2\left(-\frac{1}{\tau}\right) &= \sum_{d\in\mathbb{Z}}\sum_{\substack{c\in\mathbb{Z}\\ c\neq 0}} \frac{1}{(c\tau+d)^2(c\tau+d+1)} - \sum_{d\in\mathbb{Z}}\sum_{\substack{c\in\mathbb{Z}\\ c\neq 0}} \frac{1}{(c\tau+d)^2} \\
&= -\sum_{d\in\mathbb{Z}}\sum_{\substack{c\in\mathbb{Z}\\ c\neq 0}} \frac{1}{(c\tau+d)(c\tau+d+1)} \\
&= -\lim_{N\to\infty} \sum_{d=-N}^{N-1}\sum_{\substack{c\in\mathbb{Z}\\ c\neq 0}} \left(\frac{1}{c\tau+d} - \frac{1}{c\tau+d+1}\right) \\
&= -\lim_{N\to\infty} \sum_{\substack{c\in\mathbb{Z}\\ c\neq 0}}\sum_{d=-N}^{N-1} \left(\frac{1}{c\tau+d} - \frac{1}{c\tau+d+1}\right) \\
&= -\lim_{N\to\infty} \sum_{\substack{c\in\mathbb{Z}\\ c\neq 0}} \left(\frac{1}{c\tau-N} - \frac{1}{c\tau+N}\right).
\end{aligned}$$

The two sums on the final line may be converted into sum over positive c:

$$\sum_{\substack{c\in\mathbb{Z}\\ c\neq 0}} \left(\frac{1}{c\tau-N} - \frac{1}{c\tau+N}\right) = -2\left(\sum_{c=1}^{\infty} \frac{1}{N-c\tau} + \sum_{c=1}^{\infty} \frac{1}{N+c\tau}\right)$$

Upon using the partial fraction decomposition of the cotangent (which is obtained differentiating the logarithm of the infinite product for the sine function [MM97, Chapter 3 Section 5]),

$$\pi\cot(\pi z) = \frac{1}{z}\sum_{n=1}^{\infty}\left(\frac{1}{z-n} + \frac{1}{z+n}\right),$$

we have the claimed equation:

$$G_2(\tau) - \tau^{-2}G_2\left(-\frac{1}{\tau}\right) = \frac{2}{\tau}\lim_{N\to\infty}\left(\pi\cot\left(\frac{\pi N}{\tau}\right) - \frac{\tau}{N}\right) = \frac{2\pi i}{\tau}. \qquad \square$$

Proof of Proposition 2.3.2.1. We have now verified the proposition for the generators S and T (and T^n for all $n \in \mathbb{Z}$) of $SL(2,\mathbb{Z})$. We just need to check that this implies 2.22 for all $\gamma \in SL(2,\mathbb{Z})$. Since any matrix in $SL(2,\mathbb{Z})$ can be written as a word in S and T^n, it suffices to check that if the proposition is true for γ_1 and γ_2, then it is true for $\gamma_1\gamma_2$.

Indeed, set $\gamma_1 = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ and $\gamma_2 = \left(\begin{smallmatrix} x & y \\ z & w \end{smallmatrix}\right)$. Then

$$\begin{aligned} G_2(\gamma_1\gamma_2\tau) &= (c\gamma_2\tau + d)^2 G_2(\gamma_2\tau) - 2\pi i c(c\gamma_2\tau + d) \\ &= \frac{1}{(z\tau+w)^2}(c(x\tau+y)+d(z\tau+w))^2\left((z\tau+w)^2 G_2(\tau) - 2\pi i z(z\tau+w)\right) \\ &\quad - \frac{2\pi i c}{z\tau+w}(c(x\tau+y)+d(z\tau+w)) \\ &= (c(x\tau+y)+d(z\tau+w))^2 G_2(\tau) - \frac{2\pi i z}{z\tau+w}(c(x\tau+y)+d(z\tau+w))(z(c(x\tau+y)+d(z\tau+w))+c). \end{aligned}$$

The quantity $c(x\tau+y)+d(z\tau+w)$ is $j(\gamma_1\gamma_2,\tau)^2$, and the lower left entry of $\gamma_1\gamma_2$ is $cx+dz$, so we only need to show that

$$\frac{z(c(x\tau+y)+d(z\tau+w))+c}{z\tau+w} = cx+dz.$$

But this is immediate upon rearranging and using the determinant condition $xw - zy = 1$. □

We need to produce from G_2 at least two modular forms of weight 2 for $\Gamma_0(4)$. This is afforded by the following lemma.

Lemma 2.3.2.3. *For every $N \geq 1$, the function*

$$G_{2,N}(\tau) = G_2(\tau) - NG_2(N\tau)$$

is a modular form for $\mathscr{M}_2(\Gamma_0(N))$.

Proof. If $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(4)$, then $ad - bc = 1$ and $c = Nc'$ for some integer c'. One easily checks that $N\gamma\tau = \gamma' N\tau$, where $\gamma = \left(\begin{smallmatrix} a & bN \\ c' & d \end{smallmatrix}\right) \in SL(2,\mathbb{Z})$. Therefore,

$$\begin{aligned} G_2(\gamma\tau) - NG_2(N\gamma\tau) &= (c\tau+d)^2\left(G_2(\tau) - \frac{2\pi i c}{c\tau+d}\right) - N(Nc'\tau+d)^2\left(G_2(N\tau) - \frac{2\pi i c'}{Nc'\tau+d}\right) \\ &= (c\tau+d)^2\left(G_2(\tau) - NG_2(N\tau)\right) \end{aligned}$$ □

Observe that $\Gamma_0(4) \subset \Gamma_0(2)$, so $\mathscr{M}(\Gamma_0(2)) \subset \mathscr{M}(\Gamma_0(4))$. Consequently, the modular forms $G_{2,2}(\tau)$ and $G_{2,4}(\tau)$ are in $\mathscr{M}_2(\Gamma_0(4))$.

By Proposition 2.1.2.4, the Fourier expansion of G_2 is

$$G_2 = 2\zeta(2) - 8\pi^2 \sum_{n=1}^{\infty} \sigma(n) e^{2\pi i n\tau}.$$

It follows that

$$G_{2,2}(\tau) = -\frac{\pi^2}{3}\left(1 + 24\sum_{n=1}^{\infty}\left(\sum_{\substack{d|n \\ d \text{ odd},\ d>0}} d\right)e^{2\pi i n\tau}\right) = -\frac{\pi^2}{3}\left(1 + 24e^{2\pi i\tau} + \dots\right)$$

and

$$G_{2,4}(\tau) = -\pi^2\left(1 + 8\sum_{n=1}^{\infty}\left(\sum_{\substack{d|n \\ 4\nmid d, d>0}} d\right)e^{2\pi i n\tau}\right) = -\pi^2\left(1 + 8e^{2\pi i\tau} + \dots\right)$$

The modular forms $G_{2,2}(\tau)$ and $G_{2,4}(\tau)$ are obviously linearly independent, so by Corollary 2.3.1.2, they span $\mathscr{M}_2(\Gamma_0(4))$. Upon comparison with the Fourier coefficients of $\theta^4 = 1 + 8e^{2\pi i\tau} + \dots$, we obtain Jacobi's version of Lagrange's theorem, first stated near the beginning of this chapter:

Theorem 2.3.2.4 (Jacobi, 1829 [Jac29])**.** *For $n \geq 1$, let $r_4(n)$ count the number of ways that n may be represented as a sum of four squares. Then*

$$r_4(n) = 8\sum_{\substack{d|n \\ 4\nmid d, d>0}} d.$$

Clearly, $r_4(n)$ is always positive since 1 divides n and $4 \nmid 1$.

2.3.3 Quasimodularity via L-functions

In the last subsection, we proved, with tears, that G_2 transforms as a "quasimodular" form under the action of $SL(2,\mathbb{Z})$. Needless to say, we wish to find simpler, more conceptual proofs of the quasimodularity of G_2, and in Subsection 3.3.2, we present a proof of the modularity of G_2 based on the projective invariance of a certain differential operator.

A rather more common approach [Kob84, Chapter III, near the end of Section 2] is to note that G_2 is, up to an irrelevant multiple, the logarithmic derivative of Dedekind's eta function $\eta(z)$, defined as

$$\eta(z) = e^{2\pi i z/24} \prod_{n=1}^{\infty} \left(1 - e^{2\pi i n z}\right).$$

One easily shows that the quasimodularity of G_2 is equivalent to the following functional equation for the eta functions:

$$\eta\left(-\frac{1}{z}\right) = \sqrt{\frac{z}{i}}\,\eta(z). \tag{2.23}$$

The heart of the problem is then proving 2.23. One computes that the Fourier expansion of $\eta(z)$ is

$$\eta(z) = 1 + \sum_{n=1}^{\infty} \left(p_{\mathrm{o}}(n) - p_{\mathrm{e}}(n)\right) e^{2\pi i n z},$$

where $p_{\mathrm{e}}(n)$ and $p_{\mathrm{o}}(n)$ denote the number of partitions of n into odd and even parts, respectively. Then an elementary proof [Sha51] of the following famous theorem elucidates the Fourier coefficients of η:

Theorem 2.3.3.1 (Euler, 1748 [Eul48])**.** *A positive integer n is called a* pentagonal number *if there exists another integer m such that $n = \frac{m(3m\pm 1)}{2}$. We have*

$$p_{\mathrm{o}}(n) - p_{\mathrm{e}}(n) = \begin{cases} (-1)^n & n \text{ is a pentagonal number,} \\ 0 & \text{otherwise.} \end{cases}$$

Upon using this result to identify explicitly the Fourier coefficients of η, one may rewrite η as a theta function:

$$\eta(z) = e^{\pi i z/12} \sum_{n\in\mathbb{Z}} (-1)^n e^{\pi i (3n+1) z},$$

and one proves the functional equation 2.23 upon employing Poisson summation.

A particularly pessimistic observer might claim that this is not a significant improvement over the more well-known proof presented in Subsection 2.3.2. The point of this subsection is to prove the quasimodularity of G_2 using functional equations of L-functions. This way, we avoid the contortions associated with the conditionally convergent series we encountered in Subsection 2.3.2.

Indeed, observe that by the definition of the divisor function, we have

$$\sum_{n=1}^{\infty} n^{-s} \sum_{m=1}^{\infty} m \cdot m^{-s} = \sum_{k=1}^{\infty} \left(\sum_{\substack{n+m=k \\ n,m>0}} 1 \cdot m \right) k^{-s} = \sum_{k=1}^{\infty} \sigma_1(k) k^{-s},$$

which implies by analytic continuation that

$$L(G_2, s) = \zeta(s)\zeta(s-1),$$

where we define

$$L(G_2, s) = \sum_{n=1}^{\infty} \sigma_1(n) n^{-s}.$$

Since the Riemann zeta function satisfies the functional equation

$$\pi^{-\frac{s}{2}} \Gamma\left(\frac{s}{2}\right) \zeta(s) = \pi^{-\frac{1-s}{2}} \Gamma\left(\frac{1-s}{2}\right) \zeta(1-s), \tag{2.24}$$

upon using Legendre's duplication formula for the gamma function,

$$\Gamma(z)\Gamma\left(z+\frac{1}{2}\right) = 2^{1-2z}\sqrt{\pi}\Gamma(2z),$$

we obtain

$$L(G_2, s) = -L(G_2, 2-s). \tag{2.25}$$

One expects that this functional equation implies the desired statement for G_2:

$$G_2\left(-\frac{1}{z}\right) = z^2 G_2(z) - 2\pi i z,$$

for $z \in \mathscr{H}$. This is in fact the case, as we shall prove below. For some perspective, this kind of philosophy, by which one may obtain modular (and more generally, automorphic) forms from L-functions, is known as a *converse theorem*. The first such result is due to Hecke, and is modelled on Riemann's proof that the functional equation for the Jacobi theta function implies the functional equation for his zeta function. The analogue of this result for modular forms on congruence subgroups is more involved, and is due to Weil [Wei67].

Theorem 2.3.3.2 (Hecke, 1936 [Hec36]). *Let f and g be 1-periodic functions on $\mathscr{H}$:*

$$f(z) = \sum_{n=0}^{\infty} a_n e^{2\pi i n z}, \qquad g(z) = \sum_{n=0}^{\infty} b_n e^{2\pi i n z}$$

such that $a_n, b_n = O(n^\alpha)$ for some positive constant α. Set

$$L(f,s) = \sum_{n=1}^{\infty} a_n n^{-s}, \qquad \Lambda(f,s) = \frac{\sqrt{q}}{2\pi}^s \Gamma(s) L(f,s)$$
$$L(g,s) = \sum_{n=1}^{\infty} b_n n^{-s}, \qquad \Lambda(g,s) = \frac{\sqrt{q}}{2\pi}^s \Gamma(s) L(g,s),$$

where q is positive. Then the following are equivalent:

1. *The functions f and g are related by*

$$f\left(-\frac{1}{qz}\right) = (\sqrt{q}z)^k g(z).$$

2. *The functions $\Lambda(f,s)$ and $\Lambda(g,s)$ possess meromorphic continuations into all of the complex plane; the functions*

$$\Lambda(f,s) + a_0 s^{-1} + b_0 i^k (k-s)^{-1} \quad \text{and} \quad \Lambda(g,s) + b_0 s^{-1} + a_0 i^{-k}(k-s)^{-1}$$

are entire and bounded on vertical strips, and are related by

$$\Lambda_f(s) = i^k \Lambda_g(k-s).$$

In any case, it is easy to deduce the quasimodularity of $G_2(z)$ from the functional equation 2.25 for $L(G_2, s)$.

Recall that

$$G_2(z) = 2\zeta(2) - 8\pi^2 \sum_{n=1}^{\infty} \sigma_1(n) e^{2\pi i n z}.$$

We define

$$G_2^*(y) = G_2(iy) = 2\zeta(2) - 8\pi^2 \sum_{n=1}^{\infty} \sigma_1(n) e^{-2\pi i n y}$$

and

$$\Lambda(G_2, s) = (2\pi)^{-s}\Gamma(s) L(G_2, s) = (2\pi)^{-s}\Gamma(s) \sum_{n=1}^{\infty} \sigma_1(n) n^{-s},$$

and observe that, by Mellin's inversion theorem, for fixed σ with $\mathrm{Re}(\sigma) > 2$, we have

$$\begin{aligned}
-\frac{1}{8\pi^2}\left(G_2^*(y) - 2\zeta(2)\right) &= \frac{1}{2\pi i}\int_{\mathrm{Re}(s)=\sigma} \Lambda(G_2, s) y^{-s} ds \\
&= -\frac{1}{2\pi i}\int_{\mathrm{Re}(s)=\sigma} \Lambda(G_2, 2-s) y^{-s} ds \\
&= -y^{-2}\frac{1}{2\pi i}\int_{\mathrm{Re}(u)=2-\sigma} \Lambda(G_2, u) y^{u} du \\
&= -y^{-2}\left(\frac{1}{2\pi i}\int_{\mathrm{Re}(u)=\sigma} \Lambda(G_2, u) y^{u} du - \frac{y^2}{24} + \frac{y}{2} - \frac{1}{24}\right),
\end{aligned}$$

where we have calculated the residues of $\Lambda(G_2, s)$ at the poles at $u = 2$ (caused by the factor $\zeta(u-1)$), $u = 1$ (caused by the factor $\zeta(u)$) and $u = 0$ (caused by the gamma function). By the quintessential inversion property of the Mellin transform,

$$-\frac{1}{8\pi^2}\left(G_2^*\left(\frac{1}{y}\right) - 2\zeta(2)\right) = \frac{1}{2\pi i}\int_{\mathrm{Re}(u)=\sigma} \Lambda(G_2, u) y^{u} du,$$

and by analytic continuation the identity

$$G_2\left(-\frac{1}{z}\right) = z^2 G_2(z) - 2\pi i z$$

is immediate. Upon referring to the proof of Proposition 2.3.2.1, the transformation of $G_2(z)$ under the action of *all* of $SL(2, \mathbb{Z})$ is established.

Chapter 3

Projective geometry and Jacobi's four square theorem

We met Jacobi's theta function in Section 2.2, and we noted there that it satisfies a remarkable transformation law (Theorem 2.2.1.1). Either by taking the fourth power of both sides of Jacob's transformation law, or directly from the higher-dimensional form of Poisson summation (Proposition 1.2.3.1) with Q taken to be the 4×4 identity matrix, we proved that the function Θ, whose Fourier coefficients count the number of ways of representing an integer as a sum of four squares, satisfies:

$$\theta^4(-1/4\tau) = -4\tau^2\theta^4(\tau). \tag{3.1}$$

The aim of this chapter is to explain our attempt to find the simplest possible route to conclude from 3.1 Jacobi's theorem, first stated as Theorem 2.3.2.4:

Theorem 3.0.0.1. *Let $r_4(n)$ denote the number of ways of writing an integer n as a sum of four squares. Then*

$$r_4(n) = 8 \sum_{4 \nmid d \mid n} d.$$

Of course, we require some extra ingredients (c.f. the remarks on the positivity of the Fourier coefficients of Ramanujan's τ-function at the start of Chapter 2), but in contrast to the blunt methods of Section 2.3, we are able cut down on some extraneous details by making use of the fact that $\Theta = \theta^4(\tau)d\tau$ may be realised as a holomorphic one-form on the twice-punctured sphere.

Granted this, by Liouville's theorem, it suffices to show that the function

$$L(\tau) = 1 - 24 \sum_{n=1}^{\infty} \sigma_1(n) e^{2\pi i n \tau}$$

satisfies

$$L\left(\frac{a\tau + b}{c\tau + d}\right) = (c\tau + d)^2 L(\tau) + \frac{6}{\pi i} c(c\tau + d), \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{Z}), \tag{3.2}$$

for then $(\Theta(\tau) - \Theta(\tau + 1/2))\, d\tau$ and $(L(\tau) - L(\tau + 1/2))\, d\tau$ are meromorphic one-forms on the sphere such that their quotient is bounded, so one is a constant multiple of the other. This proves Jacobi's theorem for odd n, and a combinatorial argument suffices to establish Jacobi's theorem for all n. The difficult part is, of course, proving 3.2. Instead of grappling with non-absolutely convergent sums as in Subsection 2.3.2, we deduce 3.2 by using the fact that $L(q) = L(e^{2\pi i \tau})$ and $M(q) = 45G_4(e^{2\pi i \tau})/\pi^4$ solve a certain differential equation,

$$12q\frac{df}{dq} - f^2 + M(q) = 0, \tag{3.3}$$

first studied in connection with modular forms by Ramanujan [Ram57]. This differential equation is of Ricatti type, and the associated equation for the local potential $f(\tau) = -\frac{6}{\pi i}\frac{g'(\tau)}{g(\tau)}$,

$$\frac{d^2g}{d\tau^2} + \frac{\pi^2}{36}M(\tau)g(\tau) = 0$$

has the exciting feature of being projectively invariant, which ties in nicely with the geometric theme of this chapter.

We must admit that we use a bit of number theory in order to show that $L(\tau)$ does indeed satisfy 3.3. Comparing Fourier coefficients, this is the result that

$$5\sigma_3(n) + \sigma_1(n)(1-6n) = \sum_{k=1}^{n-1}\sigma_1(k)\sigma_1(n-k),$$

and a neat elementary proof exists due to Skoruppa [Sko93].

Any and all original results described in this chapter were obtained by joint work with Michael Eastwood.

3.1 The geometry of punctured spheres

In 1569, the Dutch cartographer Mercator discovered a projection of the sphere, punctured at the North and South poles, with the remarkable property that courses of constant bearing on the surface of the sphere correspond to straight lines on the image of the projection [Sny97, pp. 157]. Consequently, atlases produced using Mercator's projection are extremely useful in navigation.

From a mathematical point of view, Mercator's projection gives a diffeomorphism between the twice-punctured sphere and the cylinder in such a way that $q = e^{2\pi i\tau}$ is a local coordinate on $\mathbb{S}^2 \setminus \{q = \infty\}$. In the first subsection, we use this observation to give a short proof of a vital identity.

In the second subsection, we realise the thrice-punctured sphere as the quotient of the upper half plane by the ubiquitous $\Gamma_0(4)$. We use the Riemann Mapping Theorem and the Schwartz reflection principle in order to give a treatment independent of the general considerations of Subsection 1.1.1.

In the last subsection, we sketch a proof of a theorem due to Eastwood and Gover, and as an application, we determine some particular automorphisms of the thrice-punctured sphere in explicit coordinates. This allows us to characterise the meromorphic extension of Θ to the sphere by its poles and residues.

3.1.1 The twice-punctured sphere

The purpose of this subsection is to find the Fourier coefficients of the Eisenstein series G_4 using the Mercator projection.

Proposition 3.1.1.1. *If $q = e^{2\pi i\tau}$ and $|q| < 1$, then*

$$\sum_{d=-\infty}^{\infty}(\tau + d)^{-2} = -4\pi^2\sum_{m=1}^{\infty} mq^m.$$

Proof. The sum on the left hand side converges uniformly on compact subsets of $\mathbb{C} \setminus \mathbb{Z}$, and is invariant under $\tau \mapsto \tau + 1$. It therefore descends to a holomorphic function on the thrice-punctured sphere, on which we use q as a coordinate:

$$\mathbb{S}^2 \setminus \{q = 0, 1, \infty\}.$$

We call this function $F(q)$ and note that $F(1/q) = F(q)$ and $F(q) \to 0$ as $q \to 0$. By Riemann's removeable singularities theorem, $F(q)$ extends holomorphically across the punctures at $q = 0$ and $q = \infty$, and at $q = 1$ it has a double pole. By Liouville's theorem,

$$F(q) = C\frac{q}{(q-1)^2}$$

for some constant C. This constant is easily determined by evaluating $F(q)$ at $\tau = 1/2$:

$$C = -16\sum_{d=-\infty}^{\infty}(2d+1)^{-2} = -4\pi^2.$$

On the other hand, since $|q| < 1$,

$$\frac{q}{(q-1)^2} = q\frac{d}{dq}\frac{1}{1-q} = q\frac{d}{dq}\sum_{m=0}^{\infty} q^m = \sum_{m=1}^{\infty} mq^m. \tag{3.4}$$

□

Corollary 3.1.1.2. *If $q = e^{2\pi i\tau}$ and $|q| < 1$, then*

$$\sum_{d=-\infty}^{\infty} (\tau + d)^{-4} = \frac{8\pi^4}{3}\sum_{m=1}^{\infty} m^3 q^m.$$

Proof. The identity follows upon applying $\frac{d}{d\tau}$ to 3.4 twice and using the chain rule

$$\frac{d}{d\tau} = 2\pi i q\frac{d}{dq}.$$

□

We now use Corollary 3.1.1.2 to find the Fourier coefficients of the q-expansion of G_4.

$$\begin{aligned} G_4(\tau) &= \sum_{\substack{d=-\infty \\ d\neq 0}}^{\infty} \frac{1}{d^4} + \sum_{\substack{c=-\infty \\ c\neq 0}}^{\infty}\sum_{d=-\infty}^{\infty} \frac{1}{(c\tau+d)^4} \\ &= 2\zeta(4) + 2\sum_{c=1}^{\infty}\sum_{d=-\infty}^{\infty}\frac{1}{(c\tau+d)^4} \\ &= \frac{\pi^4}{45} + 2\sum_{c=1}^{\infty}\left(\frac{8\pi^4}{3}\sum_{m=1}^{\infty} m^3 e^{2\pi i c m\tau}\right) \\ &= \frac{\pi^4}{45} + \frac{16\pi^4}{3}\sum_{m=1}^{\infty}\left(\sum_{d|m} d^3\right) e^{2\pi i m\tau} \\ &= \frac{\pi^4}{45} + \frac{16\pi^4}{3}\sum_{m=1}^{\infty}\sigma_3(m) e^{2\pi i m\tau}. \end{aligned}$$

3.1.2 The thrice-punctured sphere

Recall the congruence subgroup $\Gamma_0(4)$, which featured so prominently in Section 2.3, and was shown to be inextricably linked to the fortunes of Jacobi's theta function in 2.2.3:

$$\Gamma_0(4) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{Z}) \mid c = 0 \bmod 4 \right\}.$$

As $\Gamma_0(4)$ plays a central role in the proof of the Jacobi's theorem in Section 2.3, so too is it critical to this chapter. In Subsection 1.1.1, we showed that the space of orbits of the upper half plane under the action of a congruence subgroup admits the structure of an open Riemann surface, and may be compactified to form a compact Riemann surface. In this subsection, we are able to prove a stronger result in the special case that $\Gamma = \Gamma_0(4)$: we show that there is a biholomorphism between $\Gamma_0(4)\backslash\mathscr{H}$ and $\mathbb{S}^2 \setminus \{0, 1, \infty\}$.

For this, we require the Riemann mapping theorem and the Schwartz reflection principle. In this chapter, we use the notation $\mathbb{D}$ to denote the interior of the complex unit disk:

$$\mathbb{D} = \{z \in \mathbb{C} \mid |z| < 1\}.$$

Theorem 3.1.2.1 (Riemann mapping theorem, [FB09, Theorem IV. 4. 5])**.** *Let U be a non-empty simply-connected proper open subset of $\mathbb{C}$. Then there exists a biholomorphism $f : U \to \mathbb{D}$.*

One verifies easily that the map $z \mapsto \frac{i-z}{i+z}$ is a biholomorphism from $\mathscr{H}$ to $\mathbb{D}$.

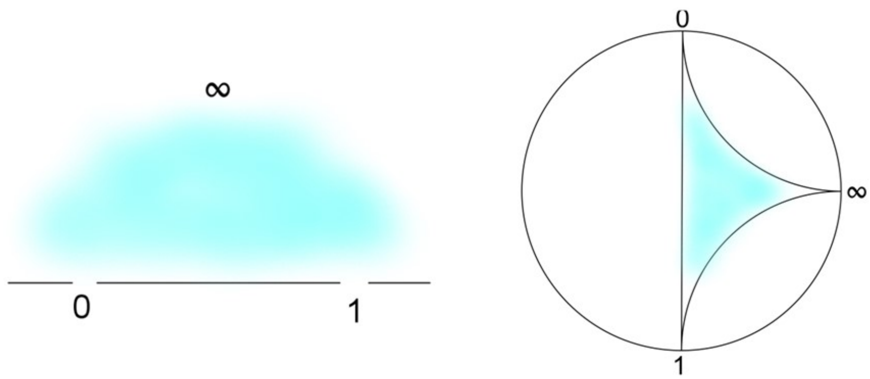
Proposition 3.1.2.2 (Schwartz reflection principle, [FB09, II.3, Exercise 12])**.** *Let F be a continuous function on $\{z \in \mathbb{C} \mid \mathrm{Im}(z) \geq 0\}$ such that it is holomorphic on $\mathscr{H}$ and real-valued on $\mathbb{R}$. Then F may be extended to a holomorphic function on $\mathbb{C}$ by $F(\bar{z}) = \overline{F(z)}$.*

Following Tony Scholl [personal communication to Michael Eastwood, 17th May 1984], we use Theorem 3.1.2.1 and Proposition 3.1.2.2 prove the following result.

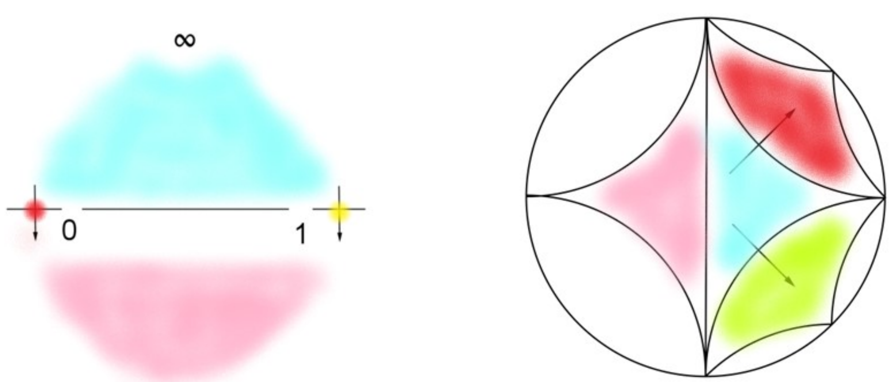
Proposition 3.1.2.3. *The universal cover of $\mathbb{S}^2 \setminus \{0, 1, \infty\}$ is $\mathscr{H}$.*

Proof. The proof is effected by pictures. We begin by using the Riemann mapping theorem to assert the existence of a biholomorphism between $\mathbb{S}^2 \setminus \{0, 1, \infty\} \cong \mathscr{C} \setminus \{0, 1\}$ and a hyperbolic triangle inscribed in $\mathbb{D}$. Since we may take any three points on the boundary of $\mathbb{D}$ to any other three points via a biholomorphism of $\mathbb{D}$, we are free to place the images of $0, 1$ and ∞ at the indicated positions.

We then map the lower half plane (minus two points) onto the reflected triangle in the $\mathbb{D}$, noting that the two maps combine to form a biholomorphism of $\mathbb{C} \setminus \{0, 1, \infty\}$ onto the union of the two triangles by the Schwartz reflection principle.

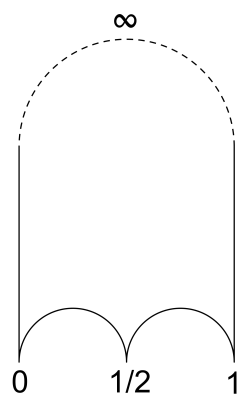


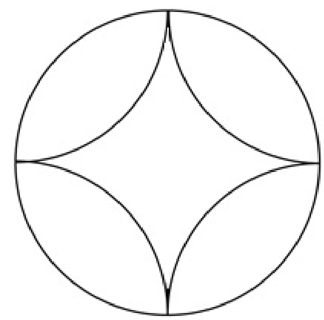
Next, we work out the image of a path in the complex plane passing around the punctures at 0 and 1. If we move along paths from the upper half plane towards the lower half plane around the punctures at 0 and 1 in the indicated directions, then the images of these paths in $\mathbb{D}$ head towards parts of the disk not in the image of our biholomorphism. Therefore, we use the Riemann mapping theorem again to place hyperbolic triangles at these new regions, each sharing the correct part of the boundary with one of the old triangles, and extend our biholomorphism of $\mathbb{C} \setminus \{0, 1, \infty\}$ to the appropriate covering.



Continuing in this fashion, we obtain a tessellation of $\mathbb{D}$ by hyperbolic triangles. Recognising the lower half plane with three points removed as biholomorphic to $\mathbb{S}^2 \setminus \{0, 1, \infty\}$, we see that we have proved that $\mathscr{H}$ is the universal cover of $\mathbb{S}^2 \setminus \{0,1,\infty\}$. In fact, we can go further and identify the quotient of $\mathscr{H}$ to which $\mathbb{S}^2 \setminus \{0,1,\infty\}$ is biholomorphic.

Indeed, instead of picturing the disk as tessellated by triangles, we may glue pairs of triangles together along the edge corresponding to the segment of the real axis joining 0 to 1 and view the tessellation as comprised of astroids. One easily write down a fractional linear transformation describing a biholomorphism between the astroid and the figure in the upper half plane with hyperbolic geodesics for edges and cusps at 0, 1, $\frac{1}{2}$ and ∞.





But the figure in the upper half plane is a fundamental domain for the action of $\Gamma_0(4)$, and the procedure for gluing together the astroids in the disk correspond exactly to the rules for joining the fundamental domains to produce a tessellation of $\mathscr{H}$ by $\Gamma_0(4)$. So we have identified the thrice-punctured sphere as the quotient of $\mathscr{H}$ by $\Gamma_0(4)$. $\square$

We have two obvious local coordinates. Indeed, by Proposition 2.1.1.2, $\Gamma_0(4)$ is generated by

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 \\ 4 & 1 \end{pmatrix},$$

and we have (as sets)

$$\Gamma_0(4)\backslash\mathscr{H} = \mathscr{H} \Big/ \left\{ \tau \sim \tau + 1, \tau \sim \frac{\tau}{4\tau + 1} \right\},$$

so one coordinate is $q = e^{2\pi i \tau}$. On the other hand, we have the natural inclusion $z : \mathbb{C} \hookrightarrow \mathbb{C} \cup \{\infty\}$.

A fundamental class of sections of bundles on $\mathbb{S}^2 \setminus \{0,1,\infty\}$ is the space of meromorphic one-forms on $\mathbb{S}^2$ with at worst poles at the three marked points $\{0,1,\infty\}$. In terms of the coordinate z, there is a special one dz/z, which is holomorphic away from simple poles at 0 and ∞, with residues 1 and -1 respectively.

The coordinate τ has an important geometric interpretation too: from the discussion above, it is defined up to real Möbius transformations. As we demonstrated in Subsection 1.1.3, this means that τ confers a projective structure upon $\mathbb{S}^2 \setminus \{0,1,\infty\}$.

A summary of the coordinates we have introduced and their actions on the marked points is as follows:

q	0	-1	1
τ	∞	$\frac{1}{2}$	0
z	0	1	∞

3.1.3 Puncture repair for compact Riemann surfaces

In this section, we outline a proof of the result due to Eastwood and Gover that there is no difference between a marked compact connected conformal manifold and a punctured compact connected conformal manifold. We outlined in 1.1.3 a proof of the fact that when the dimension of the manifold is two, an oriented conformal structure is the same as a complex structure, so their result implies:

Proposition 3.1.3.1. *Suppose M is a compact connected Riemann surface and $p \in M$. Suppose that N is a compact connected Riemann surface and there is an open subset $U \subset N$ such that U is biholomorphic to $M \setminus \{p\}$. Then this biholomorphism extends to give $M \cong N$.*

We will sketch the proof of the general case.

Theorem 3.1.3.2 (Eastwood–Gover [GE18])**.** *Suppose M is a connected compact conformal manifold and $p \in M$. Suppose that N is a connected compact conformal manifold and there is an open subset U of N such that U is conformally equivalent to $M \setminus \{p\}$. Then this conformal equivalence extends to one between M and N.*

Proof. First, we reduce Theorem 3.1.3.2 to the case in which U, M and N are embedded conformal submanifolds of $\mathbb{R}^n$. Without loss of generality, we may place p at the origin, and we may select a Riemannian manifold in the conformal class of M such that its metric g_{ab} agrees with the Euclidean metric η at the origin. The volume form of g_{ab} is of the form

$$F(x^1, \dots x^n) dx^1 \wedge \cdots \wedge dx^n,$$

and $F(0) = 1$. By continuity,

$$||X||_{g(x)} \leq 2||X||_\eta \leq 4||X||_{g(x)} \tag{3.5}$$

for all $X \in T_x M$ with $||x||_\eta < \epsilon$. Let us change to spherical coordinates

$$\mathbb{R}_{>0} \times \mathbb{S}^{n-1} \ni (r, x) \mapsto rx \in \mathbb{R}^n \setminus \{0\}$$

near the origin and suppose that the conformally rescaled metric $\hat{g} = \Omega^2 g$ on U extends to N. If ∂U contains at least two points, then the concentric hyperspheres $\{r = \epsilon\}$ have diameter bounded away from 0 in the $\hat{g}$ metric, and hence also in the metric $\Omega^2 \eta$ by 3.5. If we can prove that the volume of the collar $\{0 < r < \epsilon\}$ with respect to $\hat{\eta}$ is infinite, then it follows by 3.5 that the volume of the collar with respect to $\hat{g}$ is infinite, which contradicts our assumption that $\hat{g}$ extends to N.

We now prove the theorem in Euclidean space by demonstrating that if Ω is a smooth function defined near the origin with the property that the conformally rescaled metric $\hat{\eta} = \Omega^2 \eta$ extends to N, and ∂U contains at least two points, then the volume of the collar is infinite with respect to the rescaled Euclidean metric $\hat{\eta}$.

Indeed, since ∂U contains at least two points, the concentric hypersurfaces $r = \epsilon$ have diameter bounded away from 0 in the $\hat{nu}$ metric as $\epsilon \to 0$, so for some $\epsilon > 0$ and $\ell > 0$ and any $0 < r < \epsilon$, there exist $\alpha, \beta \in \mathbb{S}^{n-1}$ such that

$$\int_\alpha^\beta \Omega(r, x) r \geq \ell,$$

where the integral is taken along any smooth path from α to β on $\mathbb{S}^{n-1}$.

Lemma 3.1.3.3. *Suppose $\Omega : \mathbb{S}^{n-1} \to \mathbb{R}_{>0}$ is smooth and that there are two points α and $\beta \in \mathbb{S}^{n-1}$ such that $\int_\alpha^\beta \geq d > 0$ for all smooth paths from α to β on $\mathbb{S}^{n-1}$. Then*

$$\int_{\mathbb{S}^{n-1}} \Omega^n \geq C_n d^n.$$

We omit the proof of the lemma, instead referring the reader to Eastwood and Gover's article [GE18]. The proof of the theorem now follows upon noting that the volume of the collar is infinite,

$$\int_0^\epsilon \int_{\mathbb{S}}^{n-1} \Omega^n r^{n-1} dr \geq \int_0^\epsilon C_n \frac{\ell n}{r^n} r^{n-1} dr = C_n \int_0^\epsilon \frac{\ell n}{r} dr = \infty,$$

so $\hat{\eta}$ cannot extend smoothly to N, which contradicts our earlier assumption. □

Remark 3.1.3.4. *In the case of dimension two, the proof sketched above becomes morally equivalent to a result of Ahlfors and Beurling [AB50], who proved that the extremal length of a family of curves winding once around a 2-dimensional puncture is zero (whereas the extremal length of the same family around viewed on $N \not\cong M$ would be positive). This was observed by Ben Warhurst.*

We now use Theorem 3.1.3.2 to investigate some of the automorphisms of $\mathbb{S}^2 \setminus \{0, 1, \infty\}$. Indeed, by Theorem 3.1.3.2, the automorphisms of $\mathbb{S}^2 \setminus \{0, 1, \infty\}$ are exactly the automorphisms of $\mathbb{S}^2 \cong \mathbb{CP}^1$ which permute the marked points. Since any automorphism of $\mathbb{S}^2$ is specified uniquely by its action on any three distinct points, the automorphisms of $\mathbb{S}^2 \setminus \{0, 1, \infty\}$ correspond to permutations of $\{0, 1, \infty\}$. By Proposition 3.1.2.3, an automorphism of $\mathbb{S}^2 \setminus \{0, 1, \infty\}$ is an automorphism of $\mathscr{H}$ such that its conjugates by $\Gamma_0(4)$ are in $\Gamma_0(4)$. In particular,

$$\begin{pmatrix} 0 & -1/2 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d & -c/4 \\ -4b & a \end{pmatrix} \begin{pmatrix} 0 & -1/2 \\ 2 & 0 \end{pmatrix},$$

we have that

$$\phi : \tau \mapsto -1/4\tau$$

is an automorphism of $\mathbb{S}^2 \setminus \{0, 1, \infty\}$. In the z-coordinate, this is the map that swaps 0 and ∞ and fixes 1, namely

$$z \mapsto 1/z.$$

Secondly,

$$\begin{pmatrix} 0 & 1/2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a + c/2 & b + (d-a)/2 - c/4 \\ -c & d - c/2 \end{pmatrix} \begin{pmatrix} 1 & 1/2 \\ 0 & 1 \end{pmatrix},$$

so

$$\psi : \tau \mapsto \tau + 1/2$$

is also an automorphism of $\mathbb{S}^2 \setminus \{0, 1, \infty\}$. In z-coordinates, this is the automorphism that swaps 1 and ∞ and fixes 0, which we recognise as

$$z \mapsto z/(z-1).$$

3.2 A digression into function theory

In this section we state and prove everything we need to know about the fourth power θ^4 of Jacobi's theta function in order to conclude Lagrange's theorem from general geometric considerations. For θ^4, this means that we characterise it as a particular kind of holomorphic one-form on the thrice punctured sphere with simple poles of prescribed residues at the punctures.

We will see later that $L(\tau)$ also gives rise to a one-form on the thrice punctured sphere with the same characteristics as the one-form for theta, but in this section we content ourselves with presenting the number-theoretic details involved in showing that it satisfies the differential equation mentioned earlier.

3.2.1 Jacobi's theta function

As mentioned in the introduction to this chapter, our starting point is Jacobi's transformation law, proved by taking a fourth power in Theorem 2.2.1.1 or by setting Q equal to the 4×4 identity matrix in Proposition 1.2.3.1:

Theorem 3.2.1.1. *Let $r_4(n)$ denote the number of ways of writing an integer as a sum of four squares. The function*

$$\theta^4(\tau) = \sum_{n=0}^{\infty} r_4(n) e^{2\pi i n \tau}$$

is holomorphic for $\tau \in \mathscr{H}$, and satisfies

$$\theta^4\left(-\frac{1}{4\tau}\right) = -4\tau^2\theta^4(\tau). \tag{3.6}$$

Define $\phi : \mathscr{H} \to \mathscr{H}$ by $\phi(\tau) = -1/4\tau$. Then an alternative way to state 3.6 in terms of the holomorphic one-form $\Theta = \theta^4 d\tau$ is

$$\phi^*\Theta = -\Theta.$$

If we also introduce $T : \mathscr{H} \to \mathscr{H}$ by $T(\tau) = \tau + 1$, then we have

$$T^*\Theta = \Theta.$$

Note that if we set $R = \phi \circ T^{-1} \circ \phi$, then $R(\tau) = \tau/(4\tau + 1)$ and

$$R^*\Theta = \Theta.$$

By Proposition 2.1.1.2, R and T generate $\Gamma_0(4)$, so summing up:

Theorem 3.2.1.2. *The holomorphic one-form Θ descends to the thrice-punctured sphere, and under the automorphism ϕ, $\phi^*\Theta = -\Theta$.*

Corollary 3.2.1.3. *In the z-coordinate on $\mathbb{S}^2 \setminus \{0, 1, \infty\}$,*

$$\Theta = \frac{1}{2\pi i} \frac{dz}{z}.$$

Proof. In q-coordinates,

$$\Theta = \frac{1}{2\pi i q}\left(1 + 8q + 24q^2 + 32q^3 + \dots\right) dq$$

near $q = 0$, so Θ extends meromorphically through $q = 0$, with a simple pole there of residue $1/2\pi i$. The truth of this statement is independent of the choice of local coordinates, so we have in the z-coordinate:

$$\Theta = \frac{1}{2\pi i z}\left(1 + \dots\right) dz.$$

In the z-coordinate, the automorphism ϕ interchanges $z = 0$ and $z = \infty$ whilst fixing $z = 1$, so the relation $\phi^*\Theta = -\Theta$ implies that Θ has a simple pole at $z = \infty$ with residue $-1/2\pi i$.

Finally, recall that the automorphism ψ swaps $z = 1$ and $z = \infty$ whilst fixing $z = 0$. As a consequence, we may compare the behaviour of Θ along the line $\mathrm{Re}(\tau) = 0$ with its behaviour along $\mathrm{Re}(\tau) = 1/2$. For convenience, we set $\tau = it$ so that we may deal with real positive t. With $q = e^{-2\pi t}$,

$$\begin{aligned} -i\Theta &= \left(1 + 8q + 24q^2 + 32q^3 + \dots\right) dt, \\ -i\psi^*\Theta &= \left(1 - 8q + 24q^2 - 32q^3 + \dots\right) dt. \end{aligned}$$

We know that $\Theta(it)$ has a simple pole at $t = 0$. Since $\Theta(\tau)$ is real-valued when $\mathrm{Re}(\tau) = 0$ or $\mathrm{Re}(\tau) = 1/2$, and the q-expansion coefficients are all non-negative, $\Theta(1/2 + it)$ is dominated by $\Theta(1/2 + it)$ as $t \to 0^+$. To exclude the gloomy possibility of an essential singularity, we observe that the intersection of any circle centred at $\tau = 1/2$ with a well-chosen fundamental domain containing $\{1/2 + it \mid t \in \mathbb{R}\}$ is a finite curve, so the maximal value of $\Theta(\tau)$ as τ runs over the semicircle is bounded by the value of $\Theta(is)$ for some real s. So the behaviour of Θ at $z = 1$ is certainly no worse than the behaviour at $z = 0$.

In summary, the holomorphic one-form Θ extends to a meromorphic one-form on $\mathbb{S}^2$ with simple poles at 0 and ∞ with residues $1/2\pi i$ and $-1/2\pi i$ respectively. By the residue theorem, since Θ may not have a pole of order ≥ 2 at $z = 1$, it must extend holomorphically across $z = 1$, and having identified exactly two simple poles, Θ cannot have any zeros. Thus, we have uniquely characterised Θ. □

3.2.2 A quasi-modular Eisenstein series

We have already met some Eisenstein series in Subsection 2.1.2, and we observed there that if $k \geq 4$, then the sum defining the series

$$G_k(\tau) = \sum_{(c,d)\in\mathbb{Z}^2\setminus\{(0,0)\}} (c\tau + d)^{-4} = 2\zeta(k) + 2\frac{(2\pi i)^k}{(k-1)!}\sum_{n=1}^{\infty} \sigma_{k-1}(n) e^{2\pi i n\tau}$$

converges absolutely, so we may rearrange the sum to obtain

$$G_k\left(\frac{a\tau + b}{c\tau + d}\right) = (c\tau + d)^k G_k(\tau). \tag{3.7}$$

When $k = 2$, the sum defining G_2 doesn't converge absolutely, and 3.7 isn't true. Instead, a very tedious calculation revealed that $G(\tau)$ satisfies

$$G_2\left(\frac{a\tau + b}{c\tau + d}\right) = (c\tau + d)^2 G_2(\tau) - 2\pi i c(c\tau + d) \quad \text{for all } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{Z}). \tag{3.8}$$

The purpose of the rest of the chapter is to provide a proof of 3.8 which eschews the fiddly calculations of Subsection 2.3.2 in favour of manipulations inspired by geometry. To stress that we begin with the q expansion of G_2 and *nothing else*, we make the following definition (the notation is due to Ramanujan [Ram57]):

Definition 3.2.2.1. *For $\tau \in \mathscr{H}$,*

$$L(\tau) = 1 - 24\sum_{m=1}^{\infty} \sigma_1(m) e^{2\pi i m\tau}.$$

In terms of $L(\tau)$, we may restate 3.8 as

$$L\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^2L(\tau) + \frac{6}{\pi i}c(c\tau+d) \quad \text{for all } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{Z}).$$

Now, using only elementary number theory, we prove the result which will lead us to 3.8.

Theorem 3.2.2.2. *With $L(\tau)$ as above, $q = e^{2\pi i\tau}$ as usual, and $M(\tau) = 1 + 240\sum_{n=1}^{\infty}\sigma_3(n)q^n$,*

$$12q\frac{dL}{dq} - L^2 + M(q) = 0, \tag{3.9}$$

Comparing q-expansion coefficients of both sides of 3.9, we find that we need to prove that

$$5\sigma_3(n) + \sigma_1(n)(1-6n) = 12\sum_{m=1}^{n-1}\sigma_1(m)\sigma_1(n-m). \tag{3.10}$$

The first elementary proof of 3.10 seems to be deducible from results of Liouville [Lio58], and a more recent treatment is due to Rankin [Ran76] in the context of finding elementary proofs for the analogous identities stemming from the algebraic dependencies between the Eistenstein series. A simple proof is due to Skoruppa [Sko93], and we will sketch his proof here for completeness. First, we require a lemma.

Lemma 3.2.2.3. *Let ℓ be a positive integer. For any pair (a,b) of positive integers, let*

$$\Lambda_\ell(a,b) = \#\{(x,y) \in \mathbb{Z}_{\geq 0}\times\mathbb{Z}_{\geq 0} \text{ such that } ax+by=\ell\},$$

and extend Λ_ℓ to $\mathbb{Z}^2$ by requiring that it be odd in each variable:

$$\Lambda_\ell = \begin{cases} \operatorname{sgn}(ab)\Lambda_\ell(|a|,|b|) & ab \neq 0, \\ 0 & ab = 0. \end{cases}$$

Then for any triple (a,b,c) of integers such that $a+b+c=0$,

$$\Lambda_\ell(a,b) + \Lambda_\ell(a,c) + \Lambda_\ell(b,c) = \begin{cases} 1 & abc \neq 0, k \mid \ell, \\ 1-\ell/k & abc = 0, k \neq 0, k \mid \ell, \\ 0 & \textit{otherwise}, \end{cases} \tag{3.11}$$

where $k = \max\{|a|,|b|,|b|\}$.

Proof. Both sides of 3.11 are invariant under $(a,b,c) \mapsto (-a,-b,-c)$ and any permutation of $\{a,b,c\}$. Since $a+b+c=0$, $abc \leq 0$, so we only need to deal with the cases $a=0$ or $a,b>0>c$. If $a=0$, only $\Lambda_l(b,c)$ can possibly be nonzero, and this is $1-\ell/|b|$ if $b \mid l$ and 0 otherwise.

If $a,b>0>c$, then 3.11 becomes

$$\Lambda_\ell(a,b) + \Lambda_\ell(a,c) + \Lambda_\ell(b,c) = \begin{cases} 1 & a+b \mid c. \\ 0 & \text{otherwise.} \end{cases}$$

To prove this, observe that the map $(x,y) = (z-w,w)$ puts the solutions of $az+bw=\ell$ with $z>w$ in bijective correspondence with the solutions of $ax+by=l$, whilst the map $(x,y)=(w-z,z)$ puts the solutions of $az+bw=\ell$ with $z<w$ in bijective correspondence with the solutions of $ax+(a+b)y=\ell$. □

Now for the proof of the Theorem.

Proof of Theorem 3.2.2.2. Define a linear operator U on the space of functions h on $\mathbb{Z}^2$ by

$$Uh(x,y) = h(y,y-x).$$

We begin by proving that, for any positive integer ℓ,

$$\sum_{ax+by=\ell}(h(a,b)-h(a,-b))=\sum_{k|\ell}\left(\frac{\ell}{k}h(k,0)-\sum_{j=0}^{k-1}h(k,j)\right). \tag{3.12}$$

First, for $(a,b)\in\mathbb{Z}^2$, define another function $\delta_{a,b}$ on $\mathbb{Z}^2$ by $\delta_{a,b}(x,y)=1$ if $(a,b)=(x,y)$ and 0 otherwise. If we let $h=\sum_{j=0}^{5}U^j\delta_{a,-b}$, then 3.12 reduces to 3.11. By linearity, the identity 3.12 is therefore true for any function satisfying $Uh=h$, since all such functions may be written as

$$h=\frac{1}{6}\sum_{j=0}^{5}U^j\sum_{a,b}h(a,b)\delta_{a,b}.$$

Theorem 3.2.2.2 is now a straightforward consequence of 3.12. Indeed, choose $h(x,y)=-\frac{1}{2}(x^2-xy+y^2)$, and observe that $h(a,b)=h(a,-b)=ab$. Therefore the kth term on the right hand side of 3.12 is

$$-\frac{1}{2}k\ell+\frac{5}{12}k^2+\frac{1}{12}k,$$

and 3.10 follows upon noting that

$$\sum_{k|\ell}\left(-\frac{1}{2}k\ell+\frac{5}{12}k^2+\frac{1}{12}k\right)=-\frac{1}{2}\ell\sigma_1(\ell)+\frac{5}{12}\sigma_3(\ell)+\frac{1}{12}\sigma_1(\ell),$$

and

$$\sum_{ax+by=\ell}ab=\sum_{m=1}^{\ell-1}\sum_{\substack{a|m\\ b|(\ell-m)}}ab=\sum_{m=1}^{\ell-1}\sum_{a|m}\sum_{b|(\ell-m)}=\sum_{m=1}^{\ell-1}\sigma_1(m)\sigma_1(\ell-m). \qquad \square$$

3.3 A projectively invariant differential operator

Let us remind the reader what we have proved so far. We know that

$$\Theta=\frac{1}{2\pi i}\left(1+8q+24q^2+32q^3+\dots\right)\frac{dq}{q}$$

is a holomorphic one-form on the thrice punctured sphere with a meromorphic extension across the punctures, characterised up to scale by the fact that its extension is holomorphic at one puncture with simple poles of residues $\pm\frac{1}{2\pi i}$ at the other two punctures.

We also have a holomorphic function $L(q)$, defined by

$$L(q)=1-24\sum_{m=1}^{\infty}\sigma(n)q^n,$$

which we know satisfies a certain differential equation:

$$12q\frac{dL}{dq}-L^2+M=0.$$

Our aim in this section is to prove that this implies that $L(\tau)$ is a quasi-modular form. It will follow, with a small amount of extra work, that

$$(L(\tau)-L(\tau+1/2))\,d\tau \quad \text{and} \quad \left(\theta^4(\tau)-\theta^4(\tau+1/2)\right)d\tau$$

are holomorphic $\Gamma_0(4)$-invariant one-forms, and Jacobi's theorem will be a simple consequence.

3.3.1 Projective invariance and an $SL(2,\mathbb{Z})$ representation

Upon changing variables from q to τ, the differential equation satisfies by L becomes

$$\frac{6}{\pi i}\frac{dL}{d\tau} - L^2 + M = 0.$$

This is a Ricatti equation: following standard procedure, we write

$$L(\tau) = -\frac{6}{\pi i}\frac{g'(\tau)}{g(\tau)} \tag{3.13}$$

locally and investigate the *linear* differential equation satisfied by the "potential":

$$g''(\tau) + \frac{\pi^2}{36}M(\tau)g(\tau) = 0.$$

Henceforth, let $\mathscr{V}$ denote the space of solutions f to the differential equation

$$f''(\tau) + \frac{\pi^2}{36}M(\tau)f(\tau) = 0. \tag{3.14}$$

Since $\mathscr{H}$ is simply connected, $\mathscr{V}$ must be two-dimensional. We have a representation

$$\rho : SL(2,\mathbb{Z}) \longrightarrow GL(2,\mathscr{V})$$

by $\rho(\gamma)(f)(\tau) = (c\tau+d)f(\gamma^{-1}\tau)$ for $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Indeed,

$$\frac{d^2}{d\tau^2}(c\tau+d)f\left(\frac{a\tau+b}{c\tau+d}\right) = \frac{1}{(c\tau+d)^3}f''\left(\frac{a\tau+b}{c\tau+d}\right),$$

so that if $f \in \mathscr{V}$, then

$$\begin{aligned}\frac{d^2}{d\tau^2}(c\tau+d)f\left(\frac{a\tau+b}{c\tau+d}\right) + \frac{\pi^2}{36}M(\tau)(c\tau+d)g\left(\frac{a\tau+b}{c\tau+d}\right) &= \frac{1}{(c\tau+d)^3}f''\left(\frac{a\tau+b}{c\tau+d}\right) + \frac{\pi^2}{36}M(\tau)(c\tau+d)f\left(\frac{a\tau+b}{c\tau+d}\right)\\ &= \frac{1}{(c\tau+d)}\left(-\frac{\pi^2}{36}M\left(\frac{a\tau+b}{c\tau+d}\right)f\left(\frac{a\tau+b}{c\tau+d}\right)\right) + \frac{\pi^2}{36}M(\tau)(c\tau+d)f\left(\frac{a\tau+b}{c\tau+d}\right)\\ &= 0,\end{aligned}$$

where we have used that $M(\gamma\tau) = (c\tau+d)^4M(\tau)$. So we have proved the following result.

Proposition 3.3.1.1. *The differential operator implicit at 3.14 is projectively invariant.*

It will be important to keep in mind the action on $\mathscr{V}$ of the generators of $PSL(2,\mathbb{Z})$:

$$T(f)(\tau) = f(\tau) \quad \text{and} \quad S(f)(\tau) = \tau f(-1/\tau),$$

where

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

By Proposition 3.3.1.1, if f solves 3.14 then so does

$$(Tf)(\tau) = f(\tau-1) \quad \text{and} \quad (Sf)(\tau) = -\tau f(-1/\tau).$$

We wish to prove that

$$L\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^2L(\tau) + \frac{6}{\pi i}c(c\tau+d), \tag{3.15}$$

for $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL(2,\mathbb{Z})$, which is equivalent to

$$L(\tau+1) = L(\tau) \quad \text{and} \quad L(-1/\tau) = \tau^2L(\tau) + \frac{6}{\pi i}\tau,$$

since $SL(2,\mathbb{Z})$ is generated by T and S. We can now rephrase the quasimodularity of L in terms of a local potential:

Proposition 3.3.1.2. *The function L satisfies the transformation law 3.15 if*

$$T\langle g\rangle \subseteq \langle g\rangle \quad \text{and} \quad S\langle g\rangle \subseteq \langle g\rangle, \tag{3.16}$$

where g is any local potential for L.

Proof. It is clear that if $Tg = \alpha g$ for some constant α, then $TL = L$. For the second statement, if $-\tau g(-1/\tau) = \beta g(\tau)$, then $\beta g(-1/\tau) = g(\tau)/\tau$ and $\beta g'(-1/\tau) = \tau g'(\tau) - g(\tau)$. We calculate:

$$\begin{aligned} -\frac{6}{\pi i}\frac{g'(-1/\tau)}{\tau g(-1/\tau)} &= -\frac{6}{\pi i}\frac{(\tau g'(\tau) - g(\tau)/\beta}{\tau(g(\tau)/\tau)/\beta} \\ &= -\frac{6}{\pi i}\frac{\tau g'(\tau)}{g(\tau)} + \frac{6}{\pi i}, \end{aligned}$$

which means that

$$L(-1/\tau)/\tau = \tau L(\tau) + \frac{6}{\pi i}.$$

□

In the next subsection, we will prove 3.16, thus establishing that L is quasimodular.

3.3.2 Proof of quasi-modularity

We will now finish the proof that L satisfies 3.15; by Proposition 3.3.1.2, we need to show that

$$S\langle g\rangle \subseteq \langle g\rangle \tag{3.17}$$

for some local potential g of L. We suppose to the contrary that 3.17 is not true; then we set $h = Sg$ and it follows that $\{g, h\}$ is a basis for $\mathscr{V}$. We can make this statement more explicit by choosing a particular local potential g:

Lemma 3.3.2.1. *The function*

$$g(\tau) = \exp\left(-\frac{\pi i\tau}{6} + 2\sum_{n=1}^{\infty}\frac{\sigma_1(n)}{n}q^n\right) \tag{3.18}$$

is a potential for L, and is holomorphic on $\mathscr{H}$.

Proof. We take $g(\tau) = \exp\left(-\frac{\pi i\tau}{6}\right)\psi(q)$ as an ansatz and show that we may find a holomorphic function $\psi(q)$ on the unit disk such that g is a potential for L. If we substitute our ansatz into 3.13, we find that

$$\psi - 12q\frac{d\psi}{dq} = L\psi = \psi - 24\psi\sum_{n=1}^{\infty}\sigma_1(n)q^n.$$

Therefore

$$\frac{d}{dq}\log(\psi) = 2\frac{d}{dq}\sum_{n=1}^{\infty}\frac{\sigma_1(n)}{n}q^n,$$

so normalising $\psi(0) = 1$, define ψ by

$$\log(\psi) = 2\sum_{n=1}^{\infty}\frac{\sigma_1(n)}{n}q^n,$$

which clearly converges for $|q| < 1$. □

Emboldened by this sudden appearance of explicit functions, we determine another global element of $\mathscr{V}$, linearly independent from the function at 3.18.

Lemma 3.3.2.2. *There is a power series*

$$\phi(q) = 1 + \frac{10}{7}q + \frac{365}{91}q^2 + \frac{13610}{1729}q^3 + \frac{135701}{8645}q^4 + \frac{7419742}{267995}q^5 + \dots,$$

convergent for $|q| < 1$, with all its q-expansion coefficients positive*, such that the function*

$$k(\tau) = \exp\left(\frac{\pi i\tau}{6}\right)\phi(q) \tag{3.19}$$

is a potential for L, and is holomorphic on $\mathscr{H}$.

Proof. We take 3.19 as an ansatz and substitute it into 3.14 to obtain

$$6q^2\frac{d^2\phi}{dq^2} + 7q\frac{d\phi}{dq} = 10\sum_{n=1}^{\infty}\sigma_3(n)q^n.$$

Normalising $\phi(0) = 1$, we may obtain ϕ as a formal power series with coefficients given by a recursion relation:

$$\phi(q) = \sum_{n=0}^{\infty} a_n q^n, \quad a_0 = 1, a_n = \frac{10}{n(6n+1)}\sum_{j=1}^{n}\sigma_3(j)a_{n-j}.$$

By induction, $a_n > 0$ for all n, so we only need to verify that ϕ converges on the open unit disk. Substituting the function $g(\tau) = \exp\left(\frac{\pi i}{6}\right)\psi(q)$ into 3.14, we find that

$$6q^2\frac{d^2\psi}{dq^2} + 5q\frac{d\psi}{dq} = 10\sum_{n=1}^{\infty}\sigma_3(n)q^n.$$

We already normalised in Lemma 3.3.2.1 so that $\psi(0) = 1$, and writing ψ as a formal power series,

$$\psi(q) = \sum_{n=0}^{\infty} b_n q^n, \quad b_0 = 1, \;\; b_n = \frac{10}{n(6n-1)}\sum_{j=1}^{n}\sigma_3(j)b_{n-j}.$$

Another induction argument shows that $a_n \leq b_n$, and we know that ψ converges in the open unit disk, so it follows by comparison that ψ converges there too. □

Using the particular g from 3.18 to construct the basis for $\mathscr{V}$, and reminding the reader that we are assuming $Sg \notin \langle g\rangle$, we spend the remainder of this subsection proving that it must be the case that $k_0 = h + ig$ is a nonzero multiple of k, where $h = Sg$.

Indeed, by Lemma 3.3.2.1, $Tg = e^{\pi i/6}g$, and by Lemma 3.3.2.2, $Tk = e^{-\pi i/6}k$. It follows that the actions of S and T on the *new* basis $\{h, g\}$ are

$$T\begin{pmatrix} h \\ g\end{pmatrix} = \begin{pmatrix} e^{-\pi i/6} & x \\ 0 & e^{\pi i/6}\end{pmatrix}\begin{pmatrix} h \\ g\end{pmatrix} \quad \text{and} \quad S\begin{pmatrix} h \\ g\end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0\end{pmatrix}\begin{pmatrix} h \\ g\end{pmatrix}$$

for some complex x. We also know that $(ST)^3 = 1$, because S and T satisfy this relation in $SL(2,\mathbb{Z})$, so with tears or a computer, one may determine that

$$T = \begin{pmatrix} e^{-\pi i/6} & 1 \\ 0 & e^{\pi i/6}\end{pmatrix}.$$

We set $k_0 = h + ig$ and note that $Sk_0 = ik_0$, so $k_0(i) = 0$. From the matrix representing T, we have $Tk_0 = e^{\pi i/6}k_0$. But by Lemma 3.3.2.1, $Tk = e^{\pi i/6}k$ too, and therefore k must be a multiple of k_0. By Lemma 3.3.2.1, all the q-expansion coefficients of k are positive, so in particular $k(i) > 0$. This is a contradiction, so Sg must be a multiple of g after all. Thus 3.16 is proved, and so L transforms as required under the action of $SL(2,\mathbb{Z})$.

3.3.3 Proof of Jacobi's theorem

In this final subsection, we bring together all the results so far and prove Jacobi's theorem.

First, we show that the behaviour of $L(\tau)$ under the generators of $SL(2,\mathbb{Z})$ implies that $L(\tau)d\tau$ descends to a meromorphic one-form on $\mathbb{S}^2 \setminus \{0, 1, \infty\}$. Indeed, since $\left(\begin{smallmatrix} 1 & 0 \\ 4 & 1\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 3 & -1 \\ 4 & -1\end{smallmatrix}\right)$ are elements of $SL(2,\mathbb{Z})$, we have

$$L\left(\frac{\tau}{4\tau+1}\right) = (4\tau+1)^2L(\tau) + \frac{24}{\pi i}(4\tau+1),$$

$$L\left(\frac{\tau}{4\tau+1} + \frac{1}{2}\right) = L\left(\frac{3(\tau+1/2)-1}{4(\tau+1/2)-1}\right) = (4\tau+1)^2L(\tau+1/2) + \frac{24}{\pi i}(4\tau+1),$$

and it follows that if we set $\Lambda = L(\tau)d\tau - (\psi^*L)(\tau)d\tau = L(\tau)d\tau - L(\tau+1/2)d\tau$, then

$$\Lambda(\tau+1) = \Lambda(\tau), \quad \Lambda\left(\frac{\tau}{4\tau+1}\right) = \Lambda(\tau).$$

By Proposition 3.1.2.3, Λ descends to a holomorphic one-form on the thrice-punctured sphere.

Next, we characterise the poles and zeros of Λ using the automorphism $\phi : \tau \mapsto -1/4\tau$ from Subsection 3.2.1. Since $z = 0$ corresponds to $\tau = \infty$, it is clear that Λ extends holomorphically across $z = 0$. By the residue theorem, if we can show that Λ extends to a meromorphic one form on the sphere with at worst simple poles at $z = 1$ and $z = \infty$, then the poles must have opposite residues. We require a short lemma.

Lemma 3.3.3.1. *Suppose that $\eta(\tau)$ is a holomorphic function on $\mathscr{H}$. Set $q = e^{2\pi i\tau}$, and suppose that*

1. $\eta(\tau + 1) = \eta(\tau)$,
2. $\eta(\tau)$ *is bounded on the rectangle* $\{\tau = x + iy \mid 1 \leq x \leq 1, y \geq 1\}$.

Then $\eta(\tau)d\tau$ extends to a meromorphic differential form on the unit disk $|q| < 1$ with at worst a simple pole at $q = 0$.

Proof. The effect of the first condition is obvious. For the second, observe that the translation of the boundedness condition into q coordinates is the statement that $\nu(q)$ is bounded on $\{q \mid |q| < e^{-2\pi}$. By the Riemann's removeable singularities theorem, $\nu(q)$ extends holomorphically across $q = 0$, and in τ coordinates this means that $\eta(\tau)d\tau$ has at worst a simple pole at 0. ☐

Now we investigate $\phi^* L(\tau)d\tau$. If we set $\tau' = \phi^*\tau$, then

$$L(\tau')d\tau' = \left(4L(4\tau') + \frac{6}{\pi i\tau'}\right) d\tau',$$

which satisfies the boundedness condition of Lemma 3.3.3.1 in τ' coordinates (with $x + iy$ equal to the image of our z-coordinate), but not the periodicity condition. The same is true for $\phi^*\psi^* L(\tau)d\tau$, so $\phi^*\Lambda$ satisfies both the boundedness *and* the periodicity conditions. We conclude that Λ extends to a meromorphic one-form on the sphere with at worst simple poles at the punctures at $z = 1$ and $z = \infty$.

One may easily check that the results of Subsection 3.2.1 imply that $\Theta - \psi^*\Theta$ extends holomorphically across the puncture at $z = 0$ and has simple poles of residues $\frac{1}{2\pi i}$ and $-\frac{1}{2\pi i}$ at $z = 1$ and $z = \infty$ respectively.

Therefore $\Theta - \psi^*\Theta$ and Λ are meromorphic one-forms on $\mathbb{S}^2 \setminus \{0, 1, \infty\}$ with zeros and poles in the same locations. It follows that one is a constant multiple of the other, and upon comparing Fourier coefficients we obtain

$$r_4(2n + 1) = 8\sigma_1(2n + 1).$$

It is now a matter of wading through some counting arguments to deduce Jacobi's theorem. We require two lemmas. For convenience, we will define

$$X_4(n) = \big\{(x, y, z, w) \mid x^2 + y^2 + z^2 + w^2 = n\big\} \quad \text{and as usual, } r_4(n) = |X_4(n)|.$$

Lemma 3.3.3.2. *If n is even, then*

$$r_4(n) = r_4(2n).$$

Proof. Note that if $(x, y, z, w) \in X_4(2n)$, then x, y, z and w must be all even or all odd. We define maps $\phi : X_4(2n) \to X_4(n)$ and $\psi : X_4(n) \to X_4(2n)$ by

$$\phi : (x, y, z, w) \mapsto \left(\frac{x - y}{2}, \frac{x + y}{2}, \frac{z - w}{2}, \frac{z + w}{2}\right)$$
$$\psi : (a, b, c, d) \mapsto (a + b, a - b, c + d, c - d),$$

and one easily checks that ϕ and ψ are inverses. ☐

Lemma 3.3.3.3. *If n is odd, then*

$$r_4(2n) = 3r_4(n).$$

Proof. Suppose that $(x, y, z, w) \in X_4(2n)$. Since n is odd, exactly two of x, y, z and w must be even and the other two must be odd. We use the terminology *parity pair* to mean two entries of $X_4(2n)$ with the same parity: every tuple in $X_4(2n)$ contains two parity pairs, one even and one odd. We call a parity pair *degenerate* if the entries associated to the pair are equal, and two entries associated to the same parity pair are called *parity partners*. Define

$$\begin{aligned} Y_4(2n) &= \{(x, y, z, w) \in X_4(2n) \mid x \text{ and } y \text{ have the same parity; } z \text{ and } w \text{ have the same parity}\}, \\ Z_4(n) &= \{\text{unordered pairs } \{(a, b, c, d), (a, -b, c, -d)\} \mid (a, b, c, d) \in X_4(n)\}, \end{aligned}$$

and let $\mathcal{F}$ denote the projection of $X_4(n)$ onto $Z_4(n)$:

$$\mathcal{F}(a, b, c, d) = \{(a, b, c, d), (a, -b, c, -d)\}.$$

Then the following subsets form a disjoint partition of $X_4(2n)$, and thus descend to a disjoint partition of $Y_4(2n)$:

$$\begin{aligned} S_1 &= \{\text{4-tuples in } X_4(2n) \text{ with distinct entries}\}, \\ S_2 &= \{\text{4-tuples in } X_4(2n) \text{ with exactly one degenerate parity pair}\}, \\ S_3 &= \{\text{4-tuples in } X_4(2n) \text{ with both parity pairs degenerate}\}, \end{aligned}$$

and we also define subsets effecting a convenient partition of $X_4(n)$:

$$\begin{aligned} T_1 &= \{\text{4-tuples in } X_4(n) \text{ with nonzero 2nd and 4th entries}\}, \\ T_2 &= \{\text{4-tuples in } X_4(n) \text{ with exactly one of the 2nd and 4th entries zero}\}, \\ T_3 &= \{\text{4-tuples in } X_4(n) \text{ with both the 2nd and 4th entries zero}\}. \end{aligned}$$

We now consider the map $\Pi : Y_4(2n) \to Z_4(n)$ defined by

$$\Pi(x, y, z, w) = \left\{ \left(\frac{x+y}{2}, \frac{x-y}{2}, \frac{z+w}{2}, \frac{z-w}{2} \right), \left(\frac{x+y}{2}, \frac{x-y}{2}, \frac{z+w}{2}, \frac{z-w}{2} \right) \right\}.$$

Observe that the images of T_1, T_2 and T_3 form a partition of $Z_4(n)$ under the projection $\mathcal{F}$. One easily verifies that the following restrictions of Π are surjective with the specified covering degrees:

$$\begin{aligned} \Pi\big|_{S_1} &: Y_4(2n) \cap S_1 \xrightarrow{2:1} Z_4(2n) \cap \mathcal{F}(T_1), \\ \Pi\big|_{S_2} &: Y_4(2n) \cap S_2 \xrightarrow{2:1} Z_4(2n) \cap \mathcal{F}(T_2), \\ \Pi\big|_{S_3} &: Y_4(2n) \cap S_3 \xrightarrow{1:1} Z_4(2n) \cap \mathcal{F}(T_3). \end{aligned}$$

In addition, $\mathcal{F}$ restricts to the partition of $X_4(n)$ with the following covering degrees

$$\begin{aligned} \mathcal{F}\big|_{T_1} &: X_4(n) \cap T_1 \xrightarrow{2:1} Y_4(n) \cap \mathcal{F}(T_1), \\ \mathcal{F}\big|_{T_2} &: X_4(n) \cap T_2 \xrightarrow{2:1} Y_4(n) \cap \mathcal{F}(T_2), \\ \mathcal{F}\big|_{T_3} &: X_4(n) \cap T_3 \xrightarrow{1:1} Y_4(n) \cap \mathcal{F}(T_3), \end{aligned}$$

and

$$\begin{aligned} \xi_1 &: X_4(2n) \cap S_1 \xrightarrow{12:1} Y_4(2n) \cap S_1, \\ \xi_2 &: X_4(2n) \cap S_2 \xrightarrow{12:1} Y_4(2n) \cap S_2, \\ \xi_3 &: X_4(2n) \cap S_3 \xrightarrow{3:1} Y_4(2n) \cap S_3, \end{aligned}$$

where

$$\xi_1(x, -, -, -) = (x, \Diamond, z, \Box) \qquad \begin{pmatrix} \Diamond \text{ is the parity partner of } x; z \text{ is the first element, from left to right,} \\ \text{of } (x, -, -, -) \text{ of opposite parity to } x, \Box \text{ is the parity partner of } z. \end{pmatrix}$$

$$\xi_2(x, -, -, -) = \begin{cases} (x, x, z, w) & \text{if } x \text{ is repeated,} \\ (x, y, z, z) & \text{if } x \text{ is not repeated,} \end{cases}$$

$$\xi_3(x, -, -, -) = (x, x, z, z) \quad \text{where } x \text{ and } z \text{ are the distinct entries of } (x, -, -, -).$$

Counting degrees, we have

$$|X_4(2n) \cap S_i| = 3|X_4(n) \cap T_i|$$

for $i = 1, 2, 3$, and using the fact that the S_i and T_i partition $X_4(2n)$ and $X_4(n)$ respectively, it follows that $r_4(2n) = 3r_4(n)$. □

Now we are ready to prove Jacobi's Theorem. Lemma 3.3.3.2 and Lemma 3.3.3.3 imply that $r_4(2^k n) = 3r_4(n)$ for all odd n and $k \geq 1$. We proved above that $r_4(n) = 8\sigma_1(n)$ for odd n; thus,

$$r_4(n) = \begin{cases} 8\sigma_1(n) & n \text{ odd}, \\ 24 \sum\limits_{\substack{m|n \\ m \text{ odd}}} m & m \text{ odd} \end{cases}$$

for all $n \geq 1$. One easily checks that this implies

$$r_4(n) = 8 \sum_{\substack{m|n \\ 4\nmid m}} m,$$

as required.

Chapter 4

Gauss sums over $\mathbb{Q}$

In his treatise *Vorlesungen über die Theorie der algebraischen Zahlen* [Hec23], Hecke writes

> *Es ist die Tatsache, daß die genauere Kenntnis des Verhaltens einer analytischen Funktion in der Nähe ihrer singulären Stellen eine Quelle von arithmetischen Sätzen ist.*[6]

He had in mind two particular examples. The first was Dirichlet's class number formula (Proposition 1.3.4.3), expressing the residue at $s = 1$ of the Dedekind zeta function of an algebraic number field K in terms of arithmetic invariants of K: the number-theoretic theorem here is Dirichlet's theorem on the infinitude of primes in arithmetic progressions. In more recent times, far-reaching generalisations of the class number formula have been conjectured, relating special values of the L-functions attached to arithmetic schemes to intrinsic properties of those schemes [Nek91].

The second incarnation of Hecke' philosophy occurs in the study of the law of quadratic reciprocity, and can be thought of as a sort of automorphic analogue of the first example. It indicates that we should expect number-theoretic information to be contained in the asymptotic expansions of theta functions. The simplest possible illustration of the link between the two philosophies is based on the classical statement [Doe55, 6. Kapitel, §3, pp. 115] that the asymptotic expansion of a function at a singular point is determined by the residues of its Mellin transform. In particular, in a sense which will be made precise in Subsection 6.1, the asymptotic behaviour of Jacobi's theta function at each rational point on the real line is determined by the poles and residues of the Dirichlet L-functions.

However, in this chapter, following Cauchy [Cau40], Genocchi [Gen52] and Schaar [Sch50], we use the particularly simple Fourier expansion of the theta function to obtain the asymptotic expansion using Euler–Maclaurin summation. The number-theoretic theorem which follows from a careful investigation of the asymptotics of Jacobi's theta function manifests as an identity involving two finite sums, known as quadratic Gauss sums:

$$\frac{1}{\sqrt{a}} \sum_{n=0}^{a-1} \exp\left(\frac{2\pi i n^2 b}{a}\right) = \frac{\exp\left(\frac{\pi i}{4}\right)}{\sqrt{2b}} \sum_{n=0}^{2b-1} \exp\left(-\frac{\pi i n^2 a}{2b}\right).$$

This identity is called the Landsberg–Schaar relation. We will see that for particular choices of a and b, the sums involved "transform like quadratic residue symbols", allowing us to prove quadratic reciprocity as a special case of the relation.

If one varies the ground field over all algebraic number fields and replaces Jacobi's theta function with the theta function of the number field, one sees that the asymptotics of the theta function contain enough information to prove the law of quadratic reciprocity. This analytic proof of quadratic reciprocity is due to Hecke, and in Chapter 5 we sketch Hecke's generalisation of these ideas in the context of totally real number fields.

Since we don't require calculus to state the Landsberg–Schaar relation, we shouldn't require calculus to prove it. The theme of this chapter is to provide elementary proofs. Indeed, we give two elementary proofs of the Landsberg–Schaar relation, an elementary proof of a generalised version with linear terms, an elementary proof of a twisted version, and an elementary proof of a "local" version for fourth powers.

[6] The fact is that precise knowledge of the behaviour of an analytic function in the neighborhood of its singular points is a source of number-theoretic theorems. Translation by George U. Brauer and Jay R. Goldman with the assistance of R. Kotzen [Hec81].

In Section 4.1, we compute the asymptotic expansion of Jacobi's theta function, and use the transformation law (Theorem 2.2.1.1) to derive the Landsberg–Schaar relation. We then deduce the law of quadratic reciprocity as a corollary.

In Section 4.2, we deduce the full Landsberg–Schaar relation, by induction, from a special case first proved by Gauss. This argument, published recently by the author [Moo20], avoids any techniques from analysis and thus qualifies as an elementary proof of the Landsberg–Schaar relation.

Instead of employing induction, one may give a cleaner proof of the Landsberg–Schaar relation by evaluating both sides of the identity. Following [BEW98, Sections 1.4 and 1.5] we evaluate the required Gauss sums in Section 4.3 and present the proof in Section 4.4. In fact, we evaluate a few extra Gauss sums in order to prove — without analysis — a slightly more general version of the Landsberg–Schaar relation in 4.5. This identity was first proved[7] by Genocchi in 1852 [Gen52] and a shorter proof was later given by Mordell [Mor33]. We also present a proof using theta functions.

In Section 4.6, we investigate the asymptotic expansion of Jacobi's theta function twisted by Dirichlet characters. As one might expect, a twisted version of the Landsberg–Schaar relation appears, which was first proven in 1945 by Guinand [Gui45], and generalised by Berndt in 1972 [Ber73]. Using asymptotic expansions of theta functions, we provide a different analytic proof, and using the results of Section 4.5, we provide an elementary proof. Unexpectedly, the evaluation of the twisted Gauss sums is much more complicated, and is directly related to the theory of quartic Gauss sums in the case that the Dirichlet character is a Legendre symbol. In Section 4.8 we evaluate those twisted Gauss sums with the help of Matthews' evaluation of the quartic Gauss sum [Mat79b]. Using the twisted Landsberg–Schaar relation, we prove a sort of local quartic analogue of the Landsberg–Schaar relation, and using our elementary proof of Berndt's generalisation, we prove a "generalised" local quartic Landsberg–Schaar relation.

These methods also suffice to prove some rather less attractive local Landsberg–Schaar relations for sextic and octic higher-degree Gauss sums, but as the degree increases, so too do the sightings of extraneous sums which largely resist attempts at simplification.

Throughout this chapter, we use the notation and results of Subsection 1.3.3. We remind the reader that the term Gauss sum is used to refer to both *exponential sums*, which are generally shaped like

$$\sum_{n \bmod N} \exp\left(\frac{2\pi i n^d}{N}\right),$$

and to *character sums*, which are shaped like

$$\sum_{n \bmod N} \chi(n) \exp\left(\frac{2\pi i n}{N}\right),$$

where χ is a character, often primitive, of $(\mathbb{Z}/N\mathbb{Z})^\times$. We hope that it will be clear from the context which kind of Gauss sum we are considering, as both types will appear in this chapter.

On the subject of notation, the reader should rest assured that all square roots may be taken to have positive real part unless otherwise indicated; thus, we sometimes write $\sqrt{i}$ instead of $\exp\left(\frac{\pi i}{4}\right)$ for legibility.

4.1 The Landsberg–Schaar relation over $\mathbb{Q}$

In Subsection 4.1.1, we present the modern proof of the Landsberg–Schaar relation using theta functions. This proof is essentially due to Schaar [Sch50] and Landsberg [Lan93], but a common ingredient to almost all subsequent proofs (and generalisations to theta functions of number fields) is the exploitation of the particularly simple form of the Fourier coefficients of the relevant theta function in computing the asymptotic expansion. For our proof, we use a form of Euler–Maclaurin summation:

Proposition 4.1.0.1 (Euler 1738 [Eul38], Maclaurin 1742 [Mac42], Poisson 1893 [Poi93], Murty [Mur08])**.** *Let k be a non-negative integer and f be $(k+1)$-times differentiable on $[a,b]$, with $a,b \in \mathbb{Z}$. Then*

$$\sum_{a<n\leq b} f(n) = \int_a^b f(x)dx + \sum_{r=0}^{k} \frac{(-1)^{r+1}B_{r+1}(0)}{(r+1)!}(f^{(r)}(b) - f^{(r)}(a)) + \frac{(-1)^k}{(k+1)!}\int_a^b B_{k+1}(x) f^{(k+1)}(x)dx, \quad (4.1)$$

[7] Using the theory of residues, Genocchi attempted a generalisation [Gen53], but Lindelöf claimed that the proof was flawed [Lin47].

where $B_r(x)$ is the 1-periodic extension of the restriction to $[0,1)$ of the rth Bernouilli polynomial $b_r(x)$ defined by

$$\sum_{r=0}^{\infty} b_r(x)\frac{t^r}{r!} = \frac{te^{xt}}{e^t-1}.$$

Corollary 4.1.0.2. *Suppose that f is a Schwartz function (see Definition 1.2.2.3). Then we may take the limit as $a \to -\infty$ and $b \to \infty$ in 4.1, and obtain the following, valid for all non-negative integers k:*

$$\sum_{n=-\infty}^{\infty} f(n) = \int_{\infty}^{\infty} f(x)dx + \frac{(-1)^k}{(k+1)!}\int_{-\infty}^{\infty} B_{k+1}(x)f^{(k+1)}(x)dx.$$

For a proof of Proposition 4.1.0.1, we refer the reader elsewhere [Mur08, Theorem 5.1.3]. In Subsection 6.1.2, we give a proof of 4.2.1 which does not depend on the particularly simple form of the Fourier coefficients of the theta function, but instead makes use of the correspondence with L-functions alluded in the introduction to this section.

In Subsection 4.1.2, we use the Landsberg–Schaar relation to prove the law of quadratic reciprocity. In the course of the proof, we rely on some identities for Gauss sums which are proved below in Section 4.3.

4.1.1 The asymptotic expansion of Jacobi's theta function

As in Chapter 2, we define

$$\theta(z) = \sum_{n\in\mathbb{Z}} e^{2\pi i n^2 z} = 1 + 2\sum_{n=1}^{\infty} e^{2\pi i n^2 z}.$$

Then for $a, b \in \mathbb{Z}$, $a > 0$, $\mathrm{Re}(\epsilon) > 0$,

$$\theta\left(\frac{b}{a} + i\epsilon\right) = \sum_{n=0}^{a-1} e^{2\pi i n^2 b/a} \sum_{m=n \bmod a} e^{-2\pi m^2 \epsilon}. \tag{4.2}$$

We wish to compute the asymptotic expansion of the inner sum as $\epsilon \to 0$. Notice that we may rewrite the inner sum as

$$\sum_{t\in\mathbb{Z}} e^{-2\pi(n+at)^2\epsilon},$$

where the summand is a Schwartz function of t, with t considered as a real variable. Therefore we may apply Corollary 4.1.0.2 to compute the asymptotic expansion.

Proposition 4.1.1.1. *For $a, b \in \mathbb{Z}$, $a > 0$ and $\mathrm{Re}(\epsilon) > 0$,*

$$\theta\left(\frac{b}{a} + i\epsilon\right) \sim \left(\frac{1}{a\sqrt{2}}\sum_{n=0}^{a-1}\exp\left(\frac{2\pi i n^2 b}{a}\right)\right)\epsilon^{-1/2} + O(|\epsilon|^N) \tag{4.3}$$

as $\epsilon \to 0$, for all $N > 0$.

Proof. Elementary calculus suffices to prove that

$$\int_{-\infty}^{\infty} e^{-2\pi(n+at)^2\epsilon}dt = \frac{1}{a\sqrt{2\epsilon}} \tag{4.4}$$

for real positive ϵ, and by standard arguments from complex analysis, 4.4 holds for complex ϵ with $\mathrm{Re}(\epsilon) > 0$ (the branch cut for the square root is taken along the non-positive real axis). By 4.2, we need to show that as $\epsilon \to 0$,

$$\int_{-\infty}^{\infty} B_k(t)\frac{d^k}{dt^k}\left(e^{-2\pi(n+at)^2\epsilon}\right)dt \sim O(|\epsilon|^{\frac{k-1}{2}}). \tag{4.5}$$

Note that there is some constant $C_{0,k}$ such that $|B_k(t)| \le C_{0,k}$ for all real t. Consequently, upon making the substitution $u = n + at$,

$$\left|\int_{-\infty}^{\infty} B_k(t)\frac{d^k}{dt^k}\left(e^{-2\pi(n+at)^2\epsilon}\right)dt\right| \le C_{0,k}C_{1,k}\int_{-\infty}^{\infty}\left|\frac{d^k}{du^k}\left(e^{-2\pi u^2\epsilon}\right)\right|du,$$

where $C_{1,k}$ depends on n and a. Making the substitution $v = \sqrt{\epsilon}u$ (with the obvious branch cut, as above), then using Cauchy's integral theorem together with the fact that the integrand is Schwartz,

$$\int_{-\infty}^{\infty} \left| \frac{d^k}{du^k} \left(e^{-2\pi u^2 \epsilon} \right) \right| du \leq |\epsilon|^{\frac{k-1}{2}} \int_{-\infty}^{\infty} \left| \frac{d^k}{dv^k} \left(e^{-2\pi v^2} \right) \right| dv,$$

and 4.5 is proved. □

Remark 4.1.1.2. *One easily obtains, with the same hypotheses as in Proposition 4.1.1.1,*

$$\theta\left(\frac{b}{a} + i\epsilon\right) \sim \left(\frac{1}{a\sqrt{2}} \sum_{n=0}^{a-1} \exp\left(\frac{2\pi i n^2 b}{a} \right) \right) \epsilon^{-1/2} + O(1)$$

by using Corollary 4.1.0.2 with $k = 0$, noting that $b_1(t) = t - 1/2$, so $|B_1(t)| \leq 1/2$, and making the estimate

$$\int_{-\infty}^{\infty} \left| \frac{d}{dt} \left(e^{-2\pi (n+at)^2 \epsilon} \right) \right| dt = 2 \int_{0}^{\infty} \frac{d}{dt} \left(e^{-2\pi (n+at)^2 \epsilon} \right) dt < 2.$$

We only require this weaker version of Proposition 4.1.1.1 in what follows. The purpose of Proposition 4.1.1.1 is to reassure the reader that the theta function carries no other information near the rational boundary points. Indeed, we expect "by pure thought" that Proposition 4.1.1.1 is true a priori because the Riemann zeta function has only one singularity in the complex plane: see Section 6.1.

4.1.2 The Landsberg–Schaar relation and quadratic reciprocity

We now use the asymptotic expansion of Jacobi's theta function, together with the functional equation

$$\theta(-1/4z) = \sqrt{-2iz}\,\theta(z) \tag{4.6}$$

from Theorem 2.2.1.1, to derive the Landsberg–Schaar relation. We assume that a and b are positive integers throughout this subsection. Following Murty and Pacelli [MP17],

$$-\frac{1}{4\left(\frac{b}{a} + i\epsilon\right)} = -\frac{a}{4b} + i\tau,$$

where $\mathrm{Re}(\tau) > 0$ if and only if $\mathrm{Re}(\epsilon) > 0$, $\tau \to 0^+$ if and only if $\epsilon \to 0^+$, and

$$\frac{\tau}{\epsilon} = \frac{a}{4b\left(\frac{b}{a} + i\epsilon\right)}. \tag{4.7}$$

By Proposition 4.1.1,

$$\theta\left(-\frac{a}{4b} + i\tau\right) \sim \left(\frac{1}{4b\sqrt{2}} \sum_{n=0}^{4b-1} \exp\left(-\frac{\pi i n^2 a}{b} \right) \right) \tau^{-1/2} + O(|\tau|), \tag{4.8}$$

and by the functional equation 4.6 (with $z = \frac{b}{a} + i\epsilon$),

$$\lim_{\tau \to 0} \sqrt{\tau}\,\theta\left(-\frac{a}{4b} + i\tau\right) = \lim_{\substack{\tau \to 0 \\ \epsilon \to 0}} \sqrt{\frac{\tau}{\epsilon}} \sqrt{-2i\left(\frac{b}{a} + i\epsilon\right)} \sqrt{\epsilon}\,\theta\left(\frac{b}{a} + i\epsilon\right).$$

Using the equations 4.3, 4.7 and 4.8, this last equation simplifies to

$$\frac{1}{2b\sqrt{2}} \sum_{n=0}^{2b-1} \exp\left(-\frac{\pi i n^2 a}{2b} \right) = \frac{a}{2b} \sqrt{-\frac{2bi}{a}} \frac{1}{a\sqrt{2}} \sum_{n=0}^{a-1} \exp\left(-\frac{2\pi i n^2 b}{a} \right),$$

and when the dust settles we have the Landsberg–Schaar relation:

$$\frac{1}{\sqrt{a}} \sum_{n=0}^{a-1} \exp\left(\frac{2\pi i n^2 b}{a} \right) = \frac{\exp\left(\frac{\pi i}{4}\right)}{\sqrt{2b}} \sum_{n=0}^{2b-1} \exp\left(-\frac{\pi i n^2 a}{2b} \right). \tag{4.9}$$

Now we prove the main law of quadratic reciprocity as well as the two supplementary laws. The law of quadratic reciprocity is a central result in the field of algebraic number theory, and the quest to find higher degree versions, valid over number fields, was a driving force behind the creation of some of the most beautiful fields of modern mathematics. We, however, require only the simplest version, conjectured by Euler, attempted by Lagrange and first proved by Gauss, the details of which we briefly sketch below.

Let p be an odd prime and n an integer. We recall the definition of the Legendre symbol, first defined in Subsection 1.3.2:

$$\left(\frac{n}{p}\right) = \begin{cases} 0 & p \mid n, \\ 1 & (n,p)=1 \text{ and } n = x^2 \bmod p \text{ for some integer } x, \\ -1 & (n,p)=1 \text{ and } n \neq x^2 \bmod p \text{ for any integer } x. \end{cases}$$

We now state the law of quadratic reciprocity over $\mathbb{Q}$.

Theorem 4.1.2.1 (Gauss, [Gau01a; Gau01b; Gau08; Gau11])**.** *Let p and q be distinct odd primes. Then*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{(p-1)(q-1)}{4}}. \tag{4.10}$$

We also have the two supplementary laws*:*

$$\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}, \tag{4.11}$$

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}. \tag{4.12}$$

Now we come to the point of this subsection: to prove Theorem 4.1.2.1 using the Landsberg–Schaar relation.

Proof. We deal with the main law, 4.10, first. The first item is to note the following fundamental connection between the Legendre symbol and Gauss sums, for which we let p be an odd prime, $(m,p)=1$ and $\mu m = 1 \bmod p$:

$$\begin{aligned} \sum_{n=0}^{p-1} \exp\left(\frac{2\pi i n^2 m}{p}\right) &= 1 + \frac{1}{2} \sum_{\substack{n=1 \\ n=\square \bmod p}}^{p-1} \exp\left(\frac{2\pi i n m}{p}\right) \\ &= 1 + \sum_{n=1}^{p-1} \exp\left(\frac{2\pi i n m}{p}\right) + \sum_{n=1}^{p-1} \left(\frac{n}{p}\right) \exp\left(\frac{2\pi i n m}{p}\right) \\ &= \left(\frac{\mu}{p}\right) \sum_{n=1}^{p-1} \left(\frac{n}{p}\right) \exp\left(\frac{2\pi i n}{p}\right) \\ &= \left(\frac{m}{p}\right) \sum_{n=0}^{p-1} \exp\left(\frac{2\pi i n^2}{p}\right). \end{aligned} \tag{4.13}$$

The second item we require is the "product rule" for Gauss sums. This result generalises readily for the higher-degree Gauss sums of Section 4.8. Let a and b be positive and coprime. Then as s runs from 0 to $b-1$ modulo b and t runs from 0 to $a-1$ modulo a, $as+bt$ runs from 0 to $ab-1$ modulo ab, so

$$\sum_{n=0}^{ab-1} \exp\left(\frac{2\pi i n^2}{ab}\right) = \sum_{t=0}^{a-1} \exp\left(\frac{2\pi i t^2 b}{a}\right) \sum_{s=0}^{b-1} \exp\left(\frac{2\pi i s^2 a}{b}\right). \tag{4.14}$$

We now use 4.13 and 4.14 to convert the Legendre symbols appearing in the main law of quadratic reciprocity into Gauss sums, which we evaluate using the Landsberg–Schaar relation. Indeed, taking $a=p$ and $b=1$ in the Landsberg–Schaar relation 4.9, we have

$$\sum_{n=0}^{p-1} \exp\left(\frac{2\pi i n^2}{p}\right) = \sqrt{p}\epsilon_p,$$

where

$$\epsilon_p = \begin{cases} 1 & p = 1 \bmod 4, \\ i & p = 3 \bmod 4. \end{cases}$$

Proceeding,

$$\begin{aligned}
\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) &= \frac{\sum_{n=0}^{q-1}\exp\left(\frac{2\pi i n^2 p}{q}\right)\sum_{n=0}^{p-1}\exp\left(\frac{2\pi i n^2 q}{p}\right)}{\sum_{n=0}^{q-1}\exp\left(\frac{2\pi i n^2}{q}\right)\sum_{n=0}^{p-1}\exp\left(\frac{2\pi i n^2}{p}\right)} \\
&= \frac{\sum_{n=0}^{pq-1}\exp\left(\frac{2\pi i n^2}{pq}\right)}{\sum_{n=0}^{q-1}\exp\left(\frac{2\pi i n^2}{q}\right)\sum_{n=0}^{p-1}\exp\left(\frac{2\pi i n^2}{p}\right)} \\
&= \frac{\epsilon_{pq}}{\epsilon_p\epsilon_q} \\
&= (-1)^{\frac{(p-1)(q-1)}{4}}.
\end{aligned}$$

Now we deal with the first supplementary law, 4.11. By 4.13,

$$\sum_{n=0}^{p-1}\exp\left(-\frac{2\pi i n^2}{p}\right) = \left(\frac{-1}{p}\right)\sum_{n=0}^{p-1}\exp\left(\frac{2\pi i n^2}{p}\right). \tag{4.15}$$

We will be able to use the Landsberg–Schaar relation to evaluate the sum appearing on the right hand side of 4.15, so we investigate the product of the sum on the left with the sum on the right:

$$\begin{aligned}
\sum_{n=0}^{p-1}\exp\left(-\frac{2\pi i n^2}{p}\right)\sum_{m=0}^{p-1}\exp\left(\frac{2\pi i m^2}{p}\right) &= \sum_{n=0}^{p-1}\sum_{m=0}^{p-1}\exp\left(\frac{2\pi i(m^2-n^2)}{p}\right) \\
&= \sum_{n=0}^{p-1}\sum_{m=0}^{p-1}\exp\left(\frac{2\pi i((m+n)^2-n^2)}{p}\right) \\
&= \sum_{m=0}^{p-1}\exp\left(\frac{2\pi i m^2}{p}\right)\sum_{n=0}^{p-1}\exp\left(\frac{4\pi i m n}{p}\right).
\end{aligned}$$

The inner sum is a geometric series and vanishes unless p divides $2m$. So only the $m=0$ term of the outer sum contributes, and we are left with

$$\sum_{n=0}^{p-1}\exp\left(-\frac{2\pi i n^2}{p}\right)\sum_{m=0}^{p-1}\exp\left(\frac{2\pi i m^2}{p}\right) = p.$$

Returning to 4.15, we find that

$$p = \left(\frac{-1}{p}\right)\left(\sum_{n=0}^{p-1}\exp\left(\frac{2\pi i n^2}{p}\right)\right)^2,$$

and employing the Landsberg–Schaar relation,

$$1 = \left(\frac{-1}{p}\right)\epsilon_p^2,$$

so the formula claimed for $\left(\frac{-1}{p}\right)$ follows immediately.

Finally, we deal with the second supplementary law, 4.12. As with the first supplementary law, the connection with Gauss sums is facilitated by 4.13,

$$\sum_{n=0}^{p-1}\exp\left(\frac{4\pi i n^2}{p}\right) = \left(\frac{2}{p}\right)\sum_{n=0}^{p-1}\exp\left(\frac{2\pi i n^2}{p}\right).$$

However, in this case we may use the Landsberg–Schaar relation to evaluate both Gauss sums right away:

$$\begin{aligned}\left(\frac{2}{p}\right) &= \frac{1}{2}\epsilon_p^{-1}\exp\left(\frac{\pi i}{4}\right)\left(1+\exp\left(-\frac{\pi i p}{4}\right)-1+\exp\left(-\frac{9\pi i p}{4}\right)\right)\\ &= \epsilon_p^{-1}\exp\left(\frac{\pi i}{4}\right)\exp\left(-\frac{\pi i p}{4}\right).\end{aligned}$$

If $p = 1$ or $7 \bmod 8$, then $\left(\frac{2}{p}\right) = 1$, and if $p = 3$ or $5 \bmod 8$, then $\left(\frac{2}{p}\right) = -1$, as required. □

4.2 An elementary proof of the Landsberg–Schaar relation by induction

Statement of Authorship

Title of Paper	A proof of the Landsberg-Schaar relation by finite methods
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	Published in The Ramanujan Journal, 14th January 2020; received 3rd January 2019

Principal Author

Name of Principal Author (Candidate)	Benjamin Moore		
Contribution to the Paper	Primary author		
Overall percentage (%)	100		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	28/02/2020

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author			
Contribution to the Paper			
Signature		Date	

Name of Co-Author			
Contribution to the Paper			
Signature		Date	

Please cut and paste additional co-author panels here as required.

A proof of the Landsberg–Schaar relation by finite methods

Ben Moore*

Published in The Ramanujan Journal, 14th January 2020; received 3rd January 2019.

Abstract

The Landsberg–Schaar relation is a classical identity between quadratic Gauss sums, often used as a stepping stone to prove the law of quadratic reciprocity. The Landsberg–Schaar relation itself is usually proved by carefully taking a limit in the functional equation for Jacobi's theta function. In this article we present a direct proof, avoiding any analysis.

4.2.1 Introduction

The aim of this article is to prove, using only techniques of elementary number theory, the Landsberg–Schaar relation for positive integral a and b:

$$\frac{1}{\sqrt{a}}\sum_{n=0}^{a-1}\exp\left(\frac{2\pi i n^2 b}{a}\right) = \frac{1}{\sqrt{2b}}\exp\left(\frac{\pi i}{4}\right)\sum_{n=0}^{2b-1}\exp\left(-\frac{\pi i n^2 a}{2b}\right).$$

This relation was first discovered in 1850 by Mathias Schaar [Sch50], who proved it using the Poisson summation formula, and proceeded to derive from it the law of quadratic reciprocity. In 1893 Georg Landsberg, apparently unaware of Schaar's work, rediscovered a slightly more general version of the relation [Lan93]. Although Landsberg emphasises the role of modular transformations, his proof is closer in spirit to the modern one in [MP05], in which one takes a limit of the functional equation for Jacobi's theta function towards rational points on the real line.

A few remarks are in order concerning some closely related results involving techniques differing from those in the present article. Firstly, whilst this article was under review, the author noticed that in the article [BS13], the authors prove Hecke's generalisation of the Landsberg–Schaar identity over number fields. Their argument is elementary except for an appeal to Milgram's formula, which allows for the evaluation of exponential sums over non-degenerate integer-valued symmetric bilinear forms: the cited proof [HM73, p. 127–131] uses Fourier analysis. When the number field is $\mathbb{Q}$, we recover the Landsberg–Schaar relation, and Milgram's formula in this instance is essentially Lemma 4.2.1.1 below.

Secondly, the authors of [BS13] suggest, in a parenthetical remark on the second page, that it does not seem possible to prove Hecke reciprocity by explicitly evaluating both sides. However, this does appear to work in the case of the Landsberg–Schaar relation. Indeed, in their book on Gauss and Jacobi sums [BEW98, Theorems 1.51, 1.52 and 1.54], Berndt, Evans and Williams give the first elementary evaluation of

$$\phi(a,b) := \sum_{n=0}^{a-1}\exp\left(\frac{2\pi i n^2 b}{a}\right)$$

for coprime positive integral a and b. This evaluation is deduced from Estermann's elementary evaluation [Est45] of

$$\phi(a,1) := \sum_{n=0}^{a-1}\exp\left(\frac{2\pi i n^2}{a}\right)$$

for an odd positive integer a. The equality

$$\frac{1}{\sqrt{a}}\phi(a,b) = \frac{1}{2\sqrt{2b}}\exp\left(\frac{\pi i}{4}\right)\phi(4b,-a)$$

follows by evaluating both sides using the elementary evaluation given by Berndt, Evans and Williams. This equality is precisely the Landsberg–Schaar relation. In this argument, the hard work is contained in the evaluation of $\phi(a,b)$.

*School of Mathematical Sciences, University of Adelaide, SA, Australia. Email: benjamin.moore@adelaide.edu.au

To emphasise the fact that the Landsberg–Schaar relation is an identity between Gauss sums, and to simplify the notation, we define, for a and b integers with $a > 0$,

$$\Phi(a,b) = \frac{1}{\sqrt{a}} \sum_{n=0}^{a-1} \exp\left(\frac{\pi i n^2 b}{a}\right).$$

Then the Landsberg–Schaar relation, for positive integral a and b, takes the form

$$\Phi(a,2b) = \sqrt{i}\Phi(2b,-a).$$

The starting point for our proof is the following evaluation of a quadratic Gauss sum, given by Gauss in 1811 [Gau11].

Lemma 4.2.1.1. *Let a be an integer, $a \geq 1$. Then:*

$$\Phi(a,2) = \begin{cases} 1+i & a = 0 \bmod 4 \\ 1 & a = 1 \bmod 4 \\ 0 & a = 2 \bmod 4 \\ i & a = 3 \bmod 4. \end{cases}$$

A proof of Lemma 4.2.1.1 avoiding analytical techniques may be given using linear algebra [MP17]. Stronger results, which imply Lemma 4.2.1.1 (and Propositions 4.2.3.1 and 4.2.3.2 below), are also proved using elementary methods in [BEW98, Sections 1.3 and 1.5].

One may easily check that Lemma 4.2.1.1 is exactly the Landsberg–Schaar relation for $b = 1$. Our aim is to prove the Landsberg–Schaar relation in general by induction on the number of distinct prime factors of b. The induction step follows from the next three results, and the bulk of this article is spent proving the third.

Lemma 4.2.1.2. *Let a, b and l be integers, a positive and $(a,b) = 1$. Then:*

$$\Phi(ab,l) = \Phi(a,bl)\Phi(b,al).$$

The proof is not difficult, but is hard to find in this form: usually l is assumed to be even, which simplifies matters considerably.

Proof. As s runs from 0 to $b-1$ and t runs from 0 to $a-1$, $as+bt$ runs through a complete system of representatives for elements of $\mathbb{Z}/ab\mathbb{Z}$. So

$$(as+bt)^2 = g^2 + 2gkab + k^2a^2b^2$$

for $k = 0$ or 1, $0 \leq g < ab$. It follows that:

$$\Phi(ab,l) = \frac{1}{\sqrt{ab}} \sum_{n=0}^{ab-1} \exp\left(\frac{\pi i n^2 l}{ab}\right) = \epsilon\Phi(a,bl)\Phi(b,al),$$

where

$$\epsilon = \begin{cases} 1 & a \text{ or } b \text{ even} \\ (-1)^S & a \text{ and } b \text{ both odd,} \end{cases}$$

and

$$S = \#\{(s,t) \mid as+bt > ab\}.$$

The value of S is $\frac{(a-1)(b-1)}{2}$ – the problem of determining S was set as a puzzle by Sylvester in [Syl84] and solved by W. J. Curran Sharp in the same volume. The solution runs as follows: define

$$P(x) = (1 + x^b + x^{2b} + \cdots + x^{ab})(1 + x^a + x^{2a} + \cdots + x^{ba}),$$

and note that

$$P(x) = 1 + \cdots + 2x^{ab} + \cdots + x^{2ab},$$

where the first dots comprise one term x^g for each g of the form $as + tb$ (we know that the coefficient of x^g is 1 since a and b are coprime).

Since each factor of P is a palindromic polynomial, so too is P, and it follows that the second dots comprise the same number of terms, all of coefficient 1. Therefore,

$$(1 + a)(1 + b) = P(1) = 4 + 2\#\{g < ab \mid g = as + tb\}.$$

Using the fact that

$$\#\{g < ab \mid g = as + tb\} = (ab - 1) - S,$$

the claim follows. So if a and b are both odd, then S is even, and $\epsilon = 1$ in this case too. □

The following result will not be needed until Subsection 4.2.4.

Lemma 4.2.1.3. *Suppose a, b and k are nonzero integers, a and k are positive, and at least one of a or b is even. Then*

$$\Phi(ka, kb) = \sqrt{k}\Phi(a, b).$$

Proof. This is a straightforward calculation.

$$\begin{aligned}
\Phi(ka, kb) &= \frac{1}{\sqrt{ka}} \sum_{n=0}^{ka-1} \exp\left(\frac{\pi i n^2 b}{a}\right) \\
&= \frac{1}{\sqrt{k}} \sum_{m=0}^{k-1} \frac{1}{\sqrt{a}} \sum_{n=0}^{a-1} \exp\left(\frac{\pi i (n + am)^2 b}{a}\right) \\
&= \frac{1}{\sqrt{k}} \sum_{m=0}^{k-1} \exp(\pi i a b m^2) \Phi(a, b) \\
&= \sqrt{k}\Phi(a, b).
\end{aligned}$$

□

At this point, we only need one more result to prove the Landsberg–Schaar relation in Subsection 4.2.4.

Proposition 4.2.1.4. *Let p be a prime and l an integer with $(p, l) = 1$. Then:*

$$\Phi(p^k, 2l)\Phi(p^k, -2l) = \begin{cases} 1 & p \text{ an odd prime, } k \geq 1 \\ 2 & p = 2,\ k \geq 3. \end{cases}$$

The next two subsections are devoted to proving Proposition 4.2.1.4, which is achieved by computing $\Phi(p^k, 2l)$ directly. All the results of the next two subsections are well-known in the literature, though apparently not all collected in one place. In particular, Proposition 4.2.3.1 and Proposition 4.2.3.2 are special cases of Gauss' evaluation of $\Phi(a, 2)$, and may be found in [BEW98] as mentioned above. The proof of each proposition requires one to know the number of solutions to $x^2 = a \bmod p^k$ for each a, which is the subject of the next subsection.

4.2.2 Counting solutions to $x^2 = a \bmod p^k$

The first result is reminiscent of Hensel's lemma, but is more direct.

Lemma 4.2.2.1. *Let p be a prime, not necessarily odd, and $j > i$.*

$$\#\{x \mid x^2 = kp^i \bmod p^j\} = \begin{cases} p^{i/2}\#\{x | x^2 = k \bmod p^{j-i}\} & i \text{ even} \\ 0 & i \text{ odd, } (k, p) = 1. \end{cases}$$

Proof. To dispose of the case where i is odd, note that

$$x^2 = kp^i + lp^j$$

implies p^i divides x, so p divides k. Now suppose that i is even. Define

$$A = \{z \in \mathbb{Z}/p^j\mathbb{Z} \mid z^2 = kp^i \bmod p^j\}$$
$$B = \{z \in \mathbb{Z}/p^{j-i}\mathbb{Z} \mid z^2 = k \bmod p^{j-i}\}.$$

The map $F : B \to A$ by $x \mapsto p^{i/2}x$ is surjective, so to prove that $|A| = p^{i/2}|B|$, we need only show that each fibre of F has cardinality $p^{i/2}$. Since $F(x) = F(y)$ implies $x = y + tp^{j-i/2}$, the fibre of $F(x)$ contains nothing more than the elements $x_t = x + tp^{j-i/2} \in \mathbb{Z}/p^{j-i}\mathbb{Z}$ for $t = 0, \ldots, p^{i/2} - 1$. But the x_t are contained in B:

$$(x_t)^2 = x^2 + 2xtp^{j-i/2} + t^2p^{2j-i} = x^2 \bmod p^j,$$

so the fibre of $F(x)$ is exactly the x_t. □

We can now count the solutions to $x^2 = 0 \bmod p^k$.

Lemma 4.2.2.2.

$$\#\{x | x^2 = 0 \bmod p^k\} = \begin{cases} p^{k/2} & k \textit{ even} \\ p^{(k-1)/2} & k \textit{ odd.} \end{cases}$$

Proof. For k even, put $j = k$, $i = k - 2$ in Lemma 4.2.2.1. Then

$$\#\{x | x^2 = 0 \bmod p^k\} = p^{(k-2)/2}\#\{x | x^2 = 0 \bmod p^2\},$$

and

$$\{x | x^2 = 0 \bmod p^2\} = \{0, p, 2p, \ldots, (p-1)p\}.$$

For k odd, put $j = k$, $i = k - 1$ in Lemma 4.2.2.1 to obtain

$$\#\{x | x^2 = 0 \bmod p^k\} = p^{(k-1)/2}\#\{x | x^2 = 0 \bmod p\} = p^{(k-1)/2}.$$ □

The next two results are standard: one may consult Hecke [Hec81, p. 47, Theorems 46a and 47] or Dickson [Dic57, p. 13, Theorem 17].

Lemma 4.2.2.3. *Let p be an odd prime, $j \geq 1$, $(k, p) = 1$, and write $\left(\frac{k}{p}\right)$ for the Legendre symbol. Then*

$$\#\{x \mid x^2 = k \bmod p^j\} = 1 + \left(\frac{k}{p}\right).$$

Lemma 4.2.2.4. *For $p = 2$ and $(k, 2) = 1$:*

$$\#\{x \mid x^2 = k \bmod 4\} = \begin{cases} 2 & k = 1 \bmod 4 \\ 0 & k = 3 \bmod 4. \end{cases}$$

For $j \geq 3$:

$$\#\{x \mid x^2 = k \bmod 2^j\} = \begin{cases} 4 & k = 1 \bmod 8 \\ 0 & \textit{otherwise.} \end{cases}$$

Lemma 4.2.2.1 and Lemma 4.2.2.3 taken together give us a complete picture for odd p when $k \neq 0$, as follows.

Lemma 4.2.2.5. *For $(k, p) = 1$, $j > i$ and p an odd prime:*

$$\#\{x \mid x^2 = kp^i \bmod p^j\} = \begin{cases} p^{i/2}\left(1 + \left(\frac{k}{p}\right)\right) & i \textit{ even} \\ 0 & i \textit{ odd.} \end{cases}$$

The analogue of Lemma 4.2.2.5 for $p = 2$ follows from Lemma 4.2.2.1 and Lemma 4.2.2.4. Since the exceptional cases $j = 1$ and $j = 2$ can be done by hand, we only need to consider $j - i \geq 3$.

Lemma 4.2.2.6. *For $p = 2$, $(k, 2) = 1$, $j \geq i + 3$ and i even:*

$$\#\{x \mid x^2 = 2^ik \bmod 2^j\} = \begin{cases} 4p^{i/2} & k = 1 \bmod 8 \\ 0 & \textit{otherwise.} \end{cases}$$

For i odd (and all other hypotheses unchanged):

$$\#\{x \mid x^2 = 2^ik \bmod 2^j\} = 0.$$

4.2.3 Evaluation of $\Phi(p^k, 2l)$

This subsection is devoted to evaluating $\Phi(p^k, 2l)$ for p prime and $(l, p) = 1$. We first evaluate $\Phi(p^k, 2l)$ in the case that p is an odd prime, then we proceed to the exceptional case of $p = 2$. The idea of each proof is to expand $\Phi(p^k, 2l)$ as a finite Fourier series. The coefficients have been calculated in Subsection 4.2.2, and substituting in these expressions and simplifying yields the claimed results. In these calculations we implicitly make use of Lemma 4.2.2.5, Lemma 4.2.2.6 and Lemma 4.2.2.2. We conclude this subsection with the proof of Proposition 4.2.1.4.

Proposition 4.2.3.1. *Let p be an odd prime and $(l, p) = 1$. Then:*

$$\Phi(p^k, 2l) = \begin{cases} 1 & k \text{ even, } k \geq 2 \\ \left(\frac{l}{p}\right)\Phi(p^k, 2) & k \text{ odd.} \end{cases}$$

Proof. First we treat the case of k even.

$$\begin{aligned}
\Phi(p^k, 2l) &= p^{-k/2} \sum_{n=0}^{p^k-1} \#\{x | x^2 = n \bmod p^k\} \exp\left(\frac{2\pi i n l}{p^k}\right) \\
&= p^{-k/2} \left[\sum_{n=0}^{p^k-1} \left(1 + \left(\frac{n}{p}\right)\right) \exp\left(\frac{2\pi i n l}{p^k}\right) + (p^{k/2} - 1) - \sum_{n=1}^{p^{k-1}-1} \exp\left(\frac{2\pi i n p l}{p^k}\right) \right. \\
&\qquad \left. + \sum_{i=2,4,\ldots,k-2} \sum_{\substack{n=1 \\ (n,p^k)=p^i \\ n/p^i=m}}^{p^{k-i}-1} \#\{x \mid x^2 = mp^i \bmod p^k\} \exp\left(\frac{2\pi i m l}{p^{k-i}}\right) \right].
\end{aligned}$$

We should explain each term in the last two lines: the first term gives the correct coefficients for $(n, p) = 1$, the second term makes the correct contribution for $n = 0$, the third term makes the nth coefficient 0 for any nonzero n divisible by p, and the final term restores the correct coefficient for these n. Note that the Legendre symbol $\left(\frac{n}{p}\right)$ is defined to be 0 if n is divisible by p – this implies that $\left(\frac{\cdot}{p}\right)$ is multiplicative.

The inner sum in the last term can be simplified:

$$\begin{aligned}
&\sum_{\substack{n=1 \\ (n,p^k)=p^i \\ n/p^i=m}}^{p^{k-i}-1} \#\{x \mid x^2 = mp^i \bmod p^k\} \exp\left(\frac{2\pi i m l}{p^{k-i}}\right) \\
&\qquad = p^{i/2} \left[\sum_{m=1}^{p^{k-i}-1} \left(1 + \left(\frac{m}{p}\right)\right) \exp\left(\frac{2\pi i m l}{p^{k-i}}\right) - \sum_{m=1}^{p^{k-i-1}-1} \left(1 + \left(\frac{mp}{p}\right)\right) \exp\left(\frac{2\pi i m l}{p^{k-i-1}}\right) \right] \\
&\qquad = p^{i/2} \sum_{m=1}^{p^{k-i}-1} \left(\frac{m}{p}\right) \exp\left(\frac{2\pi i m l}{p^{k-i}}\right) = p^{i/2} \left(\frac{l}{p}\right) \sum_{m=0}^{p^{k-i}-1} \left(\frac{m}{p}\right) \exp\left(\frac{2\pi i m}{p^{k-i}}\right).
\end{aligned}$$

The final equality above follows from the facts that $\left(\frac{\cdot}{p}\right)$ is multiplicative, and that $(l, p) = 1$ implies that as m runs from 0 to $p^{k-i} - 1$, so does $lm \bmod p^{k-i}$. But this last sum is zero, since $i \leq k - 2$ implies $k - i > 1$, and for $r > 1$,

$\sum_{m=0}^{p^r-1} \left(\frac{m}{p}\right) \exp\left(\frac{2\pi i m}{p^r}\right) = 0$ as follows:

$$\begin{aligned}\sum_{m=0}^{p^r-1} \left(\frac{m}{p}\right) \exp\left(\frac{2\pi i m}{p^r}\right) &= \sum_{\alpha=0}^{p^{r-1}-1} \sum_{m=\alpha p}^{(\alpha+1)p-1} \left(\frac{m}{p}\right) \exp\left(\frac{2\pi i m}{p^r}\right) \\ &= \sum_{\alpha=0}^{p^{r-1}-1} \sum_{m=0}^{p-1} \left(\frac{m+\alpha p}{p}\right) \exp\left(\frac{2\pi i (m+\alpha p)}{p^r}\right) \\ &= \sum_{\alpha=0}^{p^{r-1}-1} \exp\left(\frac{2\pi i \alpha}{p^{r-1}}\right) \sum_{m=0}^{p-1} \left(\frac{m}{p}\right) \exp\left(\frac{2\pi i m}{p^r}\right) \\ &= 0.\end{aligned}$$

Therefore, the last term in the expansion of $\Phi(p^k, 2l)$ vanishes. When we expand the factor $(1 + \left(\frac{n}{p}\right))$ multiplying the first term, we find that the sum multiplied by 1 is a geometric series, so it vanishes, and the sum multiplied $\left(\frac{n}{p}\right)$ is zero by the calculation above. The -1 in the second term combines with the third term to give another geometric series, so we are left with $\Phi(p^k, 2l) = 1$, as promised.

Now we treat odd k, and suppose for the moment that $k > 1$. Then the coefficients are very similar, apart from the contributions for $n = 0$ and $n = mp^{k-1}$ with $(m, p) = 0$. Specifically,

$$\begin{aligned}\Phi(p^k, 2l) = p^{-k/2} \Bigg[&\sum_{n=0}^{p^k-1} \left(1 + \left(\frac{n}{p}\right)\right) \exp\left(\frac{2\pi i n l}{p^k}\right) + (p^{(k-1)/2} - 1) - \sum_{n=1}^{p^{(k-1)/2}-1} \exp\left(\frac{2\pi i n p l}{p^k}\right) \\ &+ \sum_{i=2,4,\ldots,k-1} \sum_{\substack{n=1 \\ (n,p^k)=p^i \\ n/p^i = m}}^{p^{k-i}-1} \#\{x \mid x^2 = mp^i \bmod p^k\} \exp\left(\frac{2\pi i m l}{p^{k-i}}\right) \Bigg].\end{aligned}$$

The calculation above shows that each inner sum in the last term vanishes, except in the case $i = k - 1$, in which the condition $k - i > 1$ is no longer valid. So we consider this case separately:

$$\begin{aligned}\sum_{\substack{m=1 \\ m \neq \alpha p}}^{p-1} p^{(k-1)/2} (1 + \left(\frac{m}{p}\right)) \exp\left(\frac{2\pi i m l}{p}\right) &= p^{(k-1)/2} \left(-1 + \sum_{m=1}^{p-1} \left(\frac{m}{p}\right) \exp\left(\frac{2\pi i m l}{p}\right)\right) \\ &= p^{(k-1)/2} \left(-1 + \sqrt{p} \left(\frac{l}{p}\right) \Phi(p, 2)\right).\end{aligned}$$

As with the case for k even, the first term in the expansion of $\Phi(p^k, 2l)$ vanishes, the -1 in the second term helps the third term vanish, and the last term only contributes

$$p^{(k-1)/2} \left(-1 + \sqrt{p} \left(\frac{l}{p}\right) \Phi(p, 2)\right),$$

so we are left with

$$\Phi(p^k, 2l) = \left(\frac{l}{p}\right) \Phi(p, 2).$$

If $k = 1$, then it is clear that $\Phi(p, 2l) = \left(\frac{l}{p}\right) \Phi(p, 2)$ in this case too. □

Proposition 4.2.3.2. *Suppose $k \geq 3$ and $(l, 2) = 1$. Then:*

$$\Phi(2^k, 2l) = \begin{cases} \sqrt{2} \exp\left(\frac{\pi i l}{4}\right) & k \text{ odd} \\ 1 + \exp\left(\frac{\pi i l}{2}\right) & k \text{ even} \end{cases}$$

Proof. As before, for $k \geq 3$:

$$\begin{aligned}\Phi(p^k, 2l) &= p^{-k/2} \sum_{n=0}^{p^k-1} \#\{x \mid x^2 = n \bmod p^k\} \exp\left(\frac{2\pi inl}{p^k}\right) \\ &= p^{-k/2} \left[\sum_{n=1 \bmod 8}^{p^k-1} \#\{x \mid x^2 = n \bmod p^k\} \exp\left(\frac{2\pi inl}{p^k}\right) \right. \\ &\qquad \left. + \sum_{i=1}^{k-1} \sum_{(n,p^k)=p^i, n/p^i = m} \#\{x \mid x^2 = mp^i \bmod p^k\} \exp\left(\frac{2\pi imlp^i}{p^k}\right) + N \right],\end{aligned}$$

where $N = p^{k/2}$ if k is even, and $N = p^{(k-1)/2}$ if k is odd.

Suppose $k \geq 4$ is even. Then:

$$\begin{aligned}\Phi(p^k, 2l) &= p^{-k/2} \left[\sum_{n=1 \bmod 8}^{p^k-1} 4\exp\left(\frac{2\pi inl}{p^k}\right) + p^{k/2} + p^{k/2}\exp\left(\frac{\pi il}{2}\right) + \sum_{i=2,4,\ldots,k-4} \sum_{n=1 \bmod 8}^{p^{k-i}-1} 4p^{i/2} \exp\left(\frac{2\pi inlp^i}{p^k}\right) \right] \\ &= p^{-k/2} \left[\sum_{\alpha=0}^{p^{k-3}-1} 4\exp\left(\frac{2\pi i(1+p^3\alpha)l}{p^k}\right) + p^{k/2}\left(1 + \exp\left(\frac{\pi il}{2}\right)\right) + \sum_{i=2,4,\ldots,k-4} \sum_{\alpha=0}^{p^{k-i-3}-1} 4p^{i/2} \exp\left(\frac{2\pi il(1+p^3\alpha)}{p^{k-i}}\right) \right].\end{aligned}$$

Since $k \geq 4$, and the term $i = k-2$ has been treated separately (and appears as the third term in the sum), we have $p^{k-i-3} - 1 > 0$ for $i = 2, 4, \ldots, k-4$, so the final sum is a geometric series and vanishes. Similarly, $k \geq 4$ implies that the first sum vanishes too. Therefore $\Phi(2^k, 2l) = 1 + \exp\left(\frac{\pi il}{2}\right)$.

Now suppose $k > 3$ is odd. This time, the term corresponding to $i = k-1$ appears separately as the third term in the brackets, and the term corresponding to $i = k-3$ appears as the fourth term.

$$\begin{aligned}\Phi(p^k, 2l) = p^{-k/2} &\left[\sum_{\alpha=0}^{p^{k-3}-1} 4\exp\left(\frac{2\pi i(1+p^3\alpha)l}{p^k}\right) + p^{(k-1)/2} + p^{(k-1)/2}\exp(\pi il) + 4p^{(k-3)/2}\exp\left(\frac{2\pi il}{8}\right) \right. \\ &\left. + \sum_{i=2,4,\ldots,k-5} \sum_{\alpha=0}^{p^{k-i-3}-1} 4p^{i/2}\exp\left(\frac{2\pi il(1+p^3\alpha)}{p^{k-i}}\right) \right].\end{aligned}$$

Then since $p^{k-i-3} - 1 > 0$ for $i = 2, 4, \ldots, k-5$, the final sum vanishes, as does the first sum, so we are left with:

$$\Phi(2^k, 2l) = p^{-1/2}\left(1 + (-1)^l + 2\exp\left(\frac{\pi il}{4}\right)\right) = \sqrt{2}\exp\left(\frac{\pi il}{4}\right).$$

Lastly, suppose $k = 3$. Then compared to the case $k > 3$ above, the extra term for $i = k-3$ is omitted, since this is the case $i = 0$ which is already accounted for by the first term. For ease of comparison we write this expression out before explicitly setting $k = 3$:

$$\Phi(p^k, 2l) = p^{-k/2} \left[\sum_{\alpha=0}^{p^{k-3}-1} 4\exp\left(\frac{2\pi i(1+p^3\alpha)l}{p^k}\right) + p^{(k-1)/2} + p^{(k-1)/2}\exp(\pi il) \right].$$

Now we set $k = 3$ in the expression above, and simplify:

$$\Phi(p^k, 2l) = 2^{-3/2}\left(4\exp\left(\frac{\pi il}{4}\right) + 2 + 2(-1)^l\right) = \sqrt{2}\exp\left(\frac{\pi il}{4}\right).$$ □

Finally, we can prove Proposition 4.2.1.4: Let p be an odd prime, $k \geq 1$, and $(p, l) = 1$:

$$\begin{aligned}\Phi(p^k, 2l)\Phi(p^k, -2l) &= \left(\frac{l}{p}\right)\left(\frac{-l}{p}\right)\left(\Phi(p^k, 2)\right)^2 \text{ (by Proposition 4.2.3.1)} \\ &= \left(\frac{-1}{p}\right)\left(\frac{l}{p}\right)^2\left(\frac{-1}{p}\right) \text{ (by Lemma 4.2.1.1)} \\ &= 1.\end{aligned}$$

Let $p = 2$ and suppose $k \geq 3$, $(2, l) = 1$. By Proposition 4.2.3.2:

$$\begin{aligned}\Phi(2^k, 2l)\Phi(2^k, -2l) &= \begin{cases} \left(1 + \exp\left(\frac{\pi i l}{2}\right)\right)\left(1 + \exp\left(-\frac{\pi i l}{2}\right)\right) & k \text{ even} \\ 2\exp\left(\frac{\pi i l}{4}\right)\exp\left(-\frac{\pi i l}{4}\right) & k \text{ odd} \end{cases} \\ &= 2.\end{aligned}$$

4.2.4 Induction

By Lemma 4.2.1.1, the Landsberg–Schaar relation holds for $b = 1$. We proceed by induction on the number of distinct prime factors of b. We assume that $\Phi(a, 2b) = \sqrt{i}\Phi(2b, -a)$ for all b with less than n prime factors, and prove, using Proposition 4.2.1.4, that $\Phi(a, 2bp^k) = \sqrt{i}\Phi(2bp^k, -a)$ for all primes p. We may assume that $(b, p) = 1$, and also $(a, p) = 1$ by Lemma 4.2.1.3. As usual, the case for p an odd prime is treated first.

$$\begin{aligned}\Phi(a, 2bp^k) &= \frac{\Phi(p^k a, 2b)}{\Phi(p^k, 2ab)} && \text{(by Lemma 4.2.1.2)} \\ &= \frac{\sqrt{i}\Phi(2b, -p^k a)}{\Phi(p^k, 2ab)} \\ &= \frac{\sqrt{i}\Phi(2bp^k, -a)}{\Phi(p^k, 2ab)\Phi(p^k, -2ab)} && \text{(by Lemma 4.2.1.2)} \\ &= \sqrt{i}\Phi(2bp^k, -a). && \text{(by Proposition 4.2.1.4)}\end{aligned}$$

Now for $p = 2$:

$$\begin{aligned}\Phi(a, 2b.2^k) = \Phi(a, 2^{k+1}b) &= \frac{\Phi(2^{k+1}a, b)}{\Phi(2^{k+1}, ab)} && \text{(by Lemma 4.2.1.2)} \\ &= \frac{\Phi(2^{k+2}a, 2b)}{\Phi(2^{k+2}, 2ab)} && \text{(by Lemma 4.2.1.3)} \\ &= \frac{\sqrt{i}\Phi(2b, -2^{k+2}a)}{\Phi(2^{k+2}, 2ab)} \\ &= \frac{\sqrt{2}\sqrt{i}\Phi(b, -2^{k+1}a)}{\Phi(2^{k+2}, 2ab)} && \text{(by Lemma 4.2.1.3)} \\ &= \frac{\sqrt{2}\sqrt{i}\Phi(2^{k+1}b, -a)}{\Phi(2^{k+2}, 2ab)\Phi(2^{k+1}, -ab)} && \text{(by Lemma 4.2.1.2)} \\ &= \frac{\sqrt{i}\Phi(2b.2^k, -a)}{\frac{1}{2}\Phi(2^{k+2}, 2ab)\Phi(2^{k+2}, -2ab)} && \text{(by Lemma 4.2.1.3)} \\ &= \sqrt{i}\Phi(2b.2^k, -a). && \text{(by Proposition 4.2.1.4)}\end{aligned}$$

Acknowledgments

The author is extremely grateful to Mike Eastwood for his support and encouragement concerning this article, and most especially for his firm belief that an elementary proof of the Landsberg–Schaar relation should exist! The author would also like to thank Bruce Berndt for reading an earlier draft, Ram Murty for some encouraging remarks, David Roberts for tracking down Gauss' original evaluation of his eponymous sums, and the anonymous referee for suggesting valuable improvements to the article.

Bibliography

The bibliography for the article has been incorporated into the main bibliography.

4.3 Evaluating Gauss sums over $\mathbb{Q}$

Following Berndt, Evans and Williams [BEW98, Chapter 1, Section 5], we evaluate the Gauss sums appearing in the Landsberg–Schaar relation using only elementary methods together with Jacobi's version of the law of quadratic reciprocity, given below as Theorem 4.3.1.1. This version follows in a straightforward manner from the usual formulation in terms of Legendre symbols (Theorem 4.1.2.1), and the reader will find that there are many elementary proofs of Theorem 4.1.2.1 to choose from [Eis44; Gau11]. We also evaluate quadratic Gauss sums with a linear term.

4.3.1 The Jacobi symbol

Recall the definition of the Legendre symbol:

$$\left(\frac{n}{p}\right) = \begin{cases} 0 & p \mid n, \\ 1 & (n,p) = 1 \text{ and } n = x^2 \bmod p \text{ for some integer } x, \\ -1 & (n,p) = 1 \text{ and } n \neq x^2 \bmod p \text{ for any integer } x. \end{cases}$$

We have shown in Section 4.1 that the Legendre symbol satisfies a reciprocity law linking $\left(\frac{q}{p}\right)$ to $\left(\frac{p}{q}\right)$ for distinct odd primes p and q. In this section, we will define an extension of the Legendre symbol, known as the Jacobi symbol, which takes integers a as its upper argument and odd positive integers n in the lower argument. Decomposing n as a product of prime powers, $n = \prod_i p_i^{\nu_i}$, the Jacobi symbol is defined as

$$\left(\frac{a}{n}\right) = \prod_i \left(\frac{a}{p_i}\right)^{\nu_i}. \tag{4.16}$$

Since the Jacobi symbol is multiplicative, it is easily checked that the reciprocity law still holds:

Theorem 4.3.1.1 (Jacobi [Jac46])**.** *Let n and m be distinct odd positive integers. Then*

$$\left(\frac{n}{m}\right)\left(\frac{m}{n}\right) = (-1)^{\frac{(n-1)(m-1)}{4}}.$$

We also have the two supplementary laws:

$$\left(\frac{-1}{n}\right) = (-1)^{\frac{n-1}{2}}, \qquad \left(\frac{2}{n}\right) = (-1)^{\frac{n^2-1}{8}}. \tag{4.17}$$

We may extend the Jacobi symbol to negative odd n by defining

$$\left(\frac{a}{n}\right) = \left(\frac{a}{|n|}\right).$$

The reader should be aware that Theorem 4.3.1.1 does not hold for negative n without adjustments.

4.3.2 No linear term

Let a and b be integers with a and b nonzero and ab even. We wish to evaluate sums of the form

$$s(a,b) = \sum_{n=0}^{|b|-1} \exp\left(\frac{\pi i a n^2}{b}\right).$$

Since ab is even, we may use Lemma 4.3.3.3 from the next subsection to pass to a related sum, which is easier to deal with:

$$2s(a,b) = \sum_{n=0}^{2|b|-1} \exp\left(\frac{2\pi i a n^2}{2b}\right) = s'(a,2b),$$

where we define

$$s'(a,b) = \sum_{n=0}^{|b|-1} \exp\left(\frac{2\pi i a n^2}{b}\right).$$

Note that $s'(ka,kb) = ks'(a,b)$ for all $k > 0$, so if $(a,2b) = k$ with $a = ka'$ and $2b = kb'$, then $2s(a,b) = ks'(a',b')$. We state our results in full before embarking on a case-by-case proof:

Proposition 4.3.2.1. *Suppose $(a,2b) = k$, $a = ka'$ and $2b = kb'$. For $n > 0$ odd, define*

$$\epsilon_n = \begin{cases} 1 & n = 1 \bmod 4, \\ i & n = 3 \bmod 4. \end{cases} \tag{4.18}$$

Then $2s(a,b) = ks'(a',b')$ and

$$s'(a',b') = \begin{cases} \left(\frac{a' \operatorname{sgn} b'}{|b'|}\right)\epsilon_{|b'|}\sqrt{|b'|} & b' \text{ odd}, \\ 0 & b' = 2 \bmod 4, \\ \left(\frac{|b'|}{a \operatorname{sgn} b'}\right)\left(1 + i^{a' \operatorname{sgn} b'}\right)\sqrt{|b'|} & b' = 0 \bmod 4. \end{cases}$$

For the proof of Proposition 4.3.2.1, we closely follow the treatment given by Berndt, Evans and Williams [BEW98, Chapter 1, Section 5]. Our first lemma is the most important, so we state it with much fanfare, although readers may recognise it as a straightforward consequence of the Landsberg–Schaar relation. An exposition of a proof from 1945, due to Estermann 4.3.2.2, is given by Berndt, Evans and Williams [BEW98], but the first treatment (using analysis) is due to Gauss [Gau01a]. One may also consult Pólya [Pól27] for a rather underhanded proof using the central limit theorem.

Lemma 4.3.2.2 (Gauss, Estermann)**.** *Suppose b is a positive odd integer. Then*

$$s'(1,b) = \epsilon_b\sqrt{b}.$$

Now we may begin to prove Proposition 4.3.2.1 proper.

Lemma 4.3.2.3. *Suppose a and b are integers, b is odd and positive and $(a,b) = 1$. Then*

$$s'(a,b) = \left(\frac{a}{b}\right)\epsilon_b\sqrt{b}.$$

Proof. First, note that if p is an odd prime, then $\epsilon_{p^k} = \pm\epsilon_p^k$, and if a and b are odd and coprime, then $\epsilon_{ab} = \pm\epsilon_a\epsilon_b$. Write b as a product of distinct odd prime powers: $b = \prod_i p_i^{\mu_i}$. Then

$$s'(1,b) = s'(1,\prod_i p_i^{\mu_i}) = \pm\prod_i \epsilon_{p_i}^{\mu_i}\sqrt{p_i^{\mu_i}} = \pm\prod_i s'(1,p_i)^{\mu_i} \tag{4.19}$$

Now suppose $(a,b) = 1$ and let σ_a denote the automorphism of $\mathbb{Q}(e^{2\pi i/k})$ characterised by $e^{2\pi i/k} \mapsto e^{2\pi i a/k}$. Then by 4.13 and 4.19,

$$s'(a,b) = \sigma_a s'(1,b) = \pm\prod_i (\sigma_a s'(1,p_i))^{\mu_i} = \pm\prod_i\left(\left(\frac{a}{p_i}\right)s'(1,p_i)\right)^{\mu_i}$$

The proof is concluded upon using the definition of the Jacobi symbol (4.16).

$$s'(a,b) = \pm\left(\frac{a}{b}\right)\prod_i s'(1,p_i)^{\mu_i} = \left(\frac{a}{b}\right)s'(1,b).$$

□

Replacing b by $|b|$ and a by $a \operatorname{sgn} b$, we see that Lemma 4.3.2.3 implies the case of Proposition 4.3.2.1 in which b' is odd.

Lemma 4.3.2.4. *Suppose a and b are integers, $b = 2 \bmod 4$, $b > 0$ and $(a, b) = 1$. Then*

$$s'(a, b) = 0.$$

Proof. Since b is even, a is odd, so it follows that

$$\exp\left(\frac{2\pi i(n + b/2)^2 a}{b}\right) = -\exp\left(\frac{2\pi i n^2 a}{b}\right).$$

Hence $s'(a, b) = -s'(a, b)$. □

As above, replacing b by $|b|$ and a by $a \operatorname{sgn} b$, we obtain the case in which $b' = 2 \bmod 4$.

Lemma 4.3.2.5. *Suppose a and b are integers, $b = 0 \bmod 4$, $b > 0$ and $(a, b) = 1$. Then*

$$s'(a, b) = \left(\frac{b}{a}\right)(1 + i^a)\sqrt{b}.$$

Proof. Write $b = 2^k b'$, where b' is odd and $k \geq 4$. Suppose that the lemma holds for $b' = 1$. Then by 4.14,

$$\begin{aligned} s'(a, b) = s'(2^k a, b')s(ab', 2^k) &= \left(\frac{2^k a}{b'}\right)\left(\frac{2^k}{ab'}\right)\left(1 + i^{ab'}\right)\epsilon_{b'}\sqrt{b} \\ &= \left(\frac{b}{a}\right)(-1)^{(a-1)(b'-1)/4}\left(1 + i^{ab'}\right)\epsilon_{b'}\sqrt{b}. \end{aligned}$$

So we need to show that

$$(-1)^{(a-1)(b'-1)/4}\left(1 + i^{ab'}\right)\epsilon_{b'} = (1 + i^a),$$

but this is easily checked case by case modulo 4, as a and b' may only be congruent to 1 or 3 mod 4.

So we need to prove the lemma for powers of two. Suppose that $k \geq 4$. Then

$$\begin{aligned} s'(a, 2^k) &= \sum_{\substack{n=0 \\ n \text{ odd}}}^{2^k - 1} \exp\left(\frac{2\pi i n^2 a}{2^k}\right) + \sum_{\substack{n=0 \\ n \text{ even}}}^{2^k - 1} \exp\left(\frac{2\pi i n^2 a}{2^k}\right) \\ &= \sum_{n=0}^{2^{k-1}-1} \exp\left(\frac{2\pi i(2n+1)^2 a}{2^k}\right) + 2\sum_{n=0}^{2^{k-2}-1} \exp\left(\frac{2\pi i n^2 a}{2^{k-2}}\right). \end{aligned}$$

Now the first sum vanishes, since if we replace n by $n + 2^{k-3}$, the summand is negated:

$$\begin{aligned} \exp\left(\frac{2\pi i(2n + 1 + 2^{k-2})^2 a}{2^k}\right) &= \exp\left(\frac{2\pi i\left((2n+1)^2 + 2(2n+1)2^{k-2} + 2^{2k-4}\right)a}{2^k}\right) \\ &= \exp\left(\frac{2\pi i(2n+1)^2 a}{2^k}\right)\exp\left(\frac{\pi i 2^{2k-3} a}{2^k}\right), \end{aligned}$$

and we use the hypothesis that $k \geq 4$ for the first time in the last line of the equation above, to ensure that the final term is equal to one.

We now have a recurrence relation, valid for $k \geq 4$ and $2l \leq k - 2$:

$$s'(a, 2^k) = 2s'(a, 2^{k-2}) = \cdots = 2^l s'(a, 2^{k-2l}).$$

It follows that

$$s'(a, 2^k) = \begin{cases} 2^{(k-1)/2} s'(a, 4) & k \text{ even}, \\ 2^{(k-3)/2} s'(a, 8) & k \text{ odd}. \end{cases}$$

Since we may check that

$$s'(a, 4) = 2(1 + i) = \left(\frac{4}{a}\right)(1 + i^a)\sqrt{4}$$

and

$$s'(a,8) = 4e^{\pi i a/4} = \left(\frac{2}{a}\right)(1+i^a)\sqrt{8},$$

the lemma is proved in full. □

Replacing b by $|b|$ and a by $a \operatorname{sgn} b$, we obtain the case in which $b' = 0 \bmod 4$.

4.3.3 Linear terms

Let a, b and c be integers, and a and c nonzero. We wish to evaluate sums of the form

$$S(a,b,c) = \sum_{n=0}^{|c|-1} \exp\left(\frac{\pi i(an^2+bn)}{c}\right) = \sum_{n \bmod |c|} \exp\left(\frac{\pi i(an^2+bn)}{c}\right),$$

where $ac+b$ is even. This last stipulation is a natural restriction on a, b and c, for it is equivalent to the well-definedness of the rightmost sum in the definition of $S(a,b,c)$. We will evaluate these sums by "completing the square", so that we may use the results of the previous subsection. Define k by $0 < k = \gcd(a,c)$, and set $a = ka'$, $c = kc'$. For convenience, we state our results in full, before embarking on a case-by-case proof:

Proposition 4.3.3.1. *Suppose $ac+b$ is even. Then*

$$S(a,b,c) = \begin{cases} \left.\begin{array}{ll} kS(4a',2b',c') & k \mid b \ \ (b = kb') \\ 0 & k' \nmid b. \end{array}\right\} c \text{ odd}, \\ \left.\begin{array}{ll} kS(4a',2b',c') & k \mid b \ \ (b = kb'),\ \ b' \text{ odd}, \\ 0 & k' \nmid b \text{ or } k \mid b \ \ (b = kb'),\ \ b' \text{ even}. \end{array}\right\} c \text{ even, } a'c' \text{ odd}, \\ \left.\begin{array}{ll} kS(a',2b''',c') & k \mid b'' \ \ (b = 2kb''') \\ 0 & k' \nmid b''. \end{array}\right\} c \text{ even, } a'c' \text{ even } (b \text{ even, } b = 2b''). \end{cases}$$

For $a'c'$ odd,

$$S(4a',2b',c') = \exp\left(\frac{2\pi i(b')^2(2a'\mu^2-\mu)}{c'}\right)S(4a',0,c') = \sqrt{c'}\left(\frac{2a'}{c}\right)\exp\left(\frac{2\pi i(b')^2(2a'\mu^2-\mu)}{c'}\right),$$

where $4a'\mu = 1 + lc'$. For $a'c'$ even,

$$\begin{aligned} S(a',2b''',c') &= \exp\left(\frac{\pi i(b''')^2(a'\nu^2-2\nu)}{c'}\right)S(a',0,c') \\ &= \begin{cases} \left(\frac{a' \operatorname{sgn}(c)/2}{|c'|}\right)\epsilon_{|c'|}\sqrt{|c'|}\exp\left(\frac{\pi i(b''')^2(a'\nu^2-2\nu)}{c'}\right) & c' \text{ odd}, \\ \frac{1}{2}\left(\frac{|c'|/2}{a' \operatorname{sgn} c}\right)\left(1+i^{a' \operatorname{sgn} c}\right)\sqrt{|c'|}\exp\left(\frac{\pi i(b''')^2(a'\nu^2-2\nu)}{c'}\right) & c' \text{ even}, \end{cases} \end{aligned}$$

where $a'\nu = 1 + lc'$.

We require some lemmas first.

Lemma 4.3.3.2. *For a, b, c_1 and c_2 integers, a, c_1 and c_2 nonzero, $ac_1c_2 + b$ even and $(c_1, c_2) = 1$,*

$$S(a,b,c_1c_2) = S(ac_2, b \operatorname{sgn} c_2, c_1)S(ac_1, b \operatorname{sgn} c_1, c_2).$$

Proof. By assumption, $ac_1c_2 + b$ is even, so

$$S(a,b,c_1c_2) = \sum_{n \bmod |c_1c_2|} \exp\left(\frac{\pi i(an^2+bn)}{c_1c_2}\right).$$

As s runs from 0 to $|c_1|-1$ modulo $|c_1|$ and t runs from 0 to $|c_2|-1$ modulo $|c_2|$, $n = |c_2|s + |c_1|t$ modulo $|c_1c_2|$. So

$$\begin{aligned}
S(a,b,c) &= \sum_{s=0}^{|c_1|-1}\sum_{t=0}^{|c_2|-1} \exp\left(\frac{\pi i(a(|c_2|s+|c_1|t)^2 + b(|c_2|s+|c_1|t)}{c}\right) \\
&= \sum_{s=0}^{|c_1|-1} \exp\left(\frac{\pi i(ac_2s^2 + b\,\mathrm{sgn}\,(c_2)n)}{c_1}\right)\sum_{t=0}^{|c_2|-1} \exp\left(\frac{\pi i(ac_1s^2 + b\,\mathrm{sgn}\,(c_1)n)}{c_2}\right) \\
&= S(ac_2, b\,\mathrm{sgn}\,c_2, c_1)S(ac_1, b\,\mathrm{sgn}\,c_1, c_2).
\end{aligned}$$

□

Lemma 4.3.3.3. *For a, b, c and k integers, a and c nonzero and $k > 0$,*

$$S(ka,kb,kc) = \begin{cases} kS(a,b,c) & ac+b \text{ even,} \\ S(a,b,c) & ac+b \text{ and } k \text{ odd,} \\ 0 & ac+b \text{ odd and } k \text{ even.} \end{cases}$$

Proof. Partitioning the summation range, we have

$$\begin{aligned}
S(ka,kb,kc) &= \sum_{n=0}^{k|c|-1} \exp\left(\frac{\pi i(an^2+bn)}{c}\right) \\
&= \sum_{m=0}^{k-1}\sum_{n=m|c|}^{(m+1)|c|-1} \exp\left(\frac{\pi i(an^2+bn)}{c}\right) \\
&= \sum_{m=0}^{k-1}\sum_{n=0}^{|c|-1} \exp\left(\frac{\pi i(a(m|c|+n)^2 + b(m|c|+n))}{c}\right) \\
&= \sum_{m=0}^{k-1} \exp\left(\pi i(ac+b)m\right)\sum_{n=0}^{|c|-1} \exp\left(\frac{\pi i(an^2+bn)}{c}\right).
\end{aligned}$$

One easily checks that the sum over m has the value k, 1 or 0 according to the conditions given above. □

Lemma 4.3.3.4. *If*

$$\sum_{s=0}^{k-1} \exp\left(\pi i a'c'\,\mathrm{sgn}\,(c)s^2\right)\exp\left(\frac{\pi i b\,\mathrm{sgn}\,(c)s}{k}\right)$$

vanishes, then $S(a,b,c) = 0$.

Proof. Write $c = kc'$. As s runs from 0 to $k-1$ and t runs from 0 to $|c'|-1$, $n = |c'|s + t$ runs from 0 to $|c|-1$. So

$$\begin{aligned}
S(a,b,c) &= \sum_{s=0}^{k-1}\sum_{t=0}^{|c'|-1} \exp\left(\frac{\pi i(a(|c'|s+t)^2 + b(|c'|s+t))}{c}\right) \\
&= \sum_{s=0}^{k-1} \exp\left(\frac{\pi i(as^2|c'|^2 + bs|c'|)}{c}\right)\sum_{t=0}^{|c'|-1} \exp\left(\frac{\pi i(at^2+bt)}{c}\right),
\end{aligned}$$

and writing $a = ka'$,

$$\sum_{s=0}^{k-1} \exp\left(\frac{\pi i(as^2|c'|^2 + bs|c'|)}{c}\right) = \sum_{s=0}^{k-1} \exp\left(\pi i a'c'\,\mathrm{sgn}\,(c)s^2\right)\exp\left(\frac{\pi i b\,\mathrm{sgn}\,(c)s}{k}\right).$$

□

Lemma 4.3.3.5. *Suppose c is odd, $ac+b$ is even. Then $S(a,b,c) = 0$ unless $k \mid b$.*

Proof. Using the oddness of c, the evenness of $ac+b$, Lemma 4.3.3.3 and Lemma 4.3.3.2,

$$\begin{aligned} S(a,b,c) &= \frac{1}{2}S(2a,2b,2c) \\ &= \frac{1}{2}S(4a,2b,c)S(2ac,2b\,\mathrm{sgn}\,(c),2) \\ &= S(4a,2b,c). \end{aligned}$$

Applying Lemma 4.3.3.4 to $S(4a,2b,c)$, we find that $S(a,b,c)$ vanishes if

$$\sum_{s=0}^{k-1} \exp\left(\pi i 4a'c'\,\mathrm{sgn}\,(c)s^2\right)\exp\left(\frac{2\pi i b\,\mathrm{sgn}\,(c)s}{k}\right) = \sum_{s=0}^{k-1}\exp\left(\frac{2\pi i b\,\mathrm{sgn}\,(c)s}{k}\right) = 0.$$

But the last sum is a geometric series, which vanishes unless $k \mid b$. □

Lemma 4.3.3.6. *Suppose c and $ac+b$ are even: then b is even, so we write $b = 2b''$. If $a'c'$ is even, then $S(a,b,c) = 0$ unless $k \mid b''$, and if $a'c'$ is odd, $S(a,b,c) = 0$ unless $k \mid b$.*

Proof. By Lemma 4.3.3.4, it suffices to investigate the sum

$$\sum_{s=0}^{k-1} \exp\left(\pi i a'c'\,\mathrm{sgn}\,(c)s^2\right)\exp\left(\frac{2\pi i b''\,\mathrm{sgn}\,(c)s}{k}\right).$$

If $a'c'$ is even, then the sum is a geometric series and vanishes unless $k \mid b''$. If $a'c'$ is odd, then since c is even k must be even too: write $k = 2k''$. We split the sum over even and odd s:

$$\sum_{s=0}^{k-1}(-1)^s\exp\left(\frac{2\pi i b''\,\mathrm{sgn}\,(c)s}{k}\right) = \sum_{s=0}^{k''-1}\exp\left(\frac{2\pi i b''\,\mathrm{sgn}\,(c)s}{k''}\right) - \exp\left(\frac{2\pi i b''\,\mathrm{sgn}\,(c)}{k}\right)\sum_{s=0}^{k''-1}\exp\left(\frac{2\pi i b''\,\mathrm{sgn}\,(c)s}{k''}\right),$$

and both sums are geometric series which vanish unless $k'' \mid b''$, which is the same as $k \mid b$. □

Proposition 4.3.3.7. *Suppose that c is odd and $ac+b$ is even. Then $S(a,b,c) = 0$ unless $k \mid b$, whence there is some b' with $b = kb'$. In that case,*

$$S(a,b,c) = k\epsilon_{c'}\sqrt{c'}\left(\frac{2a'}{c'}\right)\exp\left(\frac{2\pi i (b')^2(2a'\mu^2-\mu)}{c'}\right),$$

where μ is defined by $4a'\mu = 1 + lc$ for some l.

Proof. In the proof of Lemma 4.3.3.5, we saw that $S(a,b,c) = S(4a,2b,c)$, and by Lemma 4.3.3.5, $S(a,b,c) = 0$ unless $k \mid b$. If $k \mid b$, then by Lemma 4.3.3.3,

$$S(a,b,c) = S(4a,2b,c) = kS(4a',2b',c') = k\sum_{n=0}^{|c'|-1}\exp\left(\frac{2\pi i(2a'n^2+b'n)}{c'}\right).$$

Since c' is odd, $(4a',c') = 1$, so we may find μ and l so that $4a'\mu = 1 + lc'$. In the last sum, we make the change of variables $n = m - \mu b' \bmod c$, so that

$$S(a,b,c) = k\exp\left(\frac{2\pi i (b')^2(2a'\mu^2-\mu)}{c'}\right)\sum_{m=0}^{|c'|-1}\exp\left(\frac{4\pi i a' m^2}{c'}\right), \tag{4.20}$$

and using Lemma 4.3.2.2 to evaluate the remaining Gauss sum, the claim follows. □

Proposition 4.3.3.8. *Suppose that c, $ac+b$ and $a'c'$ are even. Then b is even, so set $b = 2b''$. By Lemma 4.3.3.6, $S(a,b,c) = 0$ unless $k \mid b''$, whence there is some b''' with $b = 2kb'''$. In that case, there exist μ and m so that $a'\mu = 1 + lc'$, and we have*

$$S(a,b,c) = \begin{cases} \left(\frac{a' \operatorname{sgn}(c)/2}{|c'|}\right)\epsilon_{|c'|}\sqrt{|c'|}\exp\left(\frac{\pi i (b''')^2(a'\mu^2 - 2\mu)}{c'}\right) & c' \text{ odd,} \\ \frac{1}{2}\left(\frac{|c'|/2}{a' \operatorname{sgn} c}\right)\left(1 + i^{a' \operatorname{sgn} c}\right)\sqrt{|c'|}\exp\left(\frac{\pi i (b''')^2(a'\mu^2 - 2\mu)}{c'}\right) & c' \text{ even.} \end{cases}$$

Proof. Suppose we may write $b = 2kb'''$. By Lemma 4.3.3.3,

$$S(a,b,c) = S(ka', 2kb''', kc') = kS(a', 2b''', c').$$

Making the change of variables $n = m - \mu b''' \bmod c$, we complete the square:

$$S(a,b,c) = k\sum_{n=0}^{|c'|-1}\exp\left(\frac{\pi i(a'n^2 + 2b'''n)}{c'}\right) = k\exp\left(\frac{\pi i (b''')^2(a'\mu^2 - 2\mu)}{c'}\right)\sum_{m=0}^{|c'|-1}\exp\left(\frac{\pi i m^2 a'}{c'}\right). \tag{4.21}$$

The claim follows upon using Proposition 4.3.2.1 to evaluate the final Gauss sum. □

Proposition 4.3.3.9. *Suppose that c and $ac+b$ are even and $a'c'$ is odd. Then $S(a,b,c) = 0$ unless $k \mid b$ and $b' = b/k$ is odd, in which case*

$$S(a,b,c) = k\sqrt{c'}\left(\frac{a'}{c'}\right)\exp\left(\frac{2\pi i (b')^2(2a'\mu^2 - \mu)}{c'}\right),$$

where μ and m satisfy $4a'\mu = 1 + mc'$.

Proof. By Lemma 4.3.3.6, $S(a,b,c) = 0$ unless $k \mid b$. Since k is even and $a'c'$ is odd, Lemma 4.3.3.3 yields

$$S(a,b,c) = S(ka', kb', kc') = \begin{cases} kS(a', b', c') & b' \text{ odd,} \\ 0 & b' \text{ even.} \end{cases}$$

Assuming that b' is odd, and noticing that c' is odd, we may invoke Proposition 4.3.3.7 to simplify $S(a',b',c')$:

$$S(a,b,c) = k\exp\left(\frac{2\pi i (b')^2(2a'\mu^2 - \mu)}{c'}\right)\sum_{m=0}^{|c'|-1}\exp\left(\frac{4\pi i a' m^2}{c'}\right), \tag{4.22}$$

followed by Proposition 4.3.2.1 to evaluate the Gauss sum. □

4.4 Another elementary proof of the Landsberg–Schaar relation

In this section we prove, using the results of Section 4.3, a slightly more general version of the Landsberg–Schaar relation:

Proposition 4.4.0.1. *Suppose a and b are nonzero integers and ab is even. Recall that*

$$s(a,b) := \sum_{n=0}^{|b|-1}\exp\left(\frac{\pi i n^2 a}{b}\right).$$

Then

$$s(a,b) = \sqrt{\left|\frac{b}{a}\right|}\exp\left(\frac{\pi i}{4}\right)s(-b,a).$$

Actually, if we set

$$s'(a,b) = \sum_{n=0}^{|b|-1} \exp\left(\frac{2\pi i n^2 a}{b}\right),$$

then by Lemma 4.3.3.3, Proposition 4.4.0.1 is equivalent to the following:

Proposition 4.4.0.2. *Suppose a and b are nonzero integers. Then*

$$s'(a,b) = \sqrt{\left|\frac{b}{8a}\right|} \exp\left(\frac{\pi i}{4}\right) s'(-b, 4a).$$

We will prove Proposition 4.4.0.2 by evaluating the relevant Gauss sums using Proposition 4.3.2.1.

Proof. Since $s'(ka, kb) = ks'(a,b)$, we may assume that $(a,b) = 1$. Suppose first that b is odd. Then

$$s'(a,b) = \left(\frac{a \operatorname{sgn} b}{|b|}\right) \epsilon_{|b|} \sqrt{|b|} = \left(\frac{\operatorname{sgn}(a)\operatorname{sgn}(b)}{|b|}\right) \epsilon_{|b|} \left(\frac{|a|}{|b|}\right) \sqrt{|b|},$$

and

$$s'(-b, 4a) = \left(\frac{4|a|}{-a \operatorname{sgn} b}\right) \left(1 + i^{-b \operatorname{sgn} a}\right) \sqrt{4|a|} = \frac{1 + i^{-b \operatorname{sgn} a}}{\sqrt{2}} \left(\frac{|a|}{|b|}\right) \sqrt{8|a|}.$$

So we need to prove that

$$\left(\frac{\operatorname{sgn}(a)\operatorname{sgn}(b)}{|b|}\right) \epsilon_{|b|} = \frac{1 + i^{-b \operatorname{sgn} a}}{\sqrt{2}} \exp\left(\frac{\pi i}{4}\right). \tag{4.23}$$

If a and b are both positive or both negative, then the Legendre symbol disappears and we can check 4.23 using the definition of ϵ_b: if a and b are positive then it's easily checked that

$$\epsilon_b = \frac{1 + i^{-b}}{\sqrt{2}} \exp\left(\frac{\pi i}{4}\right), \tag{4.24}$$

and if a and b are both negative then we're down to

$$\epsilon_{|b|} = \frac{1 + i^{-|b|}}{\sqrt{2}} \exp\left(\frac{\pi i}{4}\right),$$

which is the same as 4.24.

If exactly one of a or b is negative, then using the first supplementary law, 4.17, the expression 4.23 becomes

$$(-1)^{\frac{b-1}{2}} \epsilon_{|b|} = \frac{1 + i^{-b \operatorname{sgn} a}}{\sqrt{2}} \exp\left(\frac{\pi i}{4}\right), \tag{4.25}$$

and regardless of which of a or b is negative, 4.25 becomes

$$(-1)^{\frac{b-1}{2}} \epsilon_{|b|} = \frac{1 + i^{|b|}}{\sqrt{2}} \exp\left(\frac{\pi i}{4}\right). \tag{4.26}$$

which is also easily checked.

Now suppose that $b = 0 \bmod 2$: then $s'(a,b) = 0$. Set $b = 2b''$, where b'' is odd, and note that since a must be odd, we have

$$s'(-b, 4a) = s'(-2b'', 4a) = 2s'(-b'', 2a) = 0.$$

Finally, suppose that $b = 0 \bmod 4$. Then $b = 4b''$ for some b'' and $(a, b'') = 1$, and a is odd. We have

$$s'(a,b) = \left(\frac{|b|}{a \operatorname{sgn} b}\right) \left(1 + i^{a \operatorname{sgn} b}\right) \sqrt{|b|} = \left(1 + i^{a \operatorname{sgn} b}\right) \left(\frac{|b''|}{|a|}\right) \sqrt{|b|},$$

and

$$\begin{aligned} s'(-b,4a) &= s'(-4b'',4a) = 4s'(-b'',a) \\ &= 4\left(\frac{-b''\operatorname{sgn} a}{|a|}\right)\epsilon_{|a|}\sqrt{|a|} \\ &= \sqrt{2}\left(\frac{-\operatorname{sgn}(a)\operatorname{sgn}(b)}{|a|}\right)\epsilon_{|a|}\left(\frac{|b''|}{|a|}\right)\sqrt{8|a|}. \end{aligned}$$

We need to prove that

$$\left(1+i^{a\operatorname{sgn} b}\right) = \sqrt{2}\left(\frac{-\operatorname{sgn}(a)\operatorname{sgn}(b)}{|a|}\right)\epsilon_{|a|}\exp\left(\frac{\pi i}{4}\right). \tag{4.27}$$

If a and b have the same sign, then using the first supplementary law 4.11 again, the expression 4.27 reduces to

$$\left(1+i^{|a|}\right) = \sqrt{2}(-1)^{\frac{a-1}{2}}\epsilon_{|a|}\exp\left(\frac{\pi i}{4}\right),$$

which has already been dealt with as 4.26. If a and b have the opposite sign, then the Legendre symbol vanishes and 4.27 reduces to

$$\left(1+i^{-|a|}\right) = \sqrt{2}\epsilon_{|a|}\exp\left(\frac{\pi i}{4}\right),$$

which we have already met as 4.24. □

Interchanging a and b, taking both to be positive and using Lemma 4.3.3.3, we see that Proposition 4.4.0.2 recovers the version of the Landsberg–Schaar relation proved earlier.

4.5 A generalised Landsberg–Schaar relation

In Section 4.3, we investigated "generalised" quadratic Gauss sums, and found that they behave very similarly to the usual quadratic Gauss sums. It is natural to wonder if there is a version of the Landsberg–Schaar relation for these sums. Such a relation was discovered and given an analytic proof by Guinand [Gui45]:

Proposition 4.5.0.1. *Let a, b and c be integers, a and c nonzero and $ac+b$ even. Then*

$$S(a,b,c) = \sqrt{\left|\frac{a}{c}\right|}\exp\left(\frac{\pi i}{4}\left(\operatorname{sgn} ac - \frac{b^2}{ac}\right)\right)S(-c,-b,a), \tag{4.28}$$

where

$$S(a,b,c) = \sum_{n=0}^{|c|-1}\exp\left(\frac{\pi i(an^2+bn)}{c}\right).$$

Even more general versions were obtained by Berndt [Ber73] and Mordell [Mor62; Mor33], but their proofs rely on analysis. It is therefore of interest to give a completely elementary proof; however, to further illustrate the point that interesting reciprocity relations between Gauss sums come from theta functions, we give an analytic proof as well, in the same style as Section 4.1.

4.5.1 Analytic proof

First, we need a theta function. We define

$$\theta(z,x) = \sum_{n=-\infty}^{\infty} e^{\pi i(n^2 z + nx)},$$

where the sum converges absolutely for $\operatorname{Im}(z) > 0$ and all $x \in \mathbb{C}$. In complete analogy with the usual theta function $\theta(z) = \theta(z,0)$, we have a transformation law.

Proposition 4.5.1.1. *With the branch cut for the square root taken along the non-positive imaginary axis, we have*

$$\theta(-1/z,-x/z) = \sqrt{z}e^{-\pi i/4}e^{\pi i x^2/4z}\theta(z,x). \tag{4.29}$$

Proof. By substituting $n \mapsto -n$ in the summation index, we may assume that $-x/z$ is replaced by x/z. Set $z = 2\tau$, $x = 2\omega$, $Q = (1)$ and $m = 1$ in Proposition 1.2.3.1. □

The idea of the proof of Proposition 4.5.0.1 is to compute the asymptotic expansion of both sides of 4.29 at rational points on the boundary of the upper half z-plane for a well-chosen x.

Lemma 4.5.1.2. *For integers a, b and c with a and c nonzero and $ac + b$ even, and $\mathrm{Re}(\epsilon) > 0$, we have*

$$\theta\left(\frac{a}{c} + i\epsilon, \frac{b}{c}\right) \sim \frac{1}{\sqrt{2}|c|} S(a,b,c)\epsilon^{-1/2} + O(|\epsilon|^N) \tag{4.30}$$

as $\epsilon \to 0$, for all $N > 0$.

Proof. Note that

$$\theta\left(\frac{a}{c} + i\epsilon, \frac{b}{c}\right) = S(a,b,c) \sum_{m = n \bmod |c|} e^{-\pi\epsilon m^2}$$

and use the proof of Proposition 4.1.1.1. □

We also need a slightly different asymptotic expansion, owing to the shape of the functional equation:

Lemma 4.5.1.3. *For integers a, b and c with a and c nonzero and $ac + b$ even, and $\mathrm{Re}(\epsilon) > 0$, we have*

$$\theta\left(\frac{a}{c} + i\epsilon, \frac{a}{c} + \frac{b}{a}\epsilon i\right) \sim \frac{1}{\sqrt{2}|c|} S(a,b,c)\epsilon^{-1/2} + O(|\epsilon|^{1/2}) \tag{4.31}$$

Proof. It is clear from the proof of Proposition 4.1.1.1 that

$$\sum_{m = n \bmod |c|} e^{-\pi\epsilon(m^2 + bm/a)} \sim \int_{-\infty}^{\infty} e^{-\pi\epsilon\left((n+|c|t)^2 + (b/a)(n+|c|t)\right)} dt + O(|\epsilon|^N)$$

as $\epsilon \to 0$, for all $N > 0$. The integral is easily computed:

$$\int_{-\infty}^{\infty} e^{-\pi\epsilon\left((n+|c|t)^2 + (b/a)(n+|c|t)\right)} dt = \frac{1}{2\sqrt{|c|}} \epsilon^{-1/2} e^{-\pi\epsilon a^2/4b^2}, \tag{4.32}$$

so the lemma is proved. □

Remark 4.5.1.4. *The term $e^{-\pi\epsilon a^2/4b^2}$ appearing in 4.32 prevents the error term in Lemma 4.5.1.3 from achieving $O(|\epsilon|^N)$ for all $N > 0$ as in Proposition 4.1.1.1, but as this term is independent of n, the higher terms in the asymptotic expansion contain no additional information about the behaviour of $\theta(z,x)$.*

We will now prove Proposition 4.5.0.1.

Proof. The functional equation from 4.35 implies that

$$e^{-\pi i/4}\sqrt{\frac{a}{c} + i\epsilon}\, \exp\left(\frac{\pi i}{4} \frac{(b/c)^2}{a/c + i\epsilon}\right) \theta\left(\frac{a}{c} + i\epsilon, \frac{b}{c}\right) = \theta\left(-\frac{1}{a/c + i\epsilon}, -\frac{b/c}{a/c + i\epsilon}\right).$$

Set

$$-\frac{1}{a/c + \epsilon} = -\frac{c}{a} + i\tau.$$

Since $\mathrm{Re}(\epsilon) > 0$, $\mathrm{Re}(\tau) > 0$, and $\epsilon \to 0$ if and only if $\tau \to 0$. Therefore

$$\sqrt{\left|\frac{a}{c}\right|} \exp\left(-\frac{\pi i}{4}\left(\mathrm{sgn}\,(ac) + \frac{b^2}{ac}\right)\right) \lim_{\epsilon \to 0} \sqrt{\epsilon}\,\theta\left(\frac{a}{c} + i\epsilon, \frac{b}{c}\right) = \left|\frac{a}{c}\right| \lim_{\tau \to 0} \sqrt{\tau}\,\theta\left(-\frac{c}{a} + i\tau, -\frac{b}{a} + \frac{bi}{c}\tau\right).$$

Employing Lemma 4.5.1.2 and Lemma 4.5.1.3 to calculate the limits, we obtain

$$\sqrt{\left|\frac{a}{c}\right|} \exp\left(-\frac{\pi i}{4}\left(\mathrm{sgn}\,(ac) + \frac{b^2}{ac}\right)\right) S(a,b,c) = S(-c,-b,a),$$

which is the same as 4.28. □

4.5.2 Elementary proof

Using the results of Section 4.3 and the elementary proof of the Landsberg–Schaar relation, we prove Proposition 4.5.0.1. We split the proof into two cases depending on the properties of a, b and c. The strategy of the proof in each case is to reduce, using Propositions 4.3.3.7, 4.3.3.8 and 4.3.3.9, to Gauss sums without linear terms, for which we may use the results of 4.3.2. Throughout the proof, we will use the notation of Subsection 4.3.3: $k = \gcd(a, c)$, $a = ka'$ and $c = kc'$.

If c is odd, or if c is even and $a'c'$ is odd, then by 4.20 and 4.22, the Gauss sums involved in 4.28 evaluate to the same expression. Consequently, we treat these two cases together first.

Proposition 4.5.2.1. *Let a, b and c be integers with a and c positive and $ac + b$ even. Suppose c is odd, or c is even and $a'c'$ is odd. Then*

$$S(a, b, c) = \sqrt{\left|\frac{a}{c}\right|} \exp\left(\frac{\pi i}{4}\left(\operatorname{sgn} ac - \frac{b^2}{ac}\right)\right) S(-c, -b, a).$$

Proof. By Lemma 4.3.3.5, both the sums in question vanish unless $k \mid b$. So suppose $k \mid b$ and set $b = kb'$. In this case, Proposition 4.3.3.7 and Proposition 4.3.3.9 yield

$$S(-c, -b, a) = k \exp\left(-\frac{2\pi i (b')^2 (2c'\mu^2 - \mu)}{a'}\right) S(-4c', 0, a'),$$

where $4c'\mu = 1 + ma'$. Similarly,

$$\begin{aligned} S(a, b, c) &= k \exp\left(\frac{2\pi i (b')^2 (2a'\nu^2 - \nu)}{c'}\right) S(4a', 0, c') \\ &= k\sqrt{\left|\frac{c'}{4a'}\right|} \exp\left(\frac{\pi i \operatorname{sgn}(a'c')}{4}\right) \exp\left(\frac{2\pi i (b')^2 (2a'\nu^2 - \nu)}{c'}\right) S(-c', 0, 4a') \\ &= \frac{k}{2}\sqrt{\left|\frac{c}{a}\right|} \exp\left(\frac{\pi i \operatorname{sgn}(ac)}{4}\right) \exp\left(\frac{2\pi i (b')^2 (2a'\nu^2 - \nu)}{c'}\right) S(-c', 0, 4a'), \end{aligned}$$

where $4a'\nu = 1 + mc'$ and we have used Proposition 4.4.0.1 (by definition, $s(a, b) = S(a, 0, b)$) to relate $S(4a', 0, c')$ to $S(-c', 0, 4a')$. By Lemma 4.3.3.3,

$$S(-c', 0, 4a') = S(-4c', 0, a') S(-a'c', 0, 4),$$

and $S(-a'c', 0, 4) = 2\exp\left(-\frac{\pi i a'c'}{4}\right)$ since $a'c'$ is odd. So it suffices to show that

$$\exp\left(-\frac{\pi i a'c'}{4}\right) \exp\left(\frac{2\pi i (b')^2 (2a'\nu^2 - \nu)}{c'}\right) \exp\left(\frac{2\pi i (b')^2 (2c'\mu^2 - \mu)}{a'}\right) = \exp\left(-\frac{\pi i (b')^2}{4a'c'}\right) = \exp\left(-\frac{\pi i (b)^2}{4ac}\right).$$

The left hand side of the last equation can be rewritten as

$$\begin{aligned} &\exp\left(-\frac{\pi i (a'c')^2}{4a'c'}\right) \exp\left(\frac{\pi i (b')^2}{4a'c'}\left((4a'\nu)^2 - 8a'\nu + (4c\mu)^2 - 8c\mu\right)\right) \\ &\qquad = \exp\left(-\frac{\pi i (b')^2}{4a'c'}\right) \exp\left(\frac{\pi i (b')^2}{4a'c'}\left(-1 - (a'c')^2 + (lc)^2 + (ma)^2\right)\right), \end{aligned}$$

we need only show that

$$8a'c' \mid -1 - (a'c')^2 + (lc)^2 + (ma)^2.$$

Since $(8, a', c') = 1$ we may proceed one factor at a time. First, note that a', c', l and m are all odd, each of $a'c', lc$ and ma are congruent to 1 mod 8, so 8 divides the expression. To show that a' divides the expression we need to verify that $a' \mid l^2c^2 - 1$. But $l^2c^2 - 1 = (4a\nu - 1)^2 - 1 = 16a^2\nu^2 - 8a\nu$. The same argument works for c'. □

Now we treat the remaining case, in which c and $a'c'$ are both even.

Proposition 4.5.2.2. *Let a, b and c be integers with a and c positive and $ac+b$ even. Suppose c and $a'c'$ are even. Then*

$$S(a,b,c)=\sqrt{\left|\frac{a}{c}\right|}\exp\left(\frac{\pi i}{4}\left(\operatorname{sgn}ac-\frac{b^2}{ac}\right)\right)S(-c,-b,a).$$

Proof. We may assume that $k \mid b/2$, otherwise both sums in question are zero, so that there is some b''' with $b = 2kb'''$. The expression 4.21 yields

$$S(-c,-b,a)=k\exp\left(-\frac{\pi i(b''')^2(c'\mu^2-2\mu)}{a'}\right)S(-c',0,a'),$$

where $c'\mu = 1 + ma'$. Similarly,

$$\begin{aligned}
S(a,b,c) &= k\exp\left(\frac{\pi i(b''')^2(a'\nu^2-2\nu)}{c'}\right)S(a',0,c')\\
&= k\sqrt{\left|\frac{c'}{a'}\right|}\exp\left(\frac{\pi i\operatorname{sgn}(a'c')}{4}\right)\exp\left(\frac{\pi i(b''')^2(a'\nu^2-2\nu)}{c'}\right)S(-c',0,a')\\
&= k\sqrt{\left|\frac{c}{a}\right|}\exp\left(\frac{\pi i\operatorname{sgn}(ac)}{4}\right)\exp\left(\frac{\pi i(b''')^2(a'\nu^2-2\nu)}{c'}\right)S(-c',0,a')
\end{aligned}$$

where $a'\nu = 1 + lc'$ and we have used Proposition 4.4.0.1 to relate $(a', 0, c')$ to $S(-c', 0, a')$. So it suffices to show that

$$\exp\left(\frac{\pi i(b''')^2(a'\nu^2-2\nu)}{c'}\right)\exp\left(\frac{\pi i(b''')^2(c'\mu^2-2\mu)}{a'}\right)=\exp\left(-\frac{\pi i(b''')^2}{a'c'}\right)=\exp\left(-\frac{\pi ib^2}{4ac}\right).$$

Rewriting the left hand side,

$$\exp\left(\frac{\pi i(b''')^2}{a'c'}\left((a'\nu)^2-2a'\nu+(c'\mu)^2-2c'\mu\right)\right)=\exp\left(-\frac{\pi i(b''')^2}{a'c'}\right)\exp\left(\frac{\pi i(b''')^2}{a'c'}\left(-1+l^2(c')^2+m^2(a')^2\right)\right),$$

we find that this is the same as showing that

$$2a'c' \text{ divides } -1+l^2(c')^2+m^2(a')^2. \tag{4.33}$$

Recall that $a'c'$ is even and by construction $(a', c') = 1$, so exactly one of a' or c' is even. The expression 4.33 is invariant if we interchange a' and c' and swap l and m accordingly, so without loss of generality we suppose that a' is even and c' is odd. Now we need only show that $2a'$ and c' divide $-1 + l^2(c')^2 + m^2(a')^2$.

It's clear that c' divides $-1 + m^2(a')^2 = -1 + (c'\mu - 1)^2$. On the other hand,

$$-1+(lc')^2+(ma')^2=-1+(a'\nu-1)^2+(ma')^2=\left(a'(\nu^2+m^2)-2\nu\right)a'.$$

Since a' is even, the expression in the brackets on the right hand side is even, so the whole expression is divisible by $2a'$. □

4.6 A generalised twisted Landsberg–Schaar relation

Let χ be a primitive Dirichlet character modulo m (see Definition 1.3.3.1). Then we may define a theta function twisted by χ as follows, with $\operatorname{Im}(z) > 0$ and $x \in \mathbb{C}$:

$$\theta_\chi(z,x)=\sum_{n=-\infty}^{\infty}\chi(n)e^{\pi i(n^2z/m+nx/m)}.$$

We may write this twisted theta function as a sum of the generalised theta functions of Subsection 4.5.1, and the functional equations proved there provide a functional equation for $\theta_\chi(z, x)$. After we have proved the functional equation, we may compute the asymptotic expansion of $\theta_\chi(z, x)$ and deduce a twisted Landsberg–Schaar relation as in Section 4.1:

Proposition 4.6.0.1 (Berndt, [Ber73]). *Let a, b and c be integers, a and c nonzero, $acm+b$ even and χ a primitive Dirichlet character modulo m. Then*

$$\frac{G(\overline{\chi})}{\sqrt{|c|m}}\sum_{n=0}^{|c|m-1}\chi(n)\exp\left(\frac{\pi i(an^2+bn)}{cm}\right)=\frac{1}{\sqrt{|a|}}\exp\left(\frac{\pi i}{4}\left(\operatorname{sgn}a-\frac{b^2}{acm}\right)\right)\sum_{n=0}^{|a|m-1}\overline{\chi(n)}\exp\left(-\frac{\pi i(cn^2+bn)}{am}\right). \tag{4.34}$$

If we define

$$S_\chi(a,b,c)=\sum_{n=0}^{|c|m-1}\chi(n)\exp\left(\frac{\pi i(an^2+bn)}{cm}\right),$$

then Proposition 4.6.0.1 assumes the form

$$\frac{G(\overline{\chi})}{\sqrt{|c|m}}S_\chi(a,b,c)=\sqrt{\frac{1}{|a|}}\exp\left(\frac{\pi i}{4}\left(\operatorname{sgn}a-\frac{b^2}{acm}\right)\right)S_{\overline{\chi}}(-c,-b,a).$$

In the case that $b=0$, Proposition 4.6.0.1 was first proved in 1944 by Guinand [Gui45] using Fourier-analytic techniques.

Alternatively, we may rewrite the left hand side of 4.34 as a sum of generalised Gauss sums from Subsection 4.5.2, and then Proposition 4.6.0.1 follows from 4.28. This method of proof entirely avoids analytical techniques.

4.6.1 Analytic proof

As usual, we begin by proving a transformation law for $\theta_\chi(z,x)$.

Proposition 4.6.1.1. *With the branch cut for the square root taken along the non-positive imaginary axis, we have*

$$\theta_\chi(-1/z,-x/z)=\sqrt{\frac{z}{m}}e^{-\pi i/4}G(\chi)e^{\pi ix^2/4mz}\theta_{\overline{\chi}}(z,x). \tag{4.35}$$

Proof. The strategy of the proof is to rewrite $\theta_\chi(z,x)$ as a sum of m generalised theta functions from Subsection 4.5.1, and then apply Proposition 4.5.1.1 to each one.

$$\begin{aligned}\theta_\chi(z,x)&=\sum_{k=0}^{m-1}\chi(k)\sum_{n=k\bmod m}e^{\pi i(n^2z+nx)}\\&=\sum_{k=0}^{m-1}\chi(k)e^{\pi i(k^2z+kx)/m}\sum_{n=-\infty}^{\infty}e^{\pi i(n^2mz+(2kz+x)n)}\\&=\sum_{k=0}^{m-1}\chi(k)e^{\pi i(k^2z+kx)/m}\theta(mz,2kz+x).\end{aligned}$$

So we have

$$\theta_\chi\left(-\frac{1}{z},-\frac{x}{z}\right)=\sum_{k=0}^{m-1}\chi(k)\exp\left(-\frac{\pi i(k^2+kx)}{mz}\right)\theta\left(-\frac{1}{z/m},-\frac{(2k+x)/m}{z/m}\right). \tag{4.36}$$

By Proposition 4.5.1.1,

$$\theta\left(-\frac{1}{z/m},-\frac{(2k+x)/m}{z/m}\right)=\sqrt{\frac{z}{m}}e^{-\pi i/4}\exp\left(\frac{\pi i(2k+x)^2}{4mz}\right)\theta(z/m,(2k+x)/m), \tag{4.37}$$

Now we substitute 4.37 into 4.36 and head towards $\theta_{\overline{\chi}}(z,x)$:

$$\exp\left(-\frac{\pi i(k^2+kx)}{mz}\right)\exp\left(\frac{\pi i(2k+x)^2}{4mz}\right)\theta(z/m,(2kz+x)/m)=\exp\left(\frac{\pi ix^2}{4mz}\right)\sum_{n=-\infty}^{\infty}e^{\pi i(n^2z/m+n(2k+x)/m)},$$

so

$$\theta_\chi\left(-\frac{1}{z}, -\frac{x}{z}\right) = \sqrt{\frac{z}{m}} e^{-\pi i/4} e^{\pi i x^2/4mz} \sum_{n=-\infty}^{\infty} \left(\sum_{k=0}^{m-1} \chi(k) e^{2\pi i k n/m}\right) e^{\pi i (n^2 z/m + nx/m)}$$
$$= \sqrt{\frac{z}{m}} e^{-\pi i/4} G(\chi) e^{\pi i x^2/4mz} \theta_{\overline{\chi}}(z, x),$$

where we have used Lemma 1.3.3.3 in the final line. □

As in Subsection 4.5.1, the idea of the proof of Proposition 4.6.0.1 is to compute the asymptotic expansion of both sides of 4.6.1.1 for suitable choices of z and x.

Lemma 4.6.1.2. *For integers a, b and c with a and c nonzero, $acm + b$ even and* $\mathrm{Re}(\epsilon) > 0$, *we have*

$$\theta_\chi\left(\frac{a}{c} + i\epsilon, \frac{b}{c}\right) \sim \frac{1}{\sqrt{2m|c|}} S_\chi(a, b, c) \epsilon^{-1/2} + O(|\epsilon|^N) \tag{4.38}$$

as $\epsilon \to 0$, for all $N > 0$.

Proof. Note that

$$\theta_\chi\left(\frac{a}{c} + i\epsilon, \frac{b}{c}\right) = S_\chi(a, b, c) \sum_{l=n \bmod |c|m} e^{-\pi\epsilon m^2},$$

and by the proof of Proposition 4.1.1.1,

$$\sum_{l=n \bmod |c|m} e^{-\pi\epsilon l^2} = \frac{1}{\sqrt{2m|c|}} \epsilon^{-1/2} + O(|\epsilon|^N)$$

as $\epsilon \to 0$, for all $N > 0$. □

We require one other asymptotic expansion:

Lemma 4.6.1.3. *For integers a, b and c with a and c nonzero and $acm + b$ even, and* $\mathrm{Re}(\epsilon) > 0$, *we have*

$$\theta_\chi\left(\frac{a}{c} + i\epsilon, \frac{a}{c} + \frac{bi\epsilon}{a}\right) \sim \frac{1}{\sqrt{2m|c|}} S_\chi(a, b, c) \epsilon^{-1/2} + O(|\epsilon|^{1/2})$$

Proof. It is clear from the proof of Proposition 4.1.1.1 that

$$\sum_{l=n \bmod |c|m} e^{-\pi\epsilon(l^2 + bl/a)} \sim \int_{-\infty}^{\infty} e^{-\pi\epsilon\left((n+|c|mt)^2 + (b/a)(n+|c|mt)\right)} dt + O(|\epsilon|^N)$$

as $\epsilon \to 0$, for all $N > 0$. The integral is easily computed:

$$\int_{-\infty}^{\infty} e^{-\pi\epsilon\left((n+|c|mt)^2 + (b/a)(n+|c|mt)\right)} dt = \frac{1}{\sqrt{2m|c|}} \epsilon^{-1/2} e^{-\pi\epsilon b^2/4ma^2}, \tag{4.39}$$

so the lemma is proved. □

Remark 4.6.1.4. *As in Proposition 4.1.1.1, the term $e^{-\pi\epsilon b^2/4ma^2}$ appearing in 4.39 prevents the error term in Lemma 4.6.1.3 from achieving $O(|\epsilon|^N)$ for all $N > 0$. But since this term is independent of n, the higher terms in the asymptotic expansion contain no additional information about the behaviour of $\theta_\chi(z, x)$.*

We will now prove Proposition 4.6.0.1.

Proof. The functional equation from Proposition 4.6.1.1 implies that

$$\frac{e^{-\pi i/4}}{\sqrt{m}} G(\overline{\chi}) \sqrt{\frac{a}{c} + i\epsilon} \exp\left(\frac{\pi i}{4} \frac{(b/c)^2}{a/c + i\epsilon}\right) \theta_\chi\left(\frac{a}{c} + i\epsilon, \frac{b}{c}\right) = \theta_{\overline{\chi}}\left(-\frac{1}{a/c + i\epsilon}, -\frac{b/c}{a/c + i\epsilon}\right).$$

Set
$$-\frac{1}{a/c+i\epsilon} = -\frac{c}{a} + i\tau.$$
Since $\mathrm{Re}(\epsilon) > 0$, $\mathrm{Re}(\tau) > 0$, and $\epsilon \to 0$ if and only if $\tau \to 0$. Therefore
$$\sqrt{\left|\frac{a}{c}\right|}G(\overline{\chi})\exp\left(-\frac{\pi i}{4}\left(\mathrm{sgn}\,(ac) - \frac{b^2}{ac}\right)\right)\lim_{\epsilon\to 0}\sqrt{\epsilon}\theta_\chi\left(\frac{a}{c}+i\epsilon, \frac{b}{c}\right) = \left|\frac{a}{c}\right|\lim_{\tau\to 0}\sqrt{\tau}\theta_{\overline{\chi}}\left(-\frac{c}{a}+i\tau, -\frac{b}{a}+\frac{bi}{c}\tau\right).$$
Employing Lemma 4.6.1.2 and Lemma 4.6.1.3 to calculate the limits, we obtain
$$\sqrt{\left|\frac{a}{c}\right|}G(\overline{\chi})\exp\left(-\frac{\pi i}{4}\left(\mathrm{sgn}\,(ac) - \frac{b^2}{ac}\right)\right)S_\chi(a,b,c) = S_{\overline{\chi}}(-c,-b,a),$$
which, upon using Lemma 1.3.3.3, is the same as 4.34. □

4.6.2 Elementary proof

We begin, as promised, by simplifying the left hand side of 4.34:
$$\begin{aligned}
G(\overline{\chi})\sum_{n=0}^{|c|m-1}\chi(n)\exp\left(\frac{\pi i(an^2+bn)}{cm}\right) &= \sum_{l=0}^{m-1}\overline{\chi(l)}\exp\left(\frac{2\pi il}{m}\right)\sum_{n=0}^{|c|m-1}\chi(n)\exp\left(\frac{\pi i(an^2+bn)}{cm}\right)\\
&= \sum_{l=0}^{m-1}\overline{\chi(l)}\sum_{n=0}^{|c|m-1}\exp\left(\frac{2\pi iln}{m}\right)\exp\left(\frac{\pi i(an^2+bn)}{cm}\right)\\
&= \sum_{l=0}^{m-1}\overline{\chi(l)}\sum_{n=0}^{|c|m-1}\exp\left(\frac{\pi i(an^2+(b+2cl)n)}{cm}\right)\\
= \sqrt{\frac{|c|m}{|a|}}\exp\left(\frac{\pi i\,\mathrm{sgn}\,a}{4}\right)&\sum_{l=0}^{m-1}\overline{\chi(l)}\exp\left(-\frac{\pi i(b+2cl)^2}{4acm}\right)\sum_{n=0}^{|a|-1}\exp\left(-\frac{\pi i(cmn^2+(b+2cl)n)}{a}\right)\\
= \sqrt{\frac{|c|m}{|a|}}\exp\left(\frac{\pi i}{4}\left(\mathrm{sgn}\,a - \frac{b^2}{acm}\right)\right)&\sum_{l=0}^{m-1}\overline{\chi(l)}\exp\left(-\frac{\pi i(bl+cl^2)}{am}\right)\sum_{n=0}^{|a|-1}\exp\left(-\frac{\pi i(cmn^2+(b+2cl)n)}{a}\right),
\end{aligned} \tag{4.40}$$
where the critical step was the deployment of 4.28 to arrive at the second-to-last line. On the other hand, as s runs from 0 to $|a|-1$ and t runs from 0 to $m-1$, $n = ms+t$ runs from 0 to $|a|m-1$, so we may rewrite the right hand side of 4.34:
$$\begin{aligned}
\sum_{n=0}^{|a|m-1}\overline{\chi(n)}\exp\left(-\frac{\pi i(cn^2+bn)}{am}\right) &= \sum_{s=0}^{|a|-1}\sum_{t=0}^{m-1}\overline{\chi(ms+t)}\exp\left(-\frac{\pi i(c(ms+t)^2+b(ms+t))}{am}\right)\\
&= \sum_{t=0}^{m-1}\overline{\chi(t)}\exp\left(-\frac{\pi i(ct^2+bt)}{am}\right)\sum_{s=0}^{|a|-1}\exp\left(-\frac{\pi i(cm^2s^2+(b+2ct)ms)}{am}\right).
\end{aligned} \tag{4.41}$$
The proof of Proposition 4.6.0.1 is complete upon substituting 4.41 into 4.40.

4.6.3 An elementary proof of Berndt's reciprocity relation

In the same article in which he gave an analytic proof of 4.60, Berndt stated a generalisation of 4.60, in which the Dirichlet character is replaced by a finite periodic sequence.

Indeed, let (a_n) be an m-periodic sequence of complex numbers. Let (b_n) be the $\mathbb{Z}/m\mathbb{Z}$ Fourier transform of (a_n):
$$a_n = \sum_{j=0}^{m-1} b_j e^{2\pi ijn/m}$$
or
$$b_n = \frac{1}{m}\sum_{j=0}^{m-1} a_j e^{-2\pi ijn/m}.$$

Now we may state Berndt's theorem:

Proposition 4.6.3.1 (Berndt [Ber73]). *Suppose a, b and c are integers, with a and c nonzero and $acm + b$ even. Then*

$$\sum_{n=0}^{|c|m-1} a_n \exp\left(\frac{\pi i(an^2+bn)}{cm}\right) = \sqrt{\frac{|c|m}{|a|}} \exp\left(\frac{\pi i}{4}\left(\operatorname{sgn} a - \frac{b^2}{acm}\right)\right) \sum_{n=0}^{|a|m-1} b_n \exp\left(-\frac{\pi i(cn^2+bn)}{am}\right). \tag{4.42}$$

It turns out that Proposition 4.6.3.1 admits an elementary proof in exactly the same style as in Proposition 4.6.0.1.

Proof. We begin with the left hand side of 4.42.

$$\begin{aligned}\sum_{n=0}^{|c|m-1} a_n \exp\left(\frac{\pi i(an^2+bn)}{cm}\right) &= \sum_{n=0}^{|c|m-1}\sum_{j=0}^{m-1} b_j e^{2\pi ij/m} \exp\left(\frac{\pi i(an^2+bn)}{cm}\right)\\ &= \sum_{j=0}^{m-1} b_j \sum_{n=0}^{|c|m-1} \exp\left(\frac{\pi i(an^2+(b+2cj)n)}{cm}\right)\end{aligned}$$

$$= \sqrt{\frac{|c|m}{|a|}} \exp\left(\frac{\pi i \operatorname{sgn} a}{4}\right) \sum_{j=0}^{m-1} b_j \exp\left(-\frac{\pi i(b+2cj)^2}{4acm}\right) \sum_{n=0}^{|a|-1} \exp\left(-\frac{\pi i(cmn^2+(b+2cj)n)}{a}\right)$$

$$= \sqrt{\frac{|c|m}{|a|}} \exp\left(\frac{\pi i}{4}\left(\operatorname{sgn} a - \frac{b^2}{acm}\right)\right) \sum_{j=0}^{m-1} b_j \exp\left(-\frac{\pi i(bj+cj^2)}{am}\right) \sum_{n=0}^{|a|-1} \exp\left(-\frac{\pi i(cmn^2+(b+2cj)n)}{a}\right), \tag{4.43}$$

where we used Proposition 4.5.0.1 to arrive at the second-to-last line. On the other hand, as s runs from 0 to $|a|-1$ and t runs from 0 to $m-1$, $n = ms + t$ runs from 0 to $|a|m - 1$, so we may rewrite the right hand side of 4.42:

$$\begin{aligned}\sum_{n=0}^{|a|m-1} b_n \exp\left(-\frac{\pi i(cn^2+bn)}{am}\right) &= \sum_{s=0}^{|a|-1}\sum_{t=0}^{m-1} b_{ms+t} \exp\left(-\frac{\pi i(c(ms+t)^2+b(ms+t))}{am}\right)\\ &= \sum_{t=0}^{m-1} b_t \exp\left(-\frac{\pi i(ct^2+bt)}{am}\right) \sum_{s=0}^{|a|-1} \exp\left(-\frac{\pi i(cm^2s^2+(b+2ct)ms)}{am}\right).\end{aligned} \tag{4.44}$$

The proof is complete upon substituting 4.44 into 4.43. □

As one should expect [Ber73], Proposition 4.6.3.1 implies Proposition 4.6.0.1.

4.7 Gauss sums of higher degree

For the purposes of the following subsections it turns out to be more expedient to express higher-degree Gauss sums in terms of sums of lower degree, if possible. In anticipation, we define

$$s_d(a, b) = \sum_{n=0}^{b-1} \exp\left(\frac{\pi i a n^d}{b}\right). \tag{4.45}$$

The rule of thumb is that $s_d(a, b)$ is "maximally difficult" only when b is a prime congruent to 1 modulo d. In all other cases, $s_d(a, b)$ may be expressed in terms of products of $s_{d'}(a', p)$ for some $d' < d$, where p is a prime congruent to 1 modulo d'. First, we show that the evaluation of $s_d(a, b)$ may be reduced to the case where b is a prime power.

Lemma 4.7.0.1. *Suppose a, b_1 and b_2 are nonzero and $(b_1, b_2) = 1$. Then*

$$s_d(2a, b_1b_2) = s_d(2ab_2^{d-1}, b_1)s_d(2ab_1^{d-1}, b_2). \tag{4.46}$$

Proof. Rewriting $s_d(a, b_1b_2)$ as a sum over a residue class,

$$s_d(2a, b_1b_2) = \sum_{n \bmod b_1b_2} \exp\left(\frac{2\pi i a n^d}{b_1b_2}\right).$$

Now, as s runs over a complete residue class modulo b_2 and t runs over a complete residue class modulo b_1, $n = b_1 s + b_2 t$ runs over a complete residue class modulo $b_1 b_2$. The claim follows from the fact that

$$n^d = b_1^{d-1} s^d + b_2^{d-1} t^d \bmod b_1 b_2.$$

□

Conjecture 4.7.0.2. *Suppose p is an odd prime and $(d,p) = 1$. Write $\alpha = dk + l$, where $k \geq 0$ and $1 \leq l \leq d$. Then*

$$s_d(2, p^\alpha) = \begin{cases} p^{(d-1)k} s_d(2,p) & l = 1 \\ p^{(d-1)k+l-1} & l \neq 1. \end{cases} \tag{4.47}$$

Suppose $p = 2$, b is odd and write $\alpha = dk + l$, where $k \geq 0$ and $1 \leq l \leq d$. If d is odd, then

$$s_d(1, 2^\alpha) = \begin{cases} 0 & l = 1 \\ 2^{(d-1)k+l-1} & l \neq 1. \end{cases}$$

If d is even, then

$$s_d(1, 2^\alpha) = \begin{cases} 0 & l = 1 \\ 2^{(d-1)k+l-1} \left(1 + \exp\left(\frac{\pi i}{2^l}\right)\right) & 2 \leq l \leq 2\nu_2(d) \\ 2^{(d-1)k+l-1} & 2\nu_2(d) < l \leq d, \end{cases}$$

where $\nu_2(d)$ is the 2-adic valuation of d: the highest power of 2 which divides d.

If p is odd, $p \mid d$ and $\alpha = dk + l$, where $k \geq 0$ and $1 \leq l \leq d$, then

$$s_d(2, p^\alpha) = p^{(d-1)k+l-1}$$

for $1 + \nu_p(d) < l \leq d$, where there is not, at first glance, an obvious simplification for $s_d(b, p^l)$, $2 \leq l \leq 1 + \nu_p(d)$.

The analogous results for which $s_d(1, \dots)$ (resp. $s_d(2, \dots)$) is replaced by $s_d(b,k)$ (resp. $s_d(2b,k)$), are easily obtained by applying the $\mathbb{Q}\left(\exp\left(\frac{\pi i}{k}\right)\right)$-automorphism $\exp\left(\frac{\pi i}{k}\right) \mapsto \exp\left(\frac{b\pi i}{k}\right)$ (resp. $\exp\left(\frac{2\pi i}{k}\right) \mapsto \exp\left(\frac{2b\pi i}{k}\right)$) to the stated equations.

All we require for what follows in this text is 4.47, in addition to a result which allows us to determine how far we can reduce the degree of a given exponential sum. We prove these statements in the next two subsections.

4.7.1 Reduction of exponent for odd prime moduli coprime to the degree

The object of this subsection is to prove the following result:

Proposition 4.7.1.1. *Suppose p is an odd prime and $(d,p) = 1$. Write $\alpha = dk + l$, where $k \geq 0$ and $1 \leq l \leq d$. Then*

$$s_d(2, p^\alpha) = \begin{cases} p^{(d-1)k} s_d(2,p) & l = 1 \\ p^{(d-1)k+l-1} & l \neq 1. \end{cases}$$

Before we begin the proof, we require some results about the number of solutions to the equation $x^d = k \bmod p^\alpha$. The proofs involve counting arguments similar to those employed earlier in Subsection 4.2.

Lemma 4.7.1.2. *Let p be any prime, not necessarily odd or coprime to d, and suppose that $j > i$.*

$$\#\{x \mid x^d = kp^i \bmod p^j\} = \begin{cases} p^{i/d} \#\{x \mid x^d = k \bmod p^{j-i}\} & i = 0 \bmod d \\ 0 & i \neq 0 \bmod d, (k,p) = 1. \end{cases}$$

Proof. The case of $i \neq 0 \bmod d$ is immediate upon noting that

$$x^d = kp^i + k'p^j$$

implies p^i divides x, so p divides k. So we may suppose that $i = 0 \bmod d$. Define

$$A = \{z \in \mathbb{Z}/p^j\mathbb{Z} \mid z^2 = kp^i \bmod p^j\}$$
$$B = \{z \in \mathbb{Z}/p^{j-i}\mathbb{Z} \mid z^2 = k \bmod p^{j-i}\}.$$

The map $F : B \to A$ defined by $x \mapsto p^{i/d}x$ is surjective, so in order to prove that $|A| = p^{i/d}|B|$, we need to show that each fibre of F has cardinality $p^{i/d}$. One may easily check that the fibre of F over $F(x)$ is contained in

$$\{x_t = x + tp^{j-i/d} \mid t = 0, \dots, p^{i/d} - 1\} \subset \mathbb{Z}/p^{j-i}\mathbb{Z}.$$

But the x_t are contained in B, since

$$(x + tp^{j-i/d})^d = x^d + O(p^{j-i/d}) = x^d \bmod p^{j-i},$$

so the fibre of F over x consists of exactly the x_t. □

As a consequence, we may count solutions to $x^d = 0 \bmod p^\alpha$.

Lemma 4.7.1.3.

$$\#\{x \mid x^d = 0 \bmod p^\alpha\} = \begin{cases} p^{\alpha/d+d-2}, & \alpha = 0 \bmod d. \\ p^{\alpha-s-1} & \alpha = l + sd, 2 \le l \le d-1. \end{cases}$$

Proof. If $\alpha = 0 \bmod d$, set $j = \alpha$, $i = \alpha - d$ and $k = 0$ in Lemma 4.7.1.2. Then

$$\begin{aligned}\#\{x \mid x^d = 0 \bmod p^\alpha\} &= p^{(\alpha-d)/d}\#\{x \mid x^d = 0 \bmod p^d\} \\ &= p^{(\alpha-d)/d+d-1} \\ &= p^{\alpha/d+d-2}.\end{aligned}$$

If $\alpha = l + sd$, where $1 \le l \le d-1$, then $p^\alpha \mid x^d$ if and only if $p^{(\alpha+d-l)/d} \mid x$, which is the same as $p^{s+1} \mid x$. So

$$\#\{x \mid x^d = 0 \bmod p^\alpha\} = \#\{tx \mid t = 0, \dots, p^{\alpha-s-1} - 1\} = p^{\alpha-s-1}.$$ □

The next result is standard; one may consult Ireland and Rosen [IR90, Proposition 4.2.3, pp. 46].

Lemma 4.7.1.4. *Suppose that p is an odd prime, $(p, d) = 1$, $j \ge 1$ and $(p, k) = 1$. Then*

$$\#\{x \mid x^d = k \bmod p^j\} = \#\{x \mid x^d = k \bmod p\}.$$

Combining Lemmas 4.7.1.2, 4.7.1.3 and 4.7.1.4, we have:

Lemma 4.7.1.5. *Suppose that p is an odd prime, $(p, d) = 1$, $j > i \ge 0$ and $(p, k) = 1$. Then*

$$\#\{x \mid x^d = kp^i \bmod p^j\} = \begin{cases} p^{i/d}\#\{x \mid x^d = k \bmod p\} & i = 0 \bmod d \\ 0 & i \ne 0 \bmod d. \end{cases}$$

Proof of Proposition 4.7.1.1. We begin by reducing to the case $k = 0$. Indeed, consider

$$s_d(2, p^\alpha) = \sum_{n=0}^{p^\alpha - 1} \exp\left(\frac{2\pi i n^d}{p^\alpha}\right),$$

and write $n = m + m'p^{\alpha-1}$, where m runs from 0 to $p^{\alpha-1} - 1$ and m' runs from 0 to $p - 1$. Then

$$n^d = m^d + dm^{d-1}m'p^{\alpha-1} \bmod p^\alpha,$$

so

$$s_d(2, p^\alpha) = \sum_{m=0}^{p^{\alpha-1}-1} \exp\left(\frac{2\pi i m^d}{p^\alpha}\right) \sum_{m'=0}^{p-1} \exp\left(\frac{2\pi i d m^{d-1} m'}{p}\right),$$

and the inner sum vanishes unless $p \mid dm^{d-1}$. Since $p \nmid d$, we must have $p \mid m$. It follows that

$$s_d(2, p^\alpha) = p \sum_{m=0}^{p^{\alpha-d}-1} \exp\left(\frac{2\pi i m^d}{p^{\alpha-d}}\right) = ps_d(2, p^{\alpha-d}),$$

so Proposition 4.7.1.1 reduces to the claim that

$$s_d(2, p^l) = p^{l-1} \tag{4.48}$$

for $2 \le l \le d$.

Now we prove 4.48 for $2 \le l \le d$.

$$\begin{aligned}
s_d(2,p^l) &= \sum_{n=0}^{p^l-1} \#\{x \mid x^d = n \bmod p^l\} \exp\left(\frac{2\pi i n}{p^l}\right) \\
&= \sum_{n=0}^{p^l-1} \#\{x \mid x^d = n \bmod p\} \exp\left(\frac{2\pi i n}{p^l}\right) + (p^{l-1}-1) - \sum_{n=1}^{p^{l-1}-1} \exp\left(\frac{2\pi i n p}{p^l}\right) \\
&\quad + \sum_{i=1}^{l-1} \sum_{\substack{n=1 \\ (n,p^l)=p^i \\ n/p^i = m}}^{p^{k-i}-1} \#\{x \mid x^d = mp^i \bmod p^l\} \exp\left(\frac{2\pi i m}{p^{k-i}}\right).
\end{aligned}$$

We should explain the expansion at the second line of the equation above: by Lemma 4.7.1.4, the first term gives the correct coefficients for $(n, p) = 1$, the second term employs Lemma 4.7.1.3 to ensure the correct contribution for $n = 0$, the third term makes sure that the coefficient of any nonzero index n divisible by p is equal to zero and the last term restores the correct solutions counts for nonzero n divisible by p.

By Lemma 4.7.1.2, the inner sum occurring in the last term of the preceding display vanishes unless $i = 0 \bmod d$. Since $l \le d$, this situation never occurs.

We now turn to the first term. By Proposition 1.3.3.4, since $(n, p) = 1$, we have

$$\#\{x \mid x^d = n \bmod p\} = \sum_{\substack{|\chi|=p \\ \chi^d=1}} \chi(n),$$

where the sum on the right hand side takes place over the Dirichlet characters modulo p of order d. The first term may be rewritten as

$$\sum_{n=0}^{p^l-1} \#\{x \mid x^d = n \bmod p\} \exp\left(\frac{2\pi i n}{p^l}\right) = \sum_{\substack{|\chi|=p \\ \chi^d=1}} \sum_{n=0}^{p^l-1} \chi(n) \exp\left(\frac{2\pi i n}{p^l}\right),$$

and each of the inner sums vanish:

$$\begin{aligned}
\sum_{n=0}^{p^l-1} \chi(n) \exp\left(\frac{2\pi i n}{p^l}\right) &= \sum_{s=0}^{p^{l-1}-1} \sum_{n=sp}^{(s+1)p-1} \chi(n) \exp\left(\frac{2\pi i n}{p^l}\right) \\
&= \sum_{s=0}^{p^{l-1}-1} \sum_{n=0}^{p-1} \chi(n+sp) \exp\left(\frac{2\pi i (n+sp)}{p^l}\right) \\
&= \sum_{s=0}^{p^{l-1}-1} \exp\left(\frac{2\pi i s}{p^{l-1}}\right) \sum_{n=0}^{p-1} \chi(n) \exp\left(\frac{2\pi i n}{p^l}\right),
\end{aligned}$$

as the outer sum is a geometric series. So we are left with

$$s_d(2,p^l) = p^{l-1} - \sum_{n=0}^{p^{l-1}-1} \exp\left(\frac{2\pi i n p}{p^l}\right) = p^{l-1}. \qquad \square$$

4.7.2 Reduction to sums of lower degree

Given a sum of the form $s_d(2,p)$ for a prime p, the next lemma tells us the smallest d' such that $s_d(2,p) = s_{d'}(2,p)$.

Lemma 4.7.2.1. *Suppose that p is a prime $d \geq 2$. Then*

$$s_d(2,p) = \begin{cases} s_q(2,p) & \text{for the largest } q \geq 1 \text{ such that } q \mid d \text{ and } p = 1 \bmod q, \\ 0 & \text{no such } q \text{ exists.} \end{cases} \tag{4.49}$$

In fact, we prove a slightly stronger statement about dth powers in the finite fields $\mathbb{F}_{p^k}$.

Proof. Consider the map $\Upsilon : \mathbb{F}_{p^k}^\times \to \mathbb{F}_{p^k}^\times$ defined by

$$\Upsilon(x) = x^d.$$

We begin by determining the cardinality of the fibre of Υ over 1. It is well-known that $\mathbb{F}_{p^k}^\times$ is cyclic of order $p^k - 1$, so let g be a generator. For any $x \in \mathbb{F}_{p^k}^\times$, $x^d = 1$ if and only if $x = g^m$ for some m such that dm divides $|\mathbb{F}_{p^k}^\times| = p^k - 1$, which is the same as $m \mid (p^k - 1)/\gcd(d, p^k - 1)$. Consequently, the fibre of Υ over 1 is the cyclic group generated by

$$g^{(p^k-1)/\gcd(d,p^k-1)},$$

so it has order exactly $\gcd(d, p^k - 1) = q$.

Therefore Υ is a $q : 1$ map, and $\#\{d\text{th powers in } \mathbb{F}_{p^k}^\times\} = (p^k - 1)/q$. Upon repeating the argument in the paragraph above with d replaced by q, we find that $x \mapsto x^q$ is a $\gcd(q, p^k - 1) : 1$ map. But $\gcd(q, p^k - 1) = q$, so since $q \mid d$,

$$\#\left\{d\text{th powers in } \mathbb{F}_{p^k}^\times\right\} = \#\left\{q\text{th powers in } \mathbb{F}_{p^k}^\times\right\},$$

and the result follows upon setting $k = 1$. □

4.8 Twisted Gauss sums

We invite the reader to ponder the twisted Landsberg–Schaar relation from Section 4.6, for a primitive Dirichlet character χ of modulus m, a, b and c integers with $acm + b$ even and a and c nonzero:

$$\frac{G(\overline{\chi})}{\sqrt{|c|m}} \sum_{n=0}^{|c|m-1} \chi(n) \exp\left(\frac{\pi i(an^2 + bn)}{cm}\right) = \sqrt{\frac{1}{|a|}} \exp\left(\frac{\pi i}{4}\left(\operatorname{sgn} a - \frac{b^2}{acm}\right)\right) \sum_{n=0}^{|a|m-1} \overline{\chi(n)} \exp\left(-\frac{\pi i(cn^2 + bn)}{am}\right).$$

Naturally, we wish to evaluate the twisted Gauss sums occurring on each side of the expression above, as we evaluated the usual Gauss sums, but this turns out to be a difficult task. To simplify matters, for the remainder of this section we deal with the sums obtained by taking a and c to be positive and making the substitutions $a \mapsto 2a$ and $b \mapsto 2b$, for which the reciprocity relation reads:

$$\frac{G(\overline{\chi})}{\sqrt{|c|m}} \sum_{n=0}^{|c|m-1} \chi(n) \exp\left(\frac{2\pi i(an^2 + bn)}{cm}\right) = \sqrt{\frac{1}{|a|}} \exp\left(\frac{\pi i}{4}\left(\operatorname{sgn} a - \frac{b^2}{4acm}\right)\right) \sum_{n=0}^{2|a|m-1} \overline{\chi(n)} \exp\left(-\frac{\pi i(cn^2 + 2bn)}{2am}\right). \tag{4.50}$$

Note that the sum at 4.50 vanishes unless χ is even. We will often require the twisted Gauss sums in the case that χ is a Legendre symbol: therefore, in analogy with Section 4.3, we make the following definition for $p = 1 \bmod 4$:

$$\phi_p(a,b) = \sum_{n=0}^{bp-1} \left(\frac{n}{p}\right) \exp\left(\frac{2\pi i n^2 a}{bp}\right).$$

In Subsection 4.8.1 we rewrite $\phi_p(a,b)$ in terms of more well-known Gauss sums, and conclude the evaluation using the following beautiful result (originally a conjecture of Loxton [Lox76; Lox78]) proved by Matthews [Mat79b] using modular forms (see also [BE81; BE82] for notation and equivalent versions):

Theorem 4.8.0.1 (Matthews, 1979)**.** *Let $p = 1 \bmod 4$ be a prime and let χ_p be a Dirichlet character modulo p of exact order 4. Write $p = \alpha^2 + \beta^2$, where α is uniquely determined by $\alpha = -1 \bmod 4$ and the sign of β is determined by the Jacobi sum of χ:*

$$J(\chi) := \sum_{n \bmod p} \chi_p(n)\chi_p(1-n) = \alpha + i\beta.$$

Then

$$G(\chi_p) = C_p \left(\frac{|\beta|}{|\alpha|}\right)(-1)^{(\beta^2+2\beta)/8}\left(\sqrt{\frac{p+\alpha p}{2}} + i\,\mathrm{sgn}(\beta)\sqrt{\frac{p-\alpha\sqrt{p}}{2}}\right),$$

where the square roots are taken to have positive real part, $\left(\frac{|\beta|}{|\alpha|}\right)$ is the Legendre symbol and $C_p \in \{-1, +1\}$ is determined by:

$$C_p = \frac{|\beta|}{\alpha}\left(\frac{p-1}{2}\right)!\ (-1)^{(p-1)/4} \bmod p.$$

In Subsection 4.8.3 we use the observation that, for all integers b and odd primes p with $(b, p) = 1$,

$$\phi_p(b,1) = \sum_{n=0}^{p-1} \exp\left(\frac{2\pi i n^4 b}{p}\right) - \sum_{n=0}^{p-1} \exp\left(\frac{2\pi i n^2 b}{p}\right), \tag{4.51}$$

to link $\phi_p(a, b)$, for certain values of a and b, to the fourth-order Gauss sums

$$s_4(a,b) = \sum_{n=0}^{b-1} \exp\left(\frac{\pi i n^4 a}{b}\right).$$

Using the Landsberg–Schaar relation for $\phi_p(a, b)$, we deduce a similar reciprocity-type identity for $s_4(a, b)$.

In Subsections 4.8.4 and 4.8.5, we investigate analogues of the Landsberg–Schaar relation in the cases for which the square is replaced by a sixth or eighth power. These relations are significantly more complicated than the analogous identities for their quartic cousins.

4.8.1 Evaluating twisted Gauss sums

This subsection is devoted to evaluating quadratic Gauss sums twisted by Legendre symbols. Through the rest of this section, we use the symbol ϵ_a from Proposition 4.3.2.1:

$$\epsilon_a = \begin{cases} 1 & a = 1 \bmod 4 \\ i & a = 3 \bmod 4 \end{cases}$$

Clearly there is no loss of generality if we assume that $(a, b) = 1$. We may also assume that p divides neither a nor b:

Lemma 4.8.1.1. *If $p \mid a$ or $p \mid b$, then*

$$\phi_p(b,a) = 0.$$

Proof. Suppose $p \mid a$ (so we may assume that $p \nmid b$). Write $a = p^i a'$, where $(a', p) = 1$. Then

$$\phi_p(b,a) = \sum_{n=0}^{a'p^{i+1}-1} \left(\frac{n}{p}\right) \exp\left(\frac{2\pi i n^2 b}{a'p^{i+1}}\right).$$

If $s = 0, \ldots, p^{i+1} - 1 \bmod p^{i+1}$ and $t = 0, \ldots, a' - 1 \bmod a'$, then $a's + p^{i+1}t = 0, \ldots, a'p^{i+1} - 1 \bmod a'p^{i+1}$, so

$$\phi_p(b,a) = \left(\frac{a'}{p}\right) \sum_{s=0}^{p^{i+1}-1} \left(\frac{s}{p}\right) \exp\left(\frac{2\pi i s^2 a' b}{p^{i+1}}\right) \sum_{t=0}^{a'-1} \exp\left(\frac{2\pi i p^{i+1} b t^2}{a'}\right).$$

Write $s = kp^i + l$, where k runs from 0 to p and l runs from 0 to $p^i - 1$. Then the leftmost sum may be unwrapped to reveal a geometric series:

$$\sum_{s=0}^{p^{i+1}-1} \left(\frac{s}{p}\right) \exp\left(\frac{2\pi i s^2 a' b}{p^{i+1}}\right) = \sum_{k=0}^{p-1}\sum_{l=0}^{p-1} \left(\frac{l}{p}\right) \exp\left(\frac{2\pi i (kp^i + l)^2 a' b}{p^{i+1}}\right)$$
$$= \sum_{l=0}^{p^i-1} \left(\frac{l}{p}\right) \left(\sum_{k=0}^{p-1} \exp\left(\frac{4\pi i k l a' b}{p}\right)\right) \exp\left(\frac{2\pi i l^2 a' b}{p^{i+1}}\right).$$

Indeed, the inner sum is a geometric series, which vanishes unless p divides $2la'b$, which implies that p divides l. But in this case, $\left(\frac{l}{p}\right) = 0$, so the entire expression vanishes.

Now suppose $b \mid p$. Write $b = b'p$: p may divide b'. We have

$$\phi_p(b, a) = \sum_{n=0}^{ap-1} \left(\frac{n}{p}\right) \exp\left(\frac{2\pi i b' p n^2}{ap}\right),$$

and writing $n = as + pt$, where a runs from 0 to $p - 1$ and t runs from 0 to $a - 1$, this leads to

$$\phi_p(b, a) = \sum_{s=0}^{p-1} \left(\frac{s}{p}\right) \exp\left(\frac{2\pi i a b' p s^2}{p}\right) \sum_{t=0}^{a-1} \exp\left(\frac{2\pi i b' p^2 t^2}{a}\right)$$

and the outer sum is $\sum_{n \bmod p} \left(\frac{n}{p}\right) = 0$. □

Proposition 4.8.1.2. *For $p = 1 \bmod 4$ a prime, a and b positive, $(a, b, p) = 1$, and the same notation as in Theorem 4.8.0.1,*

$$\phi_p(b, a) = s'(bp, a)\left(\frac{a}{p}\right) C_p\left(\frac{|\beta|}{|\alpha|}\right)(-1)^{\frac{\beta^2+2|\beta|}{8}} \operatorname{Re}\left(\overline{\chi_p(ab)}\left(1 + \left(\frac{2}{p}\right)\operatorname{sgn}(\beta) i\right)\right)\sqrt{2\left(\frac{2}{p}\right)ap + 2\left(\frac{ab}{p}\right)\alpha a\sqrt{p}} \qquad (4.52)$$

where

$$s'(bp, a) = \begin{cases} \left(\frac{bp}{a}\right)\epsilon_a & a \text{ odd,} \\ 0 & a = 2 \bmod 4, \\ \left(\frac{a}{bp}\right)(1 + i^{bp}) & a = 0 \bmod 4. \end{cases}$$

Remark 4.8.1.3. *The reader may protest that the evaluation given in Proposition 4.8.1.2 depends on the choice of character χ_p. However, this is not so, as the only other character of exact order 4 modulo p is $\overline{\chi_p}$, and upon replacing χ_p by $\overline{\chi_p}$, β is replaced by $-\beta$, so $\operatorname{Im}(\chi_p(a)\overline{\chi_p(b)}\operatorname{sgn}(\beta))$ is unchanged.*

Remark 4.8.1.4. *The most mysterious term in Equation 4.52 under the square root sign is the coefficient of $\sqrt{p}$. In fact, for some fourth root of unity μ_4,*

$$\phi_p(1, 1) = \mu_4\sqrt{2a\left(p + \alpha\left(\frac{2}{p}\right)\sqrt{p}\right)},$$

and we may recognise $\alpha\left(\frac{2}{p}\right)$ as $-\chi(p)$, where $\chi(p)$ temporarily stands for Glaisher's chi function. Glaisher [Gla84] defined his chi function for odd n by

$$\chi(n) = (-1)^{(a_1+b_1-1)/2} 2a_1 + (-1)^{(a_2+b_2-1)/2} 2a_2 + \dots,$$

where $n = a_1^2 + b_1^2 = a_2^2 + b_2^2 = \dots$, and $a_1, a_2, \dots$ must be odd. He investigated the relationship between $\chi(n)$ and $E(n)$, the excess of the number of divisors of n congruent to 1 modulo 4 over the number of divisors congruent to 3 modulo 4, and expressed $\chi(n)$ in terms of elliptic functions. Indeed, $\chi(4n+1)$ is the coefficient of q^{4n+1} in

$$(q - 3q^9 + 5q^{25} - 7q^{49} + \dots)(1 - 2q^4 + 2q^{16} - 2q^{36} + 2q^{64} - \dots),$$

and $\sum_{n=1}^{\infty} \chi(n) q^n$ is the unique newform of weight 2 and level 32.

The proof of Proposition 4.8.1.2 rests on the following lemma, which links $\phi_p(b,a)$ to more usual Gauss sums.

Lemma 4.8.1.5. *Under the hypotheses of Proposition 4.8.1.2 together with the assumption that* $(a,p)=1$,

$$\phi_p(b,a) = \left(\frac{a}{p}\right) s'(bp,a)\left(\overline{\chi_p(ab)}G(\chi_p) + \chi_p(ab)G(\overline{\chi_p})\right).$$

Proof. By a simple counting argument, we readily establish 4.51, stated at the beginning of the section:

$$\sum_{n=0}^{p-1}\left(\frac{n}{p}\right)\exp\left(\frac{2\pi i n^2 b}{p}\right) + \sum_{n=0}^{p-1}\exp\left(\frac{2\pi i n^2 b}{p}\right) = 2\sum_{\substack{n=0\\ n=\square}}^{p-1}\exp\left(\frac{2\pi i n^2 b}{p}\right) = \sum_{n=0}^{p-1}\exp\left(\frac{2\pi i n^4 b}{p}\right). \tag{4.53}$$

As n runs through a complete residue system modulo p, n^4 runs through the set of 4th power residues 4 times, so

$$\begin{aligned}\sum_{n=0}^{p-1}\left(\frac{n}{p}\right)\exp\left(\frac{2\pi i n b}{p}\right) + \sum_{n=0}^{p-1}\left(\chi_p(n)+\overline{\chi_p(n)}\right)\exp\left(\frac{2\pi i n b}{p}\right) &= \sum_{n=0}^{p-1}\left(1+\chi_p(n)+\chi_p^2(n)+\chi_p^3(n)\right)\exp\left(\frac{2\pi i n b}{p}\right)\\ &= \sum_{n=0}^{p-1}\exp\left(\frac{2\pi i n^4 b}{p}\right) \end{aligned}\tag{4.54}$$

But

$$\sum_{n=0}^{p-1}\exp\left(\frac{2\pi i n^2 b}{p}\right) = \sum_{n=0}^{p-1}\left(\frac{n}{p}\right)\exp\left(\frac{2\pi i n b}{p}\right),$$

so combining the expressions 4.53 and 4.54,

$$\sum_{n=0}^{p-1}\left(\frac{n}{p}\right)\exp\left(\frac{2\pi i n^2 b}{p}\right) = \sum_{n=0}^{p-1}\left(\chi_p(n)+\overline{\chi_p(n)}\right)\exp\left(\frac{2\pi i n b}{p}\right). \tag{4.55}$$

Now as s runs through a complete residue system modulo p and t runs through a complete residue system modulo a, $as+pt$ runs through a complete residue system modulo ap. So using 4.55,

$$\begin{aligned}\phi_p(a,b) &= \sum_{n=0}^{ap-1}\left(\frac{n}{p}\right)\exp\left(\frac{2\pi i n^2 b}{ap}\right)\\ &= \left(\frac{a}{p}\right)\sum_{s=0}^{p-1}\left(\frac{s}{p}\right)\exp\left(\frac{2\pi i s^2 ab}{p}\right)\sum_{t=0}^{a-1}\exp\left(\frac{2\pi i t^2 bp}{a}\right)\\ &= \left(\frac{a}{p}\right)s'(bp,a)\left(\sum_{s=0}^{p-1}\chi_p(s)\exp\left(\frac{2\pi i s ab}{p}\right) + \sum_{s=0}^{p-1}\overline{\chi_p(s)}\exp\left(\frac{2\pi i s ab}{p}\right)\right)\\ &= \left(\frac{a}{p}\right)s'(bp,a)\left(\overline{\chi_p(ab)}\sum_{s=0}^{p-1}\chi_p(s)\exp\left(\frac{2\pi i s}{p}\right) + \chi_p(ab)\sum_{s=0}^{p-1}\overline{\chi_p(s)}\exp\left(\frac{2\pi i s}{p}\right)\right),\end{aligned}$$

as required. □

Now that we have proved Lemma 4.8.1.5, we can apply Theorem 4.8.0.1 to evaluate the Gauss sums.

Proof of Proposition 4.8.1.2. Matthews' Theorem implies that

$$G(\chi_p) = C_p\left(\frac{|\beta|}{|\alpha|}\right)(-1)^{(\beta^2+2\beta)/8}\left(\sqrt{\frac{p+\alpha p}{2}} + i\,\mathrm{sgn}(\beta)\sqrt{\frac{p-\alpha\sqrt{p}}{2}}\right),$$

$$G(\overline{\chi_p}) = C_p\left(\frac{|\beta|}{|\alpha|}\right)(-1)^{(\beta^2-2\beta)/8}\left(\sqrt{\frac{p+\alpha p}{2}} - i\,\mathrm{sgn}(\beta)\sqrt{\frac{p-\alpha\sqrt{p}}{2}}\right),$$

and noting that

$$\left(\frac{2}{p}\right) = (-1)^{\beta/2},$$

we have

$$\overline{\chi_p(ab)}G(\chi_p) + \chi_p(ab)G_p(\overline{\chi_p}) = C_p\left(\frac{|\beta|}{|\alpha|}\right)(-1)^{\frac{\beta^2+2\beta}{8}}S(a,b,p),$$

where

$$S(a,b,p) = \overline{\chi_p(ab)}\left(\sqrt{\frac{p+\alpha p}{2}} + i\,\mathrm{sgn}(\beta)\sqrt{\frac{p-\alpha\sqrt{p}}{2}}\right) + \left(\frac{2}{p}\right)\chi_p(ab)\left(\sqrt{\frac{p+\alpha p}{2}} - i\,\mathrm{sgn}(\beta)\sqrt{\frac{p-\alpha\sqrt{p}}{2}}\right).$$

Now an examination of $S(a,b,p)$ as $\left(\frac{2}{p}\right)$ and $\overline{\chi_p(ab)}$ vary over $\{+1,-1\}$ and $\{+1,+i,-1,-i\}$ respectively, allows us to identify $S(a,b,p)$ up to a sign:

$$S(a,b,p) = T(a,b,p)\sqrt{2\left(\frac{2}{p}\right)p + 2\alpha\left(\frac{ab}{p}\right)\sqrt{p}},$$

where $T(a,b,p)$ takes values in $\{+1,-1\}$:

		$\overline{\chi_p(ab)}$: $+1$	$+i$	-1	$-i$
$\left(\frac{2}{p}\right)$	$+1$	$+$	$-\,\mathrm{sgn}(\beta)$	$-$	$+\,\mathrm{sgn}(\beta)$
	-1	$+\,\mathrm{sgn}(\beta)$	$+$	$-\,\mathrm{sgn}(\beta)$	$-$

A straightforward calculation using $\left(\frac{2}{p}\right) = (-1)^{\beta/2}$ reveals that

$$(-1)^{\frac{\beta}{4}}T(a,b,p) = (-1)^{\frac{|\beta|}{4}}U(a,b,p),$$

where $U(a,b,p)$ takes values in $\{+1,-1\}$ as described below:

		$\overline{\chi_p(ab)}$: $+1$	$+i$	-1	$-i$
$\left(\frac{2}{p}\right)$	$+1$	$+$	$-\,\mathrm{sgn}(\beta)$	$-$	$+\,\mathrm{sgn}(\beta)$
	-1	$+$	$+\,\mathrm{sgn}(\beta)$	$-$	$-\,\mathrm{sgn}(\beta)$

It's easily checked that

$$U(a,b,p) = \mathrm{Re}\left(\overline{\chi_p(ab)}\left(1+\left(\frac{2}{p}\right)\mathrm{sgn}(\beta)i\right)\right),$$

so the proposition is proved. □

As a finale, we evaluate the quartic Gauss sum in all its glory.

Corollary 4.8.1.6. *Write*

$$a = 2^\tau \prod_{\substack{p_i|a \\ p_i=1 \bmod 4}} p_i^{\mu_i} \prod_{\substack{q_j|a \\ q_j=3 \bmod 4}} q_j^{\nu_j},$$

where p_i and q_j are prime and τ, μ_i and ν_j are all positive. Furthermore, write $\tau = 4k_\tau + l_\tau$, $\mu_i = 4k_{\mu_i} + l_{\mu_i}$ and $\nu_j = 4k_{\nu_j} + l_{\nu_j}$, where $k_\tau, k_{\mu_i}, k_{\nu_j} \geq 0$ and $1 \leq l_\tau, l_{\mu_i}, l_{\nu_j} \leq 4$. Then using the notation from Theorem 4.8.0.1,

$$s_4(2b,a) = 2^{3k_\tau + l_\tau - 1} 2^{\left[3\sum_i k_{\mu_i} + 3\sum_j k_{\nu_j} + \sum_{\substack{i\\ l_{\mu_i}\neq 1}} (l_{\mu_i}-1) + \sum_{\substack{j \\ l_{\nu_j}\neq 1}} (l_{\nu_j}-1)\right]} \prod_{\substack{j\\ l_j=1}} \epsilon_{q_j}\sqrt{q_j}\left(\frac{2^{3\tau}\prod_i p_i^{3\mu_i}\prod_{t\neq j} q_t^{3\nu_t} b}{q_j}\right)$$

$$\prod_{\substack{i\\ l_i=1}} \left[\epsilon_{p_i}\sqrt{p_i}\left(\frac{2^{3\tau}\prod_{s\neq i} p_s^{3\mu_s}\prod_j q_j^{3\nu_j} b}{p_i}\right) + C_{p_i}\left(\frac{|\beta_{p_i}|}{|\alpha_{p_i}|}\right)(-1)^{\frac{{\beta_{p_i}}^2 + 2|\beta_{p_i}|}{8}} \mathrm{Re}\left(\overline{\chi_{p_i}(b)}\left(1 + \left(\frac{2}{p_i}\right)\mathrm{sgn}(\beta_{p_i}) i\right)\right) \right.$$

$$\left. \sqrt{2\left(\frac{2}{p_i}\right)p_i + 2\left(\frac{2^{3\tau}\prod_{s\neq i} p_s^{3\mu_s}\prod_j q_j^{3\nu_j} b}{p}\right)\alpha_{p_i}\sqrt{p_i}}\right].$$

We require some lemmas. The first is a straightforward generalisation of Lemma 4.3.3.2 (or a special case of Lemma 4.7.0.1), and allows us to concentrate on evaluating $s_4(2b,a)$ in the case that a is a prime power.

Lemma 4.8.1.7.

$$s_4(2b,a) = s_4\left(\prod_i p_i^{3\mu_i}\prod_j q_j^{3\nu_j} 2b, 2^\tau\right)\prod_i s_4\left(2^{3\tau}\prod_{s\neq i} p_s^{3\mu_s}\prod_j q_j^{3\nu_j} 2b, p_i^{\mu_i}\right)\prod_j s_4\left(2^{3\tau}\prod_i p_i^{3\mu_i}\prod_{t\neq j} q_t^{3\nu_t} 2b, q_j^{\nu_j}\right).$$

Proof. As l_τ runs over a complete system of residues modulo 2^τ, and for each i and j, n_i and m_j run over complete systems of residues modulo $p_i^{\mu_i}$ and $q_j^{\nu_j}$, respectively,

$$n = l_\tau \prod_i p_i^{\mu_i}\prod_j q_j^{\nu_j} + \sum_i n_i 2^\tau \prod_{s\neq i} p_s^{\mu_s}\prod_j q_j^{\nu_j} + \sum_j m_j 2^\tau \prod_i p_i^{\mu_i}\prod_{t\neq j} q_t^{\nu_t}$$

runs over a complete system of residues modulo a. Furthermore,

$$n^4 = l_\tau^4 \prod_i p_i^{4\mu_i}\prod_j q_j^{4\nu_j} + \sum_i n_i^4 2^{4\tau} \prod_{s\neq i} p_s^{4\mu_s}\prod_j q_j^{4\nu_j} + \sum_j m_j^4 2^{4\tau} \prod_i p_i^{4\mu_i}\prod_{t\neq j} q_t^{4\nu_t} \bmod a,$$

so

$$s_4(2b,a) = \sum_{n=0}^{a-1} \exp\left(\frac{2\pi i n^4 b}{a}\right)$$

splits into sums over l_τ, n_i and m_j as required. □

The reader will notice that the next lemma is a special case of the results of one of the conjectures in Section 4.7.

Lemma 4.8.1.8. *Suppose b is odd and $\tau \geq 1$. Write $\tau = 4k + l$, where $k \geq 0$, $1 \leq l \leq 4$, but one may take $l = 0$ if $k = 0$. Then*

$$s_4(2b, 2^\tau) = 2^{3k+l-1}\left(1 + \exp\left(\frac{\pi i b}{2^{l-1}}\right)\right).$$

Proof. It suffices to show that, for $\tau \geq 5$,

$$s_4(2b, 2^\tau) = 2^{3k} s_4(2b, 2^l). \tag{4.56}$$

To prove 4.56, we split the sum defining $s_4(2b, 2^\tau)$ into two sums, one over the odd indices and one over the even ones, and show that the sum over odd indices vanishes.

$$\begin{aligned}
s_4(2b, 2^\tau) &= \sum_{\substack{n=0 \\ n \text{ odd}}}^{2^\tau-1} \exp\left(\frac{2\pi i n^4 b}{2^\tau}\right) + \sum_{\substack{n=0 \\ n \text{ even}}}^{2^\tau-1} \exp\left(\frac{2\pi i n^4 b}{2^\tau}\right) \\
&= \sum_{\substack{n=0 \\ n \text{ odd}}}^{2^\tau-1} \exp\left(\frac{2\pi i (n+2^{\tau-3})^4 b}{2^\tau}\right) + \sum_{\substack{n=0 \\ n \text{ even}}}^{2^\tau-1} \exp\left(\frac{2\pi i n^4 b}{2^\tau}\right) \\
&= -\sum_{\substack{n=0 \\ n \text{ odd}}}^{2^\tau-1} \exp\left(\frac{2\pi i n^4 b}{2^\tau}\right) + 2^3 s_4(2b, 2^{4(k-1)+l}) \\
&= 2^3 s_4(2b, 2^{4(k-1)+l}).
\end{aligned}$$

To conclude,

$$s_4(2b, 2^{4k+l}) = 2^3 s_4(2b, 2^{4(k-1)+l}) = \cdots = 2^{3k} s_4(2b, 2^l). \qquad \square$$

The next lemma is also a special case of the results of Section 4.7, and for the proof, we refer the reader to Proposition 4.7.1.1.

Lemma 4.8.1.9. *Suppose p is an odd prime, and $(b,p) = 1$. Write $\alpha = 4k + l$, where $k \geq 0$ and $1 \leq l \leq 4$. Then*

$$s_4(2b, p^\alpha) = \begin{cases} p^{3k} s_4(2b, p) & l = 1 \\ p^{3k+l-1} & l \neq 1. \end{cases}$$

The next lemma is a specialisation of Lemma 4.7.2.1, but it has a quaint and well-known proof in this special case.

Lemma 4.8.1.10. *Let p be a prime with $p = 3 \bmod 4$ and $(b,p) = 1$. Then*

$$s_4(2b, p) = s_2(2b, p) = s'(b, p)$$

Proof. It suffices to show that if $p = 3 \bmod 4$, then every square in $\mathbb{Z}/p\mathbb{Z}$ is a fourth power. Consider the map $\Upsilon : (\mathbb{Z}/p\mathbb{Z})^\times \to (\mathbb{Z}/p\mathbb{Z})^\times$ defined by

$$\Upsilon(x) = x^4.$$

Then since $\left(\frac{-1}{p}\right) = -1$, Υ is a 2-1 map. So the set of all fourth powers in $(\mathbb{Z}/p\mathbb{Z})^\times$ has cardinality $(p-1)/2$. But this is exactly the cardinality of the set of squares in $(\mathbb{Z}/p\mathbb{Z})^\times$. $\square$

Proof of Corollary 4.8.1.6. Lemmas 4.8.1.8, 4.8.1.9 and 4.8.1.10 enable us to evaluate all the sums appearing in Lemma 4.8.1.7, except for those of length p, where $p = 1 \bmod 4$. Fortunately, by 4.51 these sums can be written as the sum of a quadratic Gauss sum and a term that has already been evaluated during the proof of Proposition 4.8.1.2. $\square$

Remark 4.8.1.11. *The methods employed in this subsection, together with the results of Section 4.7, may be used to evaluate the higher-degree Gauss sums*

$$s_d(2b, a) = \sum_{n=0}^{a-1} \exp\left(\frac{2\pi i n^d b}{a}\right),$$

where a factorises as $\prod_j q_j^{\nu_j}$ in terms of $s_{f_j}(2b, q_j)$, and f_j is the largest divisor of d such that $q_j = 1 \bmod f_j$. Additional effort, using techniques derived [BE81] from the observation that $s_{f_j}(2b, q_j)$ may be expressed in terms of the Gauss sums

$$G(\chi_{f_j}) = \sum_{n \bmod f_j} \chi(n) \exp\left(\frac{2\pi i n}{f_j}\right),$$

where χ_{q_j} are the Dirichlet characters modulo q_j of order f_j, allows for the explicit evaluation of each $s_{f_j}(2b, q_j)$, and hence the original sum $s_d(2b, a)$, up to the determination of some f_jth roots of unity. We have not carried this out because there is still no "explicit formula" for the $G(\chi_{f_j})$ for any f_j other than two or four.

4.8.2 From higher-degree Gauss sums to twisted quadratic Gauss sums

As in the last subsection, we set our character χ to be the Legendre symbol for an odd prime $p = 1 \bmod 4$, and concentrate on the identity

$$\frac{1}{\sqrt{a}} \sum_{n=0}^{ap-1} \left(\frac{n}{p}\right) \exp\left(\frac{2\pi i n^2 b}{ap}\right) = \sqrt{\frac{i}{2b}} \sum_{n=0}^{2bp-1} \left(\frac{n}{p}\right) \exp\left(-\frac{\pi i n^2 a}{2bp}\right). \tag{4.57}$$

Since the twisted Gauss sums involved in this identity seem to be saying something about whether squares modulo ap are themselves squares modulo p, it is natural to look for a connection with fourth powers modulo p. This connection is provided by 4.51, stated at the beginning of the section, proved in Lemma 4.8.1.5, and valid for $(b,p) = 1$:

$$\sum_{n=0}^{p-1} \left(\frac{n}{p}\right) \exp\left(\frac{2\pi i n^2 b}{p}\right) + \sum_{n=0}^{p-1} \exp\left(\frac{2\pi i n^2 b}{p}\right) = \sum_{n=0}^{p-1} \exp\left(\frac{2\pi i n^4 b}{p}\right). \tag{4.58}$$

In order to be able to apply 4.57 to produce a quartic version of the Landsberg–Schaar relation, we need to be able to convert between fourth-power Gauss sums and twisted quadratic Gauss sums. Proposition 4.8.2.3, a version of 4.58 containing a and b, provides the means to do this. We prepare the way to Proposition 4.8.2.3 with the next result.

Lemma 4.8.2.1. *Suppose $d \geq 2$ is even. Write $\alpha = dk + l$, where $k \geq 0$ and $1 \leq l \leq d$. Then if p is an odd prime such that $p \neq 1 \mod q$ for all odd primes q dividing d and $p \nmid d$,*

$$s_d(2b, p^\alpha) = f_d(2b, p^\alpha) s_2(2b, p^\alpha),$$

where

$$f_d(2b, p^\alpha) = \begin{cases} p^{(d-1)k-(\alpha-1)/2} & l = 1 \\ \left(\frac{b}{p}\right)\epsilon_p^{-1} p^{(d-1)k+l-\alpha/2} & l \neq 1, \ \ l \text{ odd}, \\ p^{(d-1)k+l-1-\alpha/2} & l \neq 1, \ \ l \text{ even}, \end{cases}$$

and if $4 \mid d$, we also stipulate that $p \neq 1 \bmod 4$. We extend $f_{2d',p}(2b,c)$ to odd c by defining

$$f_d(2b, c) = \prod_j f_d\left(2b \prod_{k \neq j} \left((q_k)_{q_j}^{-1}\right)^{\nu_k}, q_j^{\nu_j}\right),$$

where c is a product of distinct odd prime powers

$$c = \prod_j q_j^{\nu_j},$$

and $(q_k)_{q_j}^{-1}$ is a multiplicative inverse of p modulo q_j. Furthermore, if C is divisible by 2 but not by 4, then we may write $C = 2c$ where c is odd, and we define

$$f_d(b, C) = f_d(2b, c).$$

Remark 4.8.2.2. *The extension of $f_d(2b,c)$ to all odd c is motivated by the desire for a version of Lemma 4.8.2.1 valid for odd c. Indeed, with the notation of Lemma 4.8.2.1 and the added hypothesis that $c \nmid d$, together with Lemma*

4.7.0.1, we have

$$\begin{aligned}
s_{2d'}(2b,c) &= \prod_j s_{2d'}\left(2b\prod_{k\neq j} q_k^{\nu_k(2d'-1)}, q_j^{\nu_j}\right)\\
&= \prod_j s_{2d'}\left(2b\prod_{k\neq j}\left((q_k)_{q_j}^{-1}\right)^{\nu_k}, q_j^{\nu_j}\right)\\
&= \prod_j f_{2d'}\left(2b\prod_{k\neq j}\left((q_k)_{q_j}^{-1}\right)^{\nu_k}, q_j^{\nu_j}\right) s_2\left(2b\prod_{k\neq j}\left((q_k)_{q_j}^{-1}\right)^{\nu_k}, q_j^{\nu_j}\right)\\
&= f_{2d'}(2b,c)\prod_j s_2\left(2b\prod_{k\neq j} q_j^{\nu_j}, q_j^{\nu_j}\right)\\
&= f_{2d'}(2b,c)s_2(2b,c).
\end{aligned}$$

During the proof of Lemma 4.8.2.1, we will see that the definition of $f_d(b,c)$ for even c has been chosen so as to match up smoothly with a similar product relation among the degree d Gauss sums.

Proof of Lemma 4.8.2.1. Recall that we have evaluated the degree d Gauss sums for odd prime powers $p \nmid d$ in Proposition 4.7.1.1:

$$s_d(2b,p^\alpha) = \begin{cases} p^{(d-1)k}s_2(2b,p) & l=1\\ p^{(d-1)k+l-1} & l\neq 1.\end{cases}$$

If we set $d=2$, then we obtain

$$s_2(2b,p^\alpha) = \begin{cases} p^{(\alpha-1)/2}s_2(2b,p) & \alpha \text{ even},\\ p^{\alpha/2} & \alpha \text{ odd}.\end{cases}$$

By assumption, if $q>2$ divides d, then $p\neq 1 \bmod q$, so by Lemma 4.7.2.1, $s_d(2b,p)=s_2(2b,p)$. Upon scrutinising the quotient $s_d(2b,p^{\alpha)})/s_2(2b,p^{\alpha)})$, we find that we need only match up powers of p, excepting the case in which $l\neq 1$ is odd, where we find a factor of the form $(s_2(2b,p))^{-1}$. Using Proposition 4.3.2.1, we arrive at the claimed expression. □

Proposition 4.8.2.3. *Let $d'\geq 1$ and let p be a prime with $p=1 \bmod 2d'$. Let a,b and c be integers, with a and c coprime and nonzero, $(a,c,p)=1$, and suppose $4\nmid c$. Suppose that a and c have opposite parity, a and b have the same parity, none of the odd prime factors of c are congruent to $1 \bmod q$ for any odd prime q that divides d', and that $(c,d')=1$. Let χ_p be any primitive Dirichlet character modulo p of exact order d'.*

If d' is even, suppose also that no prime factor of c is congruent to $1 \bmod 4$. Then

$$\sum_{n=0}^{cp-1}\exp\left(\frac{\pi i(an^{2d'}+bcn^{d'})}{cp}\right) = f_{2d'}(ap_c^{-1},c)\sum_{k=0}^{d'-1}\sum_{n=0}^{cp-1}\chi_p^k(n)\exp\left(\frac{\pi i(an^2+bcn)}{cp}\right), \tag{4.59}$$

where $f_{2d'}(a,q_j^{\nu_j})$ is the function defined in Lemma 4.8.2.1.

Proof. First, suppose c is odd. Then we may assume that a and b are even. By Remark 4.8.2.2,

$$s_{2d'}(a,c) = f_{2d'}(a,c)s_2(a,c).$$

Then we can prove 4.59 by splitting the sum in two, one over c and one over p.

$$\begin{aligned}
\sum_{n=0}^{cp-1}\exp\left(\frac{\pi i(an^{2d'}+bcn^{d'})}{cp}\right) &= \sum_{n=0}^{cp-1}\exp\left(\frac{\pi i(ap^{2d'-1}n^{2d'}+bp^{d'-1}cn^{d'})}{c}\right)\sum_{n=0}^{p-1}\exp\left(\frac{\pi i(ac^{2d'-1}n^{2d'}+bc^{d'}n^{d'})}{p}\right)\\
&= s_{2d'}(ap_c^{-1},c)\sum_{n=0}^{p-1}\exp\left(\frac{\pi i(ac_p^{-1}n^{2d'}+bn^{d'})}{p}\right)\\
&= s_{2d'}(ap_c^{-1},c)\sum_{k=0}^{d'-1}\sum_{n=0}^{p-1}\chi_p^k(n)\exp\left(\frac{\pi i(ac_p^{-1}n^2+bn)}{p}\right)\\
&= f_{2d'}(ap_c^{-1},c)s_2(ap_c^{-1},c)\sum_{k=0}^{d'-1}\sum_{n=0}^{p-1}\chi_p^k(cn)\exp\left(\frac{\pi i(acn^2+bcn)}{p}\right)\\
&= f_{2d'}(ap_c^{-1},c)\sum_{k=0}^{d'-1}\left(\sum_{n=0}^{c-1}\exp\left(\frac{\pi i(ap_c^{-1}n^2+bcp_c^{-1}n)}{c}\right)\right)\left(\sum_{n=0}^{p-1}\chi_p^k(cn)\exp\left(\frac{\pi i(acn^2+bcn)}{p}\right)\right)\\
&= f_{2d'}(ap_c^{-1},c)\sum_{k=0}^{d'-1}\left(\sum_{n=0}^{c-1}\exp\left(\frac{\pi i(apn^2+bcn)}{c}\right)\right)\left(\sum_{n=0}^{p-1}\chi_p^k(cn)\exp\left(\frac{\pi i(acn^2+bcn)}{p}\right)\right)\\
&= f_{2d'}(ap_c^{-1},c)\sum_{k=0}^{d'-1}\sum_{n=0}^{cp-1}\chi_p^k(n)\exp\left(\frac{\pi i(an^2+bcn)}{cp}\right).
\end{aligned}$$

Lastly, we suppose that c is divisible by 2 and reduce this case to the one proved above. Suppose that $c = 2C$, where a, b, C and p satisfy all the hypotheses of Proposition 4.8.2.3. As s and t run over complete moduli classes modulo 2 and Cp respectively, $n = 2s + Cpt$ runs over a complete moduli class modulo cp. Employing Proposition 4.8.2.3 for Cp and permuting the summation range,

$$\begin{aligned}
\sum_{n=0}^{cp-1}\exp\left(\frac{\pi i(an^{2d'}+bcn^{d'})}{cp}\right) &= \sum_{n=0}^{2Cp-1}\exp\left(\frac{\pi i(as^{2d'}+2bCs^{d'})}{2Cp}\right)\\
&= \sum_{s=0}^{1}\exp\left(\frac{\pi i(a(Cp)^{2d'-1}s^{2d'}+2bC(Cp)^{d'-1}s^{d'})}{2}\right)\sum_{t=0}^{Cp-1}\exp\left(\frac{\pi i(a2^{2d'-1}t^{2d'}+2bC2^{d'-1}t^{d'})}{Cp}\right)\\
&= \sum_{s=0}^{1}\exp\left(\frac{\pi i(a(Cp)_2^{-1}s^{2d'}+2bC(Cp)_2^{-1}s^{d'})}{2}\right)\sum_{t=0}^{Cp-1}\exp\left(\frac{\pi i(a2_{Cp}^{-1}t^{2d'}+2bC2_{Cp}^{-1}t^{d'})}{Cp}\right)\\
&= f_{2d'}(2a(2^{2d'-2})p_C^{-1},C)\sum_{k=0}^{d'-1}\sum_{s=0}^{1}\exp\left(\frac{\pi i(a(Cp)_2^{-1}s^2+2bC(Cp)_2^{-1}s)}{2}\right)\sum_{t=0}^{Cp-1}\chi_p^k(t)\exp\left(\frac{\pi i(a2_{Cp}^{-1}t^2+2bC2_{Cp}^{-1}t)}{Cp}\right)\\
&= f_{2d'}(2ap_C^{-1},C)\sum_{k=0}^{d'-1}\sum_{s=0}^{1}\exp\left(\frac{\pi i(a(Cp)^{-1}s^2+2bC(Cp)s)}{2}\right)\sum_{t=0}^{Cp-1}\chi_p^k(t)\exp\left(\frac{\pi i(a2t^2+2bC2t)}{Cp}\right)\\
&= f_{2d'}(ap_c^{-1},c)\sum_{k=0}^{d'-1}\sum_{n=0}^{cp-1}\chi_p^k(n)\exp\left(\frac{\pi i(an^2+bcn)}{cp}\right).
\end{aligned}$$

□

Remark 4.8.2.4. *If, in addition to the hypotheses of Proposition 4.8.2.3, we assume that c is squarefree (or more generally, that $\alpha = 2k$), then $f_{2d'}(ap_c^{-1}, c) = 1$.*

4.8.3 A generalised quartic Landsberg–Schaar relation

Recall the following (slightly simplified) version of the generalised twisted Landsberg–Schaar relation, given at 4.50.

$$\frac{G(\overline{\chi})}{\sqrt{|c|m}}\sum_{n=0}^{|c|m-1}\chi(n)\exp\left(\frac{2\pi i(an^2+bn)}{cm}\right) = \sqrt{\frac{1}{|a|}}\exp\left(\frac{\pi i}{4}\left(\operatorname{sgn} a - \frac{b^2}{4acm}\right)\right)\sum_{n=0}^{|2a|m-1}\overline{\chi(n)}\exp\left(-\frac{\pi i(cn^2+2bn)}{2am}\right). \quad (4.60)$$

We wish to use 4.60 to prove some sort of "generalised" analogue of the local reciprocity relation 4.65: that is, for sums of the form

$$\sum_{n=0}^{c-1} \exp\left(\frac{2\pi i}{c}(an^4 + bn^2)\right). \tag{4.61}$$

The question of finding such a reciprocity relation was listed as a research problem by Berndt, Evans and Williams in 1998 [BEW98, Research Problem 8, pp. 496]:

> *In Theorem 1.2.2, a reciprocity theorem*[8] *for generalized quadratic Gauss sums was established. Does there exist an analogous formula for generalized quartic Gauss sums, that is, sums of the type*
>
> $$\sum_{n=0}^{c-1} \exp\left(\frac{\pi i}{c}(an^4 + bn^2)\right),$$
>
> *where a, b and c are integers with $a, c > 0$?*

In 1972, Berndt used 4.60 to prove the following relation [Ber73], valid for b odd and p any odd prime:

$$\sum_{n=0}^{p-1} \exp\left(\frac{\pi i}{p}(n^4 + bn^2)\right) - \sqrt{p}\exp\left(\frac{\pi i(1-\frac{b^2}{p})}{4}\right) = \exp\left(\frac{\pi i(1-\frac{b^2}{p})}{4}\right)\sum_{n=0}^{p-1} \exp\left(-\frac{\pi i(n^4 + bn^2)}{p}\right) - \sqrt{p}\epsilon_p.$$

We now establish a *local* reciprocity relation for the sum at 4.61 in the case that ac divides b:

Proposition 4.8.3.1. *Let a and c be odd coprime integers, with all prime factors congruent to* 3 *modulo* 4*, let p be a prime congruent to* 1 *modulo* 4*, let b be any integer, and let τ be a multiplicative inverse for* 2 *modulo cp. Then*

$$\begin{aligned}\frac{f_4^{-1}(2ap_c^{-1}, c)}{\sqrt{cp}}\sum_{n=0}^{cp-1}\exp\left(\frac{2\pi i(an^4 + abcn^2)}{cp}\right) &= f_4^{-1}(-cp_{2a}^{-1}, 2a)\sqrt{\frac{i}{2ap}}\exp\left(-\frac{\pi i ab^2 c}{2p}\right)\sum_{n=0}^{2ap-1}\exp\left(-\frac{\pi i(cn^4 + 2abcn^2)}{2ap}\right) \\ &+ \epsilon_c\left(\frac{a}{c}\right)\left(\left(\frac{a}{p}\right)\exp\left(\frac{2\pi i ab^2 c}{p}\tau(\tau-1)\right) - \left(\frac{c}{p}\right)\right).\end{aligned} \tag{4.62}$$

Proof. Proposition 4.8.2.3, with $d' = 2$, yields

$$f_4^{-1}(2ap_c^{-1}, c)\sum_{n=0}^{cp-1}\exp\left(\frac{2\pi i(an^4 + abcn^2)}{cp}\right) = \sum_{n=0}^{cp-1}\exp\left(\frac{2\pi i(an^2 + abcn)}{cp}\right) + \sum_{n=0}^{cp-1}\left(\frac{n}{p}\right)\exp\left(\frac{2\pi i(an^2 + abcn)}{cp}\right), \tag{4.63}$$

along with

$$\begin{aligned} f_4^{-1}(-cp_{2a}^{-1}, 2a)&\sum_{n=0}^{2ap-1}\exp\left(-\frac{\pi i(cn^4 + 2abcn^2)}{2ap}\right) \\ &= \sum_{n=0}^{2ap-1}\exp\left(-\frac{\pi i(cn^2 + 2abcn)}{2ap}\right) + \sum_{n=0}^{2ap-1}\left(\frac{n}{p}\right)\exp\left(-\frac{\pi i(cn^2 + 2abcn)}{2ap}\right).\end{aligned} \tag{4.64}$$

The link between 4.63 and 4.64 is provided by 4.60,

$$\sum_{n=0}^{cp-1}\left(\frac{n}{p}\right)\exp\left(\frac{2\pi i(an^2 + abcn)}{cp}\right) = \sqrt{\frac{c}{2a}}\exp\left(\frac{\pi i}{4}\left(1 - \frac{2ab^2c}{p}\right)\right)\sum_{n=0}^{2ap-1}\left(\frac{n}{p}\right)\exp\left(-\frac{\pi i(cn^2 + 2bn)}{2ap}\right),$$

so if we can evaluate the quadratic Gauss sums occurring at 4.63 and 4.64, we will have a relation between the degree four Gauss sums. Indeed,

$$\begin{aligned}\sum_{n=0}^{cp-1}\exp\left(\frac{2\pi i(an^2 + abcn)}{cp}\right) &= \exp\left(\frac{2\pi i ab^2 c}{p}\tau(\tau-1)\right)\sum_{n=0}^{cp-1}\exp\left(\frac{2\pi i an^2}{cp}\right) \\ &= \epsilon_c\sqrt{cp}\left(\frac{a}{cp}\right)\exp\left(\frac{2\pi i ab^2 c}{p}\tau(\tau-1)\right),\end{aligned}$$

[8] This is Proposition 4.5.2.1 in the present text.

and

$$\sum_{n=0}^{2ap-1} \exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right) = \exp\left(\frac{\pi iab^2c}{2p}\right)\sum_{n=0}^{2ap-1}\exp\left(-\frac{\pi icn^2}{2ap}\right)$$
$$= \exp\left(\frac{\pi iab^2c}{2p}\right)\exp\left(-\frac{\pi i}{4}\right)\sum_{n=0}^{c-1}\exp\left(\frac{2\pi iapn^2}{c}\right)$$
$$= \exp\left(\frac{\pi i}{4}\right)\sqrt{2ap}\epsilon_c\left(\frac{ap}{c}\right)\exp\left(\frac{\pi iab^2c}{2p}\right).$$

Therefore

$$f_4^{-1}(2ap_c^{-1},c)\sum_{n=0}^{cp-1}\exp\left(\frac{2\pi i(an^4+abcn^2)}{cp}\right) - \sqrt{\frac{ic}{2a}}\exp\left(-\frac{\pi iab^2c}{2p}\right)f_4^{-1}(-cp_{2a}^{-1},2a)\sum_{n=0}^{2ap-1}\exp\left(-\frac{\pi i(cn^4+2abcn^2)}{2ap}\right)$$
$$= \epsilon_c\sqrt{cp}\left(\frac{a}{cp}\right)\exp\left(\frac{2\pi iab^2c}{p}\tau(\tau-1)\right) - \epsilon_c\sqrt{cp}\left(\frac{ap}{c}\right)$$
$$= \epsilon_c\sqrt{p}\left(\frac{a}{c}\right)\left(\left(\frac{a}{p}\right)\exp\left(\frac{2\pi iab^2c}{p}\tau(\tau-1)\right) - \left(\frac{p}{c}\right)\right)$$

The proposition follows upon noting that, with the specified congruence conditions on c and p, $\left(\frac{p}{c}\right) = \left(\frac{c}{p}\right)$ by quadratic reciprocity (Theorem 4.1.2.1). □

If we set $b = 0$ in 4.62 and swap the symbols a and c, we obtain a non-generalised quartic Landsberg–Schaar relation.

Corollary 4.8.3.2. *Let a, b and p be positive odd coprime integers, with all prime factors of a and b congruent to 3 modulo 4 and $p = 1 \bmod 4$ a prime. Then*

$$\frac{f_4^{-1}(2ap_c^{-1},c)}{\sqrt{ap}}\sum_{n=0}^{ap-1}\exp\left(\frac{2\pi in^4b}{ap}\right) = \sqrt{\frac{i}{2bp}}f_4^{-1}(-cp_{2a}^{-1},2a)\sum_{n=0}^{2bp-1}\exp\left(-\frac{\pi in^4a}{2bp}\right) + \epsilon_a\left(\frac{b}{a}\right)\left(\left(\frac{b}{p}\right)-\left(\frac{a}{p}\right)\right). \quad (4.65)$$

Remark 4.8.3.3. *If, in addition to the hypotheses of Corollary 4.8.3.2, we assume that a and c are squarefree, then we have a simpler relation.*

$$\frac{1}{\sqrt{ap}}\sum_{n=0}^{ap-1}\exp\left(\frac{2\pi in^4b}{ap}\right) = \sqrt{\frac{i}{2bp}}\sum_{n=0}^{2bp-1}\exp\left(-\frac{\pi in^4a}{2bp}\right) + \epsilon_a\left(\frac{b}{a}\right)\left(\left(\frac{b}{p}\right)-\left(\frac{a}{p}\right)\right).$$

4.8.4 A generalised sextic Landsberg–Schaar relation

Any attempt to prove versions of the Landsberg–Schaar relation for Gauss sums of even degree higher than four seems to result in markedly less elegant formulas. This state of affairs is essentially caused by 4.59:

$$\sum_{n=0}^{cp-1}\exp\left(\frac{\pi i(an^{2d'}+bcn^{d'})}{cp}\right) = f_{2d'}(ap_c^{-1},c)\sum_{k=0}^{d'-1}\sum_{n=0}^{cp-1}\chi_p^k(n)\exp\left(\frac{\pi i(an^2+bcn)}{cp}\right).$$

As d' increases, the number of terms to which we wish to apply the twisted reciprocity relation (4.50) accumulate, and since a "simple" evaluation of Gauss sums of degree higher than two is unknown (excepting degree 4), it is impossible to piece the resulting sums back together to form a higher-degree Gauss sum without generating inelegant "error terms". In the next two subsections, we present these inelegancies for sextic and octic higher-degree Gauss sums.

We deal with the sextic case first. As the octic case relies on the evaluation of $s_4(2,p)$ for a prime $p = 1 \bmod 4$, the sextic case relies on the evaluation of $s_3(2,p)$ for a prime $p = 1 \bmod 3$. Recall that Theorem 4.8.0.1 provides an explicit formula for $s_4(2,p)$ (but not a more computationally efficient one!). In fact, Matthews [Mat79a] proved an analogous result, originally a conjecture of Cassels [Cas69], for $s_3(2,p)$. We provide the details next.

Theorem 4.8.4.1 (Matthews, 1979 [Mat79a]). *Let $p = 1 \bmod 3$ be a prime and χ_p be a character modulo p of exact order 3. Let $J(\chi_p, \chi_p)$ be the Jacobi sum*

$$J(\chi_p, \chi_p) = \sum_{n \bmod p} \chi_p(n)\chi_p(1-n).$$

Let $\wp(z)$ be a Weierstrass $\wp$-function for the cubic lattice: that is, $\wp'(z)^2 = 4\wp(z)^3 - 1$ and

$$\wp(z) = \frac{1}{z^2} + \sum_{\substack{k \in \theta\mathbb{Z}[e^{2\pi i/3}] \\ k \neq 0}} \left(\frac{1}{(z-k)^2} - \frac{1}{k^2} \right),$$

where θ is the least positive period. Choose ω to be a primitive cube root of 1 modulo p, and let R be a set of $(p-1)/3$ residues modulo p such that $R \cup \omega R \cup \omega^2 R$ is a complete system of residues modulo p, satisfying the additional requirement that $\prod_{r \in R} r = -1 \bmod p$. Then the product

$$H(\chi_p) = -\prod_{r \in R} \wp\left(r\theta / J(\chi_p, \chi_p)\right),$$

which is independent of R, satisfies

$$G(\chi_p) = p^{1/3} J(\chi_p, \chi_p) H(\chi_p). \tag{4.66}$$

We will now prove the local sextic Landsberg–Schaar relation.

Proposition 4.8.4.2. *Let a and c be odd coprime integers, with all prime factors congruent to 3 or 5 modulo 6, let p be a prime congruent to 1 modulo 6, let b be any integer, let τ be a multiplicative inverse for 2 modulo cp and let χ_p be a primitive Dirichlet character modulo p of exact order 3. Then with the notation of Theorem 4.8.4.1,*

$$\begin{aligned} f_6^{-1}(2ap_c^{-1}, c) \sum_{n=0}^{cp-1} \exp\left(\frac{2\pi i(an^6 + abcn^3)}{cp}\right) &= \epsilon_c \sqrt{cp} \left(\frac{a}{cp}\right) \exp\left(\frac{2\pi i a b^2 c}{p} \tau(\tau-1)\right) \\ + \sqrt{\frac{ic}{ap}} \exp\left(-\frac{\pi i a b^2 c}{2p}\right) p^{1/3} \Bigg[&\left(r\mathrm{Re}\left(H(\chi_p)\right) + 3\sqrt{3}s\mathrm{Im}(H(\chi_p))\right) \Bigg[f_6^{-1}(-cp_{2a}^{-1}, 2a) \sum_{n=0}^{2ap-1} \exp\left(-\frac{\pi i(cn^6 + 2abcn^3)}{2ap}\right) \\ &- \exp\left(\frac{\pi i}{4}\right) \sqrt{2ap}\epsilon_c \left(\frac{ap}{c}\right) \exp\left(\frac{\pi i a b^2 c}{2p}\right) \Bigg] \\ - \sum_{n=0}^{2ap-1} &\left(r\mathrm{Re}(H(\chi_p)\chi_p(n)) + 3\sqrt{3}s\mathrm{Im}(H(\chi_p)\chi_p(n))\right) \exp\left(-\frac{\pi i(cn^2 + 2abcn)}{2ap}\right) \Bigg]. \end{aligned} \tag{4.67}$$

Proof. Proposition 4.8.2.3, with $d' = 3$, yields

$$f_6^{-1}(2ap_c^{-1}, c) \sum_{n=0}^{cp-1} \exp\left(\frac{2\pi i(an^6 + abcn^3)}{cp}\right) = \sum_{n=0}^{cp-1} \left(1 + \chi_p(n) + \overline{\chi_p(n)}\right) \exp\left(\frac{2\pi i(an^2 + abcn)}{cp}\right),$$

and upon using 4.60 on the sums twisted by the two nontrivial characters, we find that

$$\begin{aligned} f_6^{-1}(2ap_c^{-1}, c) \sum_{n=0}^{cp-1} \exp\left(\frac{2\pi i(an^6 + abcn^3)}{cp}\right) &= \epsilon_c \sqrt{cp} \left(\frac{a}{cp}\right) \exp\left(\frac{2\pi i a b^2 c}{p} \tau(\tau-1)\right) \\ + \sqrt{\frac{ic}{ap}} \exp\left(-\frac{\pi i a b^2 c}{2p}\right) &\left(G(\chi_p) \sum_{n=0}^{2ap-1} \overline{\chi_p(n)} \exp\left(-\frac{\pi i(cn^2 + 2abcn)}{2ap}\right) \right. \\ &\left. + G(\overline{\chi_p}) \sum_{n=0}^{2ap-1} \chi_p(n) \exp\left(-\frac{\pi i(cn^2 + 2abcn)}{2ap}\right) \right). \end{aligned} \tag{4.68}$$

Now we consider more closely the last two terms in the expression above.

$$G(\chi_p)\sum_{n=0}^{2ap-1}\overline{\chi_p(n)}\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)+G(\overline{\chi_p})\sum_{n=0}^{2ap-1}\chi_p(n)\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)$$
$$=(G(\chi_p)+G(\overline{\chi_p}))\left(\sum_{n=0}^{2ap-1}\left(\chi_p(n)+\overline{\chi_p(n)}\right)\exp\left(-\frac{\pi i(cn^2+2abcn}{2ap}\right)\right) \tag{4.69}$$
$$-\left(G(\chi_p)\sum_{n=0}^{2ap-1}\chi_p(n)\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)+G(\overline{\chi_p})\sum_{n=0}^{2ap-1}\overline{\chi_p(n)}\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)\right). \tag{4.70}$$

The expression 4.69 will end up contributing to the sextic sum on the right hand side of 4.67, so we leave it for now. We will pass off the expression at 4.70 as an "error term", but we can simplify it further by reshaping the Gauss sums with the use of 4.66:

$$G(\chi_p)\sum_{n=0}^{2ap-1}\chi_p(n)\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)+G(\overline{\chi_p})\sum_{n=0}^{2ap-1}\overline{\chi_p(n)}\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)$$
$$=\sum_{n=0}^{2ap-1}\left(G(\chi_p)\chi_p(n)+\overline{G(\chi_p)}\overline{\chi_p(n)}\right)\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)$$
$$=\sum_{n=0}^{2ap-1}\left(\frac{p^{1/3}}{2}\left(\left(r+3\sqrt{3}si\right)H(\chi_p)\chi_p(n)+\left(r-3\sqrt{3}si\right)\overline{H(\chi_p)\chi_p(n)}\right)\right)\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)$$
$$=p^{1/3}\sum_{n=0}^{2ap-1}\left(r\mathrm{Re}(H(\chi_p)\chi_p(n))+3\sqrt{3}s\mathrm{Im}(H(\chi_p)\chi_p(n))\right)\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right). \tag{4.71}$$

The last piece of the puzzle is the link between 4.69 and the sextic sum.

$$\sum_{n=0}^{2ap-1}\left(\chi_p(n)+\overline{\chi_p(n)}\right)\exp\left(-\frac{\pi i(cn^2+2abcn}{2ap}\right)$$
$$=f_6^{-1}(-cp_{2a}^{-1},2a)\sum_{n=0}^{2ap-1}\exp\left(-\frac{\pi i(cn^6+2abcn^3)}{2ap}\right)-\sum_{n=0}^{2ap-1}\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)$$
$$=f_6^{-1}(-cp_{2a}^{-1},2a)\sum_{n=0}^{2ap-1}\exp\left(-\frac{\pi i(cn^6+2abcn^3)}{2ap}\right)-\exp\left(\frac{\pi i}{4}\right)\sqrt{2ap}\epsilon_c\left(\frac{ap}{c}\right)\exp\left(\frac{\pi iab^2c}{2p}\right). \tag{4.72}$$

Substituting 4.72 and 4.71 into 4.69 and 4.70, and thence into 4.68, we arrive at 4.67. □

4.8.5 A generalised octic Landsberg–Schaar relation

As we promised the reader, we will now investigate a local octic analogue of the Landsberg–Schaar relation. The underlying strategy is the same as in Proposition 4.8.4.2 in the last subsection, but in this case the sums which appear as error terms can be "explicitly" evaluated using Theorem 4.8.0.1. We also present a slightly simpler octic version without the quartic term. In this case, a certain quartic sum may also be evaluated explicitly with Theorem 4.8.0.1.

Proposition 4.8.5.1. *Let a and c be odd coprime integers, with all prime factors congruent to 3 modulo 4, let p be a prime congruent to 1 modulo 8, let b be any integer, let τ be a multiplicative inverse for 2 modulo cp and let*

χ_p be a primitive Dirichlet character modulo p of exact order 4. With the notation of Theorem 4.8.0.1,

$$f_8^{-1}(2ap_c^{-1},c)\sum_{n=0}^{cp-1}\exp\left(\frac{2\pi i(an^8+abcn^4)}{cp}\right)=\epsilon_c\sqrt{cp}\left(\frac{a}{c}\right)\left(\left(\frac{a}{p}\right)\exp\left(\frac{2\pi iab^2c}{p}\tau(\tau-1)\right)-\left(\frac{p}{c}\right)\right)$$
$$+\sqrt{\frac{ic}{ap}}\exp\left(-\frac{\pi iab^2c}{2p}\right)\left[\sqrt{p}f_4^{-1}(-cp_{2a}^{-1},2a)\sum_{n=0}^{2ap-1}\exp\left(-\frac{\pi i(cn^4+2abcn^2)}{2ap}\right)\right.$$
$$+C_p\left(\frac{|\beta_p|}{|\alpha_p|}\right)(-1)^{\frac{\beta_p{}^2+2|\beta_p|}{8}}\left[\sqrt{2p+2\alpha_p\sqrt{p}}\left(f_8^{-1}(-cp_{2a}^{-1},2a)\sum_{n=0}^{2ap-1}\exp\left(-\frac{\pi i(cn^8+2abcn^4)}{2ap}\right)\right.\right.$$
$$\left.-f_4^{-1}(-cp_{2a}^{-1},2a)\sum_{n=0}^{2ap-1}\exp\left(-\frac{\pi i(cn^4+2abcn^2)}{2ap}\right)\right)$$
$$\left.\left.-\sum_{n=0}^{2ap-1}\mathrm{Re}(\chi_p(n)\,(1+\mathrm{sgn}\,(\beta_p)i))\sqrt{2p+2\left(\frac{n}{p}\right)\alpha_p\sqrt{p}}\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)\right]\right]. \tag{4.73}$$

Proof. Proposition 4.8.2.3, with $d'=4$, yields

$$f_8^{-1}(2ap_c^{-1},c)\sum_{n=0}^{cp-1}\exp\left(\frac{2\pi i(an^8+abcn^4)}{cp}\right)=\sum_{n=0}^{cp-1}\left(1+\chi_p(n)+\left(\frac{n}{p}\right)+\overline{\chi_p(n)}\right)\exp\left(\frac{2\pi i(an^2+abcn)}{cp}\right),$$

and upon using 4.60 on the sums twisted by the three nontrivial characters (remembering that since $p=1$ mod 8, $\chi_p(-1)=1$), we find that

$$f_8^{-1}(2ap_c^{-1},c)\sum_{n=0}^{cp-1}\exp\left(\frac{2\pi i(an^8+abcn^4)}{cp}\right)=\epsilon_c\sqrt{cp}\left(\frac{a}{cp}\right)\exp\left(\frac{2\pi iab^2c}{p}\tau(\tau-1)\right)$$
$$+\sqrt{\frac{ic}{a}}\exp\left(-\frac{\pi iab^2c}{2p}\right)\left(\sum_{n=0}^{2ap-1}\left(\frac{n}{p}\right)\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)+\frac{G(\chi_p)}{\sqrt{p}}\sum_{n=0}^{2ap-1}\overline{\chi_p(n)}\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)\right.$$
$$\left.+\frac{G(\overline{\chi_p})}{\sqrt{p}}\sum_{n=0}^{2ap-1}\chi_p(n)\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)\right). \tag{4.74}$$

Now we consider more closely the last two terms in the expression above.

$$G(\chi_p)\sum_{n=0}^{2ap-1}\overline{\chi_p(n)}\exp\left(-\frac{\pi i(cn^2+2abcn}{2ap}\right)+G(\overline{\chi_p})\sum_{n=0}^{2ap-1}\chi_p(n)\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)$$
$$=(G(\chi_p)+G(\overline{\chi_p}))\left(\sum_{n=0}^{2ap-1}\left(\chi_p(n)+\overline{\chi_p(n)}\right)\exp\left(-\frac{\pi i(cn^2+2abcn}{2ap}\right)\right) \tag{4.75}$$
$$-\left(G(\chi_p)\sum_{n=0}^{2ap-1}\chi_p(n)\exp\left(-\frac{\pi i(cn^2+2abcn}{2ap}\right)+G(\overline{\chi_p})\sum_{n=0}^{2ap-1}\overline{\chi_p(n)}\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)\right). \tag{4.76}$$

The terms at 4.75 will end up contributing to the octic sum on the right hand side of 4.73, so we leave it for now. We will pass off the terms at 4.76 as "error terms", but we can simplify it further by reshaping the Gauss sums:

$$G(\chi_p)\sum_{n=0}^{2ap-1}\chi_p(n)\exp\left(-\frac{\pi i(cn^2+2abcn}{2ap}\right)=\sum_{k=0}^{p-1}\sum_{\substack{n=1\\(n,p)=1}}^{2ap-1}\chi_p(n)\chi_p(k)\exp\left(\frac{2\pi ik}{p}\right)\exp\left(-\frac{\pi i(cn^2+2abcn}{2ap}\right)$$
$$=\sum_{\substack{n=1\\(n,p)=1}}^{2ap-1}\sum_{k=0}^{p-1}\chi_p(k)\exp\left(\frac{2\pi in_p^{-1}k}{p}\right)\exp\left(-\frac{\pi i(cn^2+2abcn}{2ap}\right),$$

so upon using 4.55,

$$G(\chi_p)\sum_{n=0}^{2ap-1}\chi_p(n)\exp\left(-\frac{\pi i(cn^2+2abcn}{2ap}\right)+G(\overline{\chi_p})\sum_{n=0}^{2ap-1}\overline{\chi_p(n)}\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)$$
$$=\sum_{\substack{n=1\\(n,p)=1}}^{2ap-1}\sum_{k=0}^{p-1}\left(\chi_p(k)+\overline{\chi_p(k)}\right)\exp\left(\frac{2\pi i n_p^{-1}k}{p}\right)\exp\left(-\frac{\pi i(cn^2+2abcn}{2ap}\right)$$
$$=\sum_{\substack{n=1\\(n,p)=1}}^{2ap-1}\phi_p(n_p^{-1},1)\exp\left(-\frac{\pi i(cn^2+2abcn}{2ap}\right). \tag{4.77}$$

Recall that since $p=1 \bmod 8$, $\left(\frac{2}{p}\right)=1$ by the second supplementary law of quadratic reciprocity (4.12). The reader will not be surprised to find that we apply Proposition 4.8.1.2 to explicate 4.77:

$$\sum_{\substack{n=1\\(n,p)=1}}^{2ap-1}\phi_p(n_p^{-1},1)\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)$$
$$=C_p\left(\frac{|\beta_p|}{|\alpha_p|}\right)(-1)^{\frac{\beta_p{}^2+2|\beta_p|}{8}}\sum_{\substack{n=1\\(n,p)=1}}^{2ap-1}\mathrm{Re}(\chi_p(n)\,(1+\mathrm{sgn}\,(\beta_p)i))\sqrt{2p+2\left(\frac{n}{p}\right)\alpha_p\sqrt{p}}\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)$$
$$=C_p\left(\frac{|\beta_p|}{|\alpha_p|}\right)(-1)^{\frac{\beta_p{}^2+2|\beta_p|}{8}}\sum_{n=0}^{2ap-1}\mathrm{Re}(\chi_p(n)\,(1+\mathrm{sgn}\,(\beta_p)i))\sqrt{2p+2\left(\frac{n}{p}\right)\alpha_p\sqrt{p}}\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right).$$

We now substitute 4.8.5 into 4.76, and thence into 4.74. We rearrange the resulting expression in order to create the character sums necessary to give rise to an octic sum: the reader will note that the error terms generated during this process can be evaluated explicitly with Proposition 4.8.1.2.

$$f_8^{-1}(2ap_c^{-1},c)\sum_{n=0}^{cp-1}\exp\left(\frac{2\pi i(an^8+abcn^4)}{cp}\right)=\epsilon_c\sqrt{cp}\left(\frac{a}{cp}\right)\exp\left(\frac{2\pi iab^2c}{p}\tau(\tau-1)\right)$$
$$+\sqrt{\frac{ic}{ap}}\exp\left(-\frac{2\pi iab^2c}{p}\right)\left[\sqrt{p}\sum_{n=0}^{2ap-1}\left(\frac{n}{p}\right)\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)\right.$$
$$+\left(G(\chi_p)+G(\overline{\chi_p})\right)\left(\sum_{n=0}^{2ap-1}\left(\chi_p(n)+\overline{\chi_p(n)}\right)\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)\right)$$
$$\left.-C_p\left(\frac{|\beta_p|}{|\alpha_p|}\right)(-1)^{\frac{\beta_p{}^2+2|\beta_p|}{8}}\sum_{n=0}^{2ap-1}\mathrm{Re}(\chi_p(n)\,(1+\mathrm{sgn}\,(\beta_p)i))\sqrt{2p+2\left(\frac{n}{p}\right)\alpha_p\sqrt{p}}\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)\right] \tag{4.78}$$
$$=\epsilon_c\sqrt{cp}\left(\frac{a}{cp}\right)\exp\left(\frac{2\pi iab^2c}{p}\tau(\tau-1)\right)$$
$$+\sqrt{\frac{ic}{ap}}\exp\left(-\frac{2\pi iab^2c}{p}\right)\left[(\sqrt{p}-(G(\chi_p)+G(\overline{\chi_p})))\sum_{n=0}^{2ap-1}\left(\frac{n}{p}\right)\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)\right.$$
$$+\left(G(\chi_p)+G(\overline{\chi_p})\right)\left(\sum_{n=0}^{2ap-1}\left(\chi_p(n)+\left(\frac{n}{p}\right)+\overline{\chi_p(n)}\right)\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)\right)$$
$$\left.-C_p\left(\frac{|\beta_p|}{|\alpha_p|}\right)(-1)^{\frac{\beta_p{}^2+2|\beta_p|}{8}}\sum_{n=0}^{2ap-1}\mathrm{Re}(\chi_p(n)\,(1+\mathrm{sgn}\,(\beta_p)i))\sqrt{2p+2\left(\frac{n}{p}\right)\alpha_p\sqrt{p}}\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)\right].$$

$$= \epsilon_c\sqrt{cp}\left(\frac{a}{cp}\right)\exp\left(\frac{2\pi iab^2c}{p}\tau(\tau-1)\right)$$

$$+\sqrt{\frac{ic}{ap}}\exp\left(-\frac{2\pi iab^2c}{p}\right)\left[\left(\sqrt{p}-C_p\left(\frac{|\beta_p|}{|\alpha_p|}\right)(-1)^{\frac{\beta_p{}^2+2|\beta_p|}{8}}\sqrt{2p+2\alpha_p\sqrt{p}}\right)\sum_{n=0}^{2ap-1}\left(\frac{n}{p}\right)\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)\right.$$
$$+C_p\left(\frac{|\beta_p|}{|\alpha_p|}\right)(-1)^{\frac{\beta_p{}^2+2|\beta_p|}{8}}\left[\sqrt{2p+2\alpha_p\sqrt{p}}\left(\sum_{n=0}^{2ap-1}\left(\chi_p(n)+\left(\frac{n}{p}\right)+\overline{\chi_p(n)}\right)\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)\right)\right.$$
$$\left.\left.-\sum_{n=0}^{2ap-1}\mathrm{Re}(\chi_p(n)\,(1+\mathrm{sgn}\,(\beta_p)i))\sqrt{2p+2\left(\frac{n}{p}\right)\alpha_p\sqrt{p}}\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right)\right]\right]. \tag{4.79}$$

We have almost arrived at the reciprocal octic sum in the line above 4.79: note that

$$f_8^{-1}(-cp_{2a}^{-1},2a)\sum_{n=0}^{2ap-1}\exp\left(-\frac{\pi i(cn^8+2abcn^4)}{2ap}\right)=\sum_{n=0}^{2ap-1}\left(1+\chi_p(n)+\left(\frac{n}{p}\right)+\overline{\chi_p(n)}\right)\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right).$$

After evaluating the quadratic Gauss sum by completing the square, using the Landsberg–Schaar relation, then evaluating the resulting Gauss sum by Proposition 4.3.2.1,

$$f_8^{-1}(-cp_{2a}^{-1},2a)\sum_{n=0}^{2ap-1}\exp\left(-\frac{\pi i(cn^8+2abcn^4)}{2ap}\right)-\exp\left(\frac{\pi i}{4}\right)\sqrt{2ap}\epsilon_c\left(\frac{ap}{c}\right)\exp\left(\frac{\pi iab^2c}{2p}\right)$$
$$=\sum_{n=0}^{2ap-1}\left(\chi_p(n)+\left(\frac{n}{p}\right)+\overline{\chi_p(n)}\right)\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right). \tag{4.80}$$

Our last manoeuvre in the quest to prevent Dirichlet characters from appearing under summation symbols in our octic relation is achieved using 4.64:

$$f_4^{-1}(-cp_{2a}^{-1},2a)\sum_{n=0}^{2ap-1}\exp\left(-\frac{\pi i(cn^4+2abcn^2)}{2ap}\right)-\exp\left(\frac{\pi i}{4}\right)\sqrt{2ap}\epsilon_c\left(\frac{ap}{c}\right)\exp\left(\frac{\pi iab^2c}{2p}\right)$$
$$=\sum_{n=0}^{2ap-1}\left(\frac{n}{p}\right)\exp\left(-\frac{\pi i(cn^2+2abcn)}{2ap}\right). \tag{4.81}$$

The claim follows after substituting 4.80 and 4.81 into 4.79, then doing a bit of rearranging with the help of quadratic reciprocity. □

In the case that $b=0$, the quartic Gauss sum at 4.81 may be explicitly evaluated, and the statement of Proposition 4.8.5.1 may be simplified.

Corollary 4.8.5.2. *Let a and b be odd coprime integers, with all prime factors congruent to 3 modulo 4, let p be a prime congruent to 1 modulo 8 and let χ_p be a primitive Dirichlet character modulo p of exact order 4. With the notation of Theorem 4.8.0.1,*

$$f_8^{-1}(2bp_a^{-1},a)\sum_{n=0}^{ap-1}\exp\left(\frac{2\pi ibn^8}{ap}\right)=\epsilon_a\sqrt{ap}\left(\frac{b}{ap}\right)+\sqrt{\frac{ia}{bp}}C_p\left(\frac{|\beta_p|}{|\alpha_p|}\right)(-1)^{\frac{\beta_p{}^2+2|\beta_p|}{8}}\times$$
$$\left[\sqrt{bi}\epsilon_c\left(\frac{b}{a}\right)\left(\sqrt{p}-C_p\left(\frac{|\beta_p|}{|\alpha_p|}\right)(-1)^{\frac{\beta_p{}^2+2|\beta_p|}{8}}\sqrt{2p+2\alpha_p\sqrt{p}}\right)\mathrm{Re}\left(\overline{\chi_p(ab)}\,(1+\mathrm{sgn}\,(\beta_p)i)\right)\sqrt{2p+2\left(\frac{ab}{p}\right)\alpha_p\sqrt{p}}\right.$$
$$+\left[\sqrt{2p+2\alpha_p\sqrt{p}}\left(f_8^{-1}(-ap_{2b}^{-1},2b)\sum_{n=0}^{2bp-1}\exp\left(-\frac{\pi ian^8}{2bp}\right)-\exp\left(\frac{\pi i}{4}\right)\sqrt{2bp}\epsilon_a\left(\frac{bp}{a}\right)\right)\right.$$
$$\left.\left.-\sum_{n=0}^{2bp-1}\mathrm{Re}(\chi_p(n)\,(1+\mathrm{sgn}\,(\beta_p)i))\sqrt{2p+2\left(\frac{n}{p}\right)\alpha_p\sqrt{p}}\exp\left(-\frac{\pi ian^2}{2bp}\right)\right]\right]. \tag{4.82}$$

Proof. Setting $b = 0$, then replacing a with b and c with a at 4.78, we have

$$f_8^{-1}(2bp_a^{-1}, a) \sum_{n=0}^{ap-1} \exp\left(\frac{2\pi i b n^8}{ap}\right) = \epsilon_c \sqrt{ap} \left(\frac{b}{ap}\right)$$
$$+ \sqrt{\frac{ia}{bp}} \left[\left(\sqrt{p} - C_p\left(\frac{|\beta_p|}{|\alpha_p|}\right)(-1)^{\frac{\beta_p{}^2+2|\beta_p|}{8}} \sqrt{2p+2\alpha_p\sqrt{p}}\right) \sum_{n=0}^{2bp-1} \left(\frac{n}{p}\right) \exp\left(-\frac{\pi i a n^2}{2bp}\right) \right.$$
$$+ C_p\left(\frac{|\beta_p|}{|\alpha_p|}\right)(-1)^{\frac{\beta_p{}^2+2|\beta_p|}{8}} \left[\sqrt{2p+2\alpha_p\sqrt{p}} \left(f_8^{-1}(-ap_{2b}^{-1}, 2a) \sum_{n=0}^{2bp-1} \exp\left(-\frac{\pi i a n^8}{2bp}\right) - \exp\left(\frac{\pi i}{4}\right)\sqrt{2bp}\epsilon_a\left(\frac{bp}{a}\right)\right) \right.$$
$$\left.\left. - \sum_{n=0}^{2bp-1} \mathrm{Re}(\chi_p(n)\,(1+\mathrm{sgn}\,(\beta_p)i)) \sqrt{2p+2\left(\frac{n}{p}\right)\alpha_p\sqrt{p}} \exp\left(-\frac{\pi i a n^2}{2bp}\right)\right]\right]. \tag{4.83}$$

By 4.50,

$$\sum_{n=0}^{2bp-1} \left(\frac{n}{p}\right) \exp\left(-\frac{\pi i a n^2}{2bp}\right) = \sqrt{\frac{bi}{a}} \sum_{n=0}^{ap-1} \left(\frac{n}{p}\right) \exp\left(\frac{2\pi i n^2 b}{ap}\right),$$

where we have used that $G((\frac{\cdot}{p})) = \sqrt{p}$ since $p = 1 \bmod 8$. For the same reason, $(\frac{a}{p})s'(bp, a) = (\frac{b}{a})\epsilon_a$ by quadratic reciprocity, so by Proposition 4.8.1.2,

$$\sum_{n=0}^{ap-1} \left(\frac{n}{p}\right) \exp\left(\frac{2\pi i n^2 b}{ap}\right) = \phi_p(b, a) = \epsilon_a \left(\frac{b}{a}\right) C_p\left(\frac{|\beta_p|}{|\alpha_p|}\right)(-1)^{\frac{\beta_p{}^2+2|\beta_p|}{8}} \mathrm{Re}\left(\overline{\chi_p(ab)}\,(1+\mathrm{sgn}\,(\beta_p)i)\right) \sqrt{2ap+2a\left(\frac{ab}{p}\right)\alpha_p\sqrt{p}}.$$

Substituting 4.8.5 into 4.83, we may rearrange to form 4.82. □

Remark 4.8.5.3. *If we take a and c to be squarefree in Propositions 4.8.3.1, 4.8.4.2 and 4.8.5.1, or a and b to be squarefree in Corollary 4.8.5.2, then all terms of the form $f_4^{-1}(\,\cdot\,,\,\cdot\,)$, $f_6^{-1}(\,\cdot\,,\,\cdot\,)$ and $f_8^{-1}(\,\cdot\,,\,\cdot\,)$ are equal to one.*

We do not attempt to formulate local Landsberg–Schaar relations for any higher degrees, owing to the large number of "error terms" which appear, few of which can be satisfactorily simplified due to the non-existence of analogues of Theorem 4.8.0.1 and Theorem 4.8.4.1 for Gauss sums of higher order. It appears that little progress has been made beyond the results outlined by Berndt and Evans [BE81].

In the proof of Proposition 4.8.5.1, we used Matthews' evaluation of the quartic Gauss sums (Theorem 4.8.0.1). Although Matthews' proof is mostly algebraic, there is a critical section in which analytic results relating to modular forms are employed. Until such time as a way is found to circumvent these "analytic" portions of Matthews' proof, Proposition 4.8.5.1 will lack an elementary proof.

Chapter 5

Gauss sums over algebraic number fields

In Chapter 4 we explored in detail the relationship between Jacobi's theta functions, reciprocity between Gauss sums and the law of quadratic reciprocity. This expedition took place over the rational numbers. A natural impulse is to repeat the process over algebraic number fields. This was done around 1920 by Hecke [Hec23], and as a consequence the analogues of Gauss sums over number fields are called Hecke sums[9]. We investigate asymptotic expansions of rather simple theta functions in order to state a "generalised" version of Hecke reciprocity, in analogy with 4.9. Unfortunately, it appears that the correct theta functions to consider are more complicated than those that we treat here; therefore, we stop short of proving a local quartic version of Hecke reciprocity, as the sums we obtain do not interact well without severe restrictions on the primes involved. We do demonstrate that, at the level of Hecke sums, all the necessary identities from Chapter 4, such as reduction of order (Proposition 4.7.1.1) and the higher-degree product rule (Lemma 4.7.0.1), carry over exactly as expected.

Another reason that our results here are not so complete as our results for the rational analogues of Chapter 4 is due to the fact that the Hecke sums have not yet been evaluated in full generality. Boylan and Skoruppa have indicated [BS13, pp. 111] that this can be done using Wall's theorem [Wal64]: any non-degenerate finite quadratic module is isomorphic to the discriminant module of an even integral lattice. Indeed, the map $(\mathcal{O}_K/\mathfrak{a}, \mu+\mathfrak{a}) \mapsto \mathrm{Tr}(\omega\mu^2)+\mathbb{Z}$ is a non-degenerate finite quadratic module, so by Wall's theorem each Hecke sum may be written as the *Gauss invariant* of an even integral lattice, which may then be evaluated as a known multiple of an eighth root of unity using Milgram's formula [HM73, Appendix 4] (with the proof in the 1-dimensional case replaced with an elementary proof such as Gauss' [Gau11]). Of course, in order to carry out this process one must determine exactly which even integral lattice Wall's theorem associates to each Hecke sum.

We also note that an analogue of Matthews' evaluation of the quartic (or cubic) rational Gauss sum is as yet unavailable [Patterson, personal communication], so we are unable to evaluate explicitly twisted Hecke sums of the form

$$\sum_{\mu \in \mathcal{O}_K/\mathfrak{m}\cdot \mathrm{denom}(\omega)} \chi(\mu) e^{2\pi i \mathrm{Tr}(\mu^2\omega)}$$

for any nontrivial Dirichlet characters modulo $\mathfrak{m}$, which precludes the possibility of writing down interesting local sextic or octic versions of Hecke reciprocity.

We intend, in future work, to attempt to rectify these omissions as far as possible. There is very little doubt that a local quartic version of Hecke reciprocity may be obtained by employing Hecke theta functions, and one expects that a fairly explicit evaluation of the quadratic Hecke sums may be easily obtained.

5.1 Theta functions associated to number fields

In this section we use theta functions associated to *totally real* number fields to deduce Hecke reciprocity. A similar process achieves exactly the same result over number fields with a complex embedding, but the proof is longer. We state the "product rule" and the formula connecting residue symbols to Hecke sums, which, as in Subsection 4.1.2, may be combined with Hecke reciprocity to prove the main law of quadratic reciprocity over all number fields.

[9] One usually does see the phrase "Gauss sum" used in the literature for this expression as well, but we eschew this practice in order to avoid putting extra stress on an already overused term.

We use slightly a slightly different process to that employed by Hecke in his original derivation of the asymptotic expansions in order to better harmonise the discussion with our treatment of the rational case.

5.1.1 Theta functions over totally real number fields

Let K be a totally real number field of degree n, with real embeddings $\sigma_1, \dots, \sigma_n$. For $\omega \in K$, we write $\omega^{(j)}$ for $\sigma_j(n)$. Let $\mathfrak{a}$ stand for a nonzero ideal of K, and choose a basis $\{\alpha_1, \dots, \alpha_n\}$ for $\mathfrak{a}$ over $\mathbb{Z}$. By Proposition 1.3.1.15, $\{\beta_1, \dots, \beta_n\}$ is a $\mathbb{Z}$-basis for $\mathfrak{a}^{-1}\mathfrak{d}^{-1}$, where the β_j are defined implicitly by

$$\sum_{j=1}^{n} \beta_s^{(j)} \alpha_t^{(j)} = \begin{cases} 1 & s = t, \\ 0 & s \neq t. \end{cases}$$

We define our theta function:

$$\Theta_K(\mathfrak{a}, t) = \sum_{\mu \in \mathfrak{a}} \exp\left(2\pi i \sum_{j=1}^{n} t_j \left(\mu^{(j)}\right)^2\right),$$

where $t = (t_1, \dots, t_n) \in \mathscr{H}^n$, and in subsequent subsections we use the abbreviation

$$\Theta_K(t) = \Theta_K(\mathcal{O}_K, t) = \sum_{\mu \in \mathfrak{a}} \exp\left(2\pi i \sum_{j=1}^{n} t_j \left(\mu^{(j)}\right)^2\right).$$

First we establish a functional equation for $\Theta_K(\mathfrak{a}, t)$. We set

$$f_t(x) = \exp\left(2\pi i \sum_{j=1}^{n} t_j \left(\sum_{s=1}^{n} \alpha_s^{(j)} x_s\right)^2\right)$$

for $x = (x_s) \in \mathbb{R}^n$, and calculate the Fourier transform

$$\hat{f}_t(m) = \int_{\mathbb{R}^n} \exp\left(2\pi i \left[\sum_{j=1}^{n} t_j \left(\sum_{s=1}^{n} \alpha_s^{(j)} x_s\right)^2 - \sum_{s=1}^{n} m_s x_s\right]\right) dx.$$

We make the change of variables

$$y_j = \sum_{s=1}^{n} \alpha_s^{(j)} x_s, \quad x_s = \sum_{t=1}^{n} \beta_s^{(t)} y_t$$

and note that by Proposition 1.3.1.15, the Jacobian of this coordinate transformation is $(\mathfrak{N}(\mathfrak{a})\sqrt{|d_K|})^{-1}$, where d_K is the discriminant of K. It follows that

$$\begin{aligned}\hat{f}_t(m) &= \frac{1}{\mathfrak{N}(\mathfrak{a})\sqrt{|d_K|}} \int_{\mathbb{R}^n} \exp\left(2\pi i \left[\sum_{j=1}^{n} t_j y_j^2 - \sum_{j=1}^{n} \left(\sum_{s=1}^{n} \beta_s^{(j)} m_s\right) y_j\right]\right) dx \\ &= \frac{1}{\mathfrak{N}(\mathfrak{a})\sqrt{|d_K|}} \prod_{j=1}^{n} \int_{\mathbb{R}} \exp\left(2\pi i \left[t_j y_j^2 - \left(\sum_{s=1}^{n} \beta_s^{(j)} m_s\right) y_j\right]\right) dy_j,\end{aligned}$$

and upon computing the inner integral, we have

$$\hat{f}_t(m) = \frac{1}{\mathfrak{N}(\mathfrak{a})\sqrt{|d_K|}} \prod_{j=1}^{n} \frac{1}{\sqrt{-2it_j}} \exp\left(-2\pi i \left(\sum_{s=1}^{n} \beta_s^{(j)} m_s\right)^2 \Big/ (4t_j)\right).$$

Clearly $\hat{f}_t(m)$ is Schwartz in m, so we may apply Poisson summation to conclude that

$$\Theta_K(\mathfrak{a}, t) = \frac{1}{\mathfrak{N}(\mathfrak{a})\sqrt{|d_K|}} \prod_{j=1}^{n} \frac{1}{\sqrt{-2it_j}} \Theta_K(\mathfrak{a}^{-1}\mathfrak{d}^{-1}, -1/4t), \tag{5.1}$$

where $-1/4t = (-1/4t_1, \dots, -1/4t_n)$.

5.1.2 Asymptotic expansions of theta functions

Let ω stand for a nonzero element of K and write

$$(\omega) = \mathfrak{b}\mathfrak{a}^{-1}\mathfrak{d}^{-1},$$

where $\mathfrak{a}$ and $\mathfrak{b}$ are integral ideals, uniquely determined by the condition $(\mathfrak{a}, \mathfrak{b}) = 1$. We will set $(t_j) = (\omega^j + i\epsilon_j)$, where $\mathrm{Re}(\epsilon_j) > 0$ for all j. Then

$$\begin{aligned}\Theta_K(t_j) &= \sum_{\mu \in \mathcal{O}_K/\mathfrak{a}} e^{2\pi i \mathrm{Tr}(\omega\mu^2)} \sum_{\nu \in \mathfrak{a}} \exp\left(-2\pi \sum_{j=1}^n \epsilon_j \left(\mu^{(j)} + \nu^{(j)}\right)^2\right) \\ &= \sum_{\mu \in \mathcal{O}_K/\mathfrak{a}} e^{2\pi i \mathrm{Tr}(\omega\mu^2)} \sum_{m=(m_j)\in\mathbb{Z}^n} \exp\left(-2\pi \sum_{j=1}^n \epsilon_j \left(\mu^{(j)} + \sum_{s=1}^n \alpha_s^{(j)} m_j\right)^2\right),\end{aligned}$$

where, as in the previous subsection, $(\alpha_1, \ldots, \alpha_n)$ is a $\mathbb{Z}$-basis for $\mathfrak{a}$. By Euler–Maclaurin summation,

$$\Theta_K(t_j) \sim \sum_{\mu \in \mathcal{O}_K/\mathfrak{a}} e^{2\pi i \mathrm{Tr}(\omega\mu^2)} \int_{x=(x_j)\in\mathbb{R}^n} \exp\left(-2\pi \sum_{j=1}^n \epsilon_j \left(\mu^{(j)} + \sum_{s=1}^n \alpha_s^{(j)} x_j\right)^2\right) dx + H.O.T(\epsilon),$$

and upon making the change of variables

$$y_j = \sum_{s=1}^n \alpha_s^{(j)} x_j,$$

and using Proposition 1.3.1.15, we find that

$$\begin{aligned}\int_{x=(x_j)\in\mathbb{R}^n} \exp\left(-2\pi \sum_{j=1}^n \epsilon_j \left(\mu^{(j)} + \sum_{s=1}^n \alpha_s^{(j)} x_j\right)^2\right) dx &= \frac{1}{\mathfrak{N}(\mathfrak{a})\sqrt{|d_k|}} \int_{\mathbb{R}^n} \exp\left(-2\pi \sum_{j=1}^n \epsilon_j \left(\mu_j + y_j\right)^2\right) dy \\ &= \frac{1}{\mathfrak{N}(\mathfrak{a})\sqrt{|d_k|}} \prod_{j=1}^n \int_{\mathbb{R}} \exp\left(-2\pi\epsilon_j y_j^2\right) dy_j \\ &= \frac{1}{\mathfrak{N}(\mathfrak{a})\sqrt{\prod_{j=1}^n 2\epsilon_j |d_k|}}.\end{aligned}$$

Putting it all together, we have, as $\epsilon \to 0$,

$$\Theta_K(t) \sim \frac{1}{\mathfrak{N}(\mathfrak{a})\sqrt{\prod_{j=1}^n 2\epsilon_j |d_k|}} \sum_{\mu \in \mathcal{O}_K/\mathfrak{a}} e^{2\pi i \mathrm{Tr}(\omega\mu^2)} + H.O.T(\epsilon). \tag{5.2}$$

Now we set

$$-\frac{1}{4t_j} = -\frac{1}{4\omega^{(j)} + 4i\epsilon_j} = -\frac{1}{4\omega^{(j)}} + i\tau_j,$$

and we observe that $\mathrm{Re}(\epsilon_j) > 0$ if and only if $\mathrm{Re}(\tau_j) > 0$.

Let us now choose an auxiliary ideal $\mathfrak{c}$ so that $\mathfrak{c}\mathfrak{d} = (\delta)$ is principal and $(\mathfrak{c}, 2\mathfrak{b}) = 1$. Then the elements of $\mathfrak{d}^{-1}$ may be written as μ/δ, where μ runs through the elements of $\mathfrak{c}$. We have

$$\Theta_K\left(\mathfrak{d}^{-1}, -\frac{1}{4t}\right) = \sum_{\mu\in\mathfrak{c}} \exp\left(2\pi i \sum_{j=1}^n \left(-\frac{1}{4\omega^{(j)}} + i\tau_j\right)\left(\frac{\mu^{(j)}}{\delta^{(j)}}\right)^2\right).$$

Next, we define $\mathfrak{b}_1$ to be the least common multiple of the denominator of $\mathfrak{a}/4\mathfrak{b}$ and the denominator of $\mathfrak{c}/2\mathfrak{b}$, and it follows that

$$\begin{aligned}\Theta_K\left(\mathfrak{d}^{-1}, -\frac{1}{4t}\right) &= \sum_{\substack{\mu \bmod \mathfrak{b}_1 \\ \mu = 0 \bmod \mathfrak{c}}} e^{2\pi i \mathrm{Tr}(\mu^2/4\omega\delta^2)} \sum_{\nu\in\mathfrak{b}_1\mathfrak{c}} \exp\left(-2\pi \sum_{j=1}^n \frac{\tau_j}{\left(\delta^{(j)}\right)^2}\left(\mu^{(j)} + \nu^{(j)}\right)^2\right) \\ &= \sum_{\substack{\mu \bmod \mathfrak{b}_1 \\ \mu = 0 \bmod \mathfrak{c}}} e^{2\pi i \mathrm{Tr}(\mu^2/4\omega\delta^2)} \sum_{m=(m_j)\in\mathbb{Z}^n} \exp\left(-2\pi \sum_{j=1}^n \frac{\tau_j}{\left(\delta^{(j)}\right)^2}\left(\mu^{(j)} + \sum_{s=1}^n m_s\gamma_s^{(j)}\right)^2\right),\end{aligned}$$

where $(\gamma_1, \ldots, \gamma_n)$ be a $\mathbb{Z}$-basis for $\mathfrak{b}_1\mathfrak{c}$. Now we may use Euler–Maclaurin summation to compute the asymptotic expansion of the inner sum:

$$\sum_{m\in\mathbb{Z}^n} \exp\left(-2\pi \sum_{j=1}^n \frac{\tau_j}{\left(\delta^{(j)}\right)^2}\left(\mu^{(j)} + \sum_{s=1}^n m_s\gamma_s^{(j)}\right)^2\right) \sim \int_{\mathbb{R}^n} \exp\left(-2\pi \sum_{j=1}^n \frac{\tau_j}{\left(\delta^{(j)}\right)^2}\left(\mu^{(j)} + \sum_{s=1}^n x_s\gamma_s^{(j)}\right)^2\right) dx + H.O.T(\tau).$$

Upon making the change of variables $y_j = \sum_{s=1}^n x_s\gamma_s^{(j)}$, and using Proposition 1.3.1.15,

$$\begin{aligned}\int_{x=(x_s)\in\mathbb{R}^n} \exp\left(-2\pi \sum_{j=1}^n \frac{\tau_j}{\left(\delta^{(j)}\right)^2}\left(\mu^{(j)} + \sum_{s=1}^n x_s\gamma_s^{(j)}\right)^2\right) dx &= \frac{1}{\mathfrak{N}(\mathfrak{b}_1\mathfrak{c})\sqrt{|d_K|}} \prod_{j=1}^n \delta^{(j)} \int_{\mathbb{R}} \exp\left(-2\pi\tau_j y_j^2\right) dy_j \\ &= \frac{\mathrm{N}(\delta)}{\mathfrak{N}(\mathfrak{b}_1\mathfrak{c})\sqrt{\prod_{j=1}^n 2\tau_j|d_k|}}.\end{aligned}$$

Therefore the full asymptotic expansion is

$$\Theta_K\left(\mathfrak{d}^{-1}, -\frac{1}{4t}\right) \sim \frac{\mathrm{N}(\delta)}{\mathfrak{N}(\mathfrak{b}_1\mathfrak{c})\sqrt{\prod_{j=1}^n 2\tau_j|d_k|}} \sum_{\substack{\mu \bmod \mathfrak{b}_1 \\ \mu=0 \bmod \mathfrak{c}}} e^{2\pi i \mathrm{Tr}(-\mu^2/4\omega\delta^2)} + H.O.T(\tau). \tag{5.3}$$

Comparing 5.3 with 5.2, setting

$$|\tau| = \prod_{j=1}^n \tau^j, \quad |\epsilon| = \prod_{j=1}^n \epsilon^j,$$

and using the functional equation 5.1, we find that

$$\begin{aligned}\frac{1}{\mathfrak{N}(\mathfrak{a})\sqrt{2^n|d_k|}} \sum_{\mu\in\mathcal{O}_K/\mathfrak{a}} e^{2\pi i\mathrm{Tr}(\omega\mu^2)} &= \lim_{\epsilon\to 0} \sqrt{|\epsilon|}\Theta_K(t) \\ &= \lim_{\epsilon\to 0} \sqrt{|\epsilon|}\frac{1}{\sqrt{|d_K|}} \prod_{j=1}^n \frac{1}{\sqrt{-2i(\omega^{(j)} + i\epsilon_j)}} \Theta_K(\mathfrak{d}^{-1}, -1/4t) \\ &= \lim_{\substack{\epsilon\to 0\\ \tau\to 0}} \sqrt{\frac{|\epsilon|}{|\tau|}} \frac{1}{\sqrt{|d_K|}} \prod_{j=1}^n \frac{1}{\sqrt{-2i(\omega^{(j)} + i\epsilon_j)}} \frac{\mathrm{N}(\delta)}{\mathfrak{N}(\mathfrak{b}_1\mathfrak{c})\sqrt{2^n|d_K|}} \sum_{\substack{\mu \bmod \mathfrak{b}_1 \\ \mu=0 \bmod \mathfrak{c}}} e^{-2\pi i\mathrm{Tr}(\mu^2/4\omega\delta^2)}.\end{aligned} \tag{5.4}$$

We note that

$$\lim_{\epsilon\to 0} \prod_{j=1}^n \frac{1}{\sqrt{-2i(\omega^{(j)} + i\epsilon_j)}} = |\sqrt{\mathfrak{N}(2\omega)}| \exp\left(-\frac{\pi i}{4}\mathrm{Tr}\,\mathrm{sgn}(\omega)\right),$$

where we define

$$\mathrm{Tr}\,\mathrm{sgn}(\omega) = \mathrm{sgn}\,\omega^{(1)} + \ldots \mathrm{sgn}\,\omega^{(n)}.$$

Also, we have

$$\lim_{\substack{\epsilon_j\to 0\\ \tau_j\to 0}} \frac{|\epsilon|}{|\tau|} = 4\left(\omega^j\right)^2, \quad \text{so} \quad \lim_{\substack{\epsilon\to 0\\ \tau\to 0}} \sqrt{\frac{|\epsilon|}{|\tau|}} = |\mathfrak{N}(2\omega)|.$$

We may now simplify 5.4 to

$$\frac{1}{\mathfrak{N}(\mathfrak{a})} \sum_{\mu\in\mathcal{O}_K/\mathfrak{a}} e^{2\pi i\mathrm{Tr}(\omega\mu^2)} = \frac{|\mathfrak{N}(2\omega)|}{|\sqrt{\mathfrak{N}(2\omega)}|} \frac{\mathrm{N}(\delta)}{\mathfrak{N}(\mathfrak{b}_1\mathfrak{c})\sqrt{|d_K|}} \exp\left(\frac{\pi i}{4}\mathrm{Tr}\,\mathrm{sgn}(\omega)\right) \sum_{\substack{\mu \bmod \mathfrak{b}_1 \\ \mu=0 \bmod \mathfrak{c}}} e^{2\pi i\mathrm{Tr}(\mu^2/4\omega\delta^2)},$$

and upon using $\mathfrak{N}(\mathfrak{d}) = |d_K|$, this may be rearranged to yield

$$\frac{1}{|\sqrt{\mathfrak{N}(\mathfrak{a})}|} \sum_{\mu\in\mathcal{O}_K/\mathfrak{a}} e^{2\pi i\mathrm{Tr}(\omega\mu^2)} = \left|\frac{\sqrt{\mathfrak{N}(2\mathfrak{b})}}{\mathfrak{N}(\mathfrak{b}_1)}\right| \exp\left(\frac{\pi i}{4}\mathrm{Tr}\,\mathrm{sgn}(\omega)\right) \sum_{\substack{\mu \bmod \mathfrak{b}_1 \\ \mu=0 \bmod \mathfrak{c}}} e^{-2\pi i\mathrm{Tr}(\mu^2/4\omega\delta^2)}.$$

The sum on the right hand side may be simplified as follows. Pick some integer α divisible by $\mathfrak{c}$, such that $((\alpha)\mathfrak{c}^{-1}, \mathfrak{b}_1) = 1$. Then summing over $\mu\alpha$ instead of μ and letting μ run over a system of residues modulo $\mathfrak{b}_1$, we obtain

$$\sum_{\mu \in \mathfrak{b}_1} e^{-2\pi i \mathrm{Tr}(\mu^2\alpha^2/4\omega\delta^2)}.$$

We set $\alpha/\delta = \gamma$ and obtain (a special case of) Hecke reciprocity.

Theorem 5.1.2.1. *Let $\omega \in K$ and write $\mathfrak{d}(\omega) = \mathfrak{b}/\mathfrak{a}$ where $(\mathfrak{a}, \mathfrak{b}) = 1$. Let $\mathfrak{b}_1$ stand for the denominator of $\mathfrak{a}/4\mathfrak{b}^{-1}$ and let γ be any element of K such that $\mathfrak{d}(\gamma)$ is integral and relatively prime to $\mathfrak{b}_1$. Then*

$$\frac{C(\omega)}{|\sqrt{\mathfrak{N}(\mathfrak{a})}|} = \left|\frac{\sqrt{\mathfrak{N}(2\mathfrak{b})}}{\mathfrak{N}(\mathfrak{b}_1)}\right| \exp\left(\frac{\pi i}{4}\left(\mathrm{Tr}\,\mathrm{sgn}(\omega)\right) C\left(-\frac{\gamma^2}{4\omega}\right).\right.$$

Suppose that K is an arbitrary *number field, and extend* Tr sgn *with the definition*

$$\mathrm{Tr}\,\mathrm{sgn}(\omega) = \mathrm{sgn}\,\omega^{(1)} + \ldots \mathrm{sgn}\,\omega^{(r_1)},$$

where $\sigma_1, \ldots \sigma_{r_1}$ are the real embeddings of $K \to \mathbb{C}$. Then Theorem 5.1.2.1 holds over K: this is the general version of Hecke reciprocity [Hec81, Theorem 163].

The reader may consult Hecke for his proof of quadratic reciprocity. We are content to note that the following analogue of 4.13 holds:

Lemma 5.1.2.2. *Suppose that the denominator $\mathfrak{a}$ of $\mathfrak{d}(\omega)$ is an odd ideal. Then for every integer $\mathfrak{x}$ which is relatively prime to $\mathfrak{a}$, we have*

$$C(\mathfrak{x}\omega) = \left(\frac{\mathfrak{x}}{\mathfrak{a}}\right) C(\omega).$$

and there is a "product rule" for quadratic Gauss sums, obtained by setting $d = 2$ in Proposition 5.3.2.1. A faithful mimicry of the process followed in Subsection 4.1.2, but with an abundance of auxiliary ideals, suffices to establish quadratic reciprocity:

Theorem 5.1.2.3. *Let K be an algebraic number field, with $\sigma_1, \ldots, \sigma_{r_1}$ the real embeddings. An integer in K is called* primary *if it is odd and congruent to a square in $\mathcal{O}_K$ modulo (4).*

If α and β are two odd relatively prime integers of K, and at least one is primary, then

$$\left(\frac{\alpha}{\beta}\right)\left(\frac{\beta}{\alpha}\right) = (-1)^{\sum_{p=1}^{r_1}\left(\mathrm{sgn}\,\alpha^{(p)}-1\right)\left(\mathrm{sgn}\,\beta^{(p)}-1\right)/4}.$$

5.2 Generalised Hecke reciprocity

In this section we augment the theta functions of Section 5.1 with an auxiliary complex variable. We prove a functional equation for the resulting theta series, and compute the asymptotic expansion. As a corollary, we deduce a generalised version of Hecke's reciprocity relation (Theorem 5.1.2.1), which we conjecture is valid over *arbitrary* number fields.

The results of this section are of particular use in the next section, in which we prove a twisted version of Hecke's reciprocity relation.

5.2.1 Theta functions

For K a totally real number field, $\mathfrak{a}$ an integral ideal with a $\mathbb{Z}$-basis $(\alpha_1, \ldots, \alpha_n)$. With $t = (t_1, \ldots, t_n) \in \mathscr{H}^n$ and $z = (z_1, \ldots, z_n) \in \mathbb{C}^n$, we define

$$\Theta_K(\mathfrak{a}, t, z) = \sum_{\mu \in \mathfrak{a}} \exp\left(2\pi i \sum_{j=1}^{n}\left[t_j\left(\mu^{(j)}\right)^2 + z_j \mu^{(j)}\right]\right).$$

To obtain the functional equation, we set

$$f_t(x) = \exp\left(2\pi i \sum_{j=1}^{n}\left[t_j\left(\sum_{s=1}^{n}\alpha_s^{(j)}x_s\right)^2 + z_j\sum_{s=1}^{n}\alpha_s^{(j)}x_s\right]\right),$$

and calculate the Fourier transform

$$\hat{f}_t(m) = \int_{\mathbb{R}^n}\exp\left(2\pi i\sum_{j=1}^{n}\left[t_j\left(\sum_{s=1}^{n}\alpha_s^{(j)}x_s\right)^2 + z_j\sum_{s=1}^{n}\alpha_s^{(j)}x_s - \sum_{s=1}^{n}m_s x_s\right]\right)dx.$$

We make the linear change of variables

$$y_j = \sum_{s=1}^{n}\alpha_s^{(j)}x_s, \quad x_s = \sum_{t=1}^{n}\beta_s^{(t)}y_t$$

and as in the preceding version, the Jacobian of this coordinate transformation is $(\mathfrak{N}(\mathfrak{a})\sqrt{|d_K|})^{-1}$. Then our integrand is diagonalised as follows:

$$\begin{aligned}\hat{f}_t(m) &= \frac{1}{\mathfrak{N}(\mathfrak{a})\sqrt{|d_K|}}\int_{\mathbb{R}^n}\exp\left(2\pi i\left[\sum_{j=1}^{n}t_j y_j^2 + \sum_{j=1}^{n}\left(z_j - \sum_{s=1}^{n}\beta_s^{(j)}m_s\right)y_j\right]\right)dx\\ &= \frac{1}{\mathfrak{N}(\mathfrak{a})\sqrt{|d_K|}}\prod_{j=1}^{n}\int_{\mathbb{R}}\exp\left(2\pi i\left[t_j y_j^2 + \left(z_j - \sum_{s=1}^{n}\beta_s^{(j)}m_s\right)y_j\right]\right)dy_j.\end{aligned}$$

We compute the integrals to obtain

$$\hat{f}_t(m) = \frac{1}{\mathfrak{N}(\mathfrak{a})\sqrt{|d_K|}}\frac{1}{\sqrt{(-2i)^n|t|}}\exp\left(-\frac{\pi i}{2}\sum_{j=1}^{n}\frac{z_j^2}{t_j}\right)\exp\left(2\pi i\sum_{j=1}^{n}\left[-\frac{1}{4t_j}\left(\sum_{s=1}^{n}\beta_s^{(j)}m_s\right)^2 + \frac{z_j}{2t_j}\sum_{s=1}^{n}\beta_s^{(j)}m_s\right]\right),$$

where $|t| = t_1\cdot\cdots\cdot t_n$. Clearly $\hat{f}_t(m)$ is Schwartz in m, so we may apply Poisson summation to conclude that

$$\Theta_K(\mathfrak{a},t,z) = \frac{1}{\mathfrak{N}(\mathfrak{a})\sqrt{|d_K|}}\frac{1}{\sqrt{(-2i)^n|t|}}\exp\left(-\frac{\pi i}{2}\sum_{j=1}^{n}\frac{z_j^2}{t_j}\right)\Theta_K(\mathfrak{a}^{-1}\mathfrak{d}^{-1}, -1/4t, z/2t), \tag{5.5}$$

where $-1/4t = (-1/4t_1,\ldots,-1/4t_n)$ and $z/2t = (z_1/2t_1,\ldots,z_n/2t_n)$.

5.2.2 Another asymptotic expansion

The contents of this section are predictable. We follow a similar trajectory to the path traced out in Section 4.5 to arrive at a generalised version of Theorem 5.1.2.1.

Let ω and ρ be nonzero elements of a totally real number field K such that

$$\mathfrak{d}(\omega) = \frac{\mathfrak{b}}{\mathfrak{a}}, \quad \mathfrak{d}(\rho) = \frac{\mathfrak{c}}{\mathfrak{a}},$$

where $\mathfrak{a}, \mathfrak{b}$ and $\mathfrak{c}$ are integral ideals such that $(\mathfrak{a},\mathfrak{b}) = 1 = (\mathfrak{a},\mathfrak{c})$. Set

$$t = (t_j) = (\omega^{(j)} + i\epsilon_j) = \omega + i\epsilon, \quad z = (z_j) = (\rho^{(j)}) = \rho,$$

where $\mathrm{Re}(\epsilon_j) > 0$ for all $j = 1,\ldots,n$. As in Section 5.1, we write $\Theta_K(t,z)$ instead of $\Theta_K(\mathcal{O}_K,t,z)$.

We have

$$\Theta_K(\omega + i\epsilon,\rho) = \sum_{\mu\in\mathcal{O}_K/\mathfrak{a}} e^{2\pi i\mathrm{Tr}(\omega\mu^2+\rho\mu)}\sum_{\nu\in\mathfrak{a}}\exp\left(-2\pi\sum_{j=1}^{n}\epsilon_j\left(\mu^{(j)}+\nu^{(j)}\right)^2\right),$$

and the inner sum may be rewritten as

$$\sum_{\nu\in\mathfrak{a}}\exp\left(-2\pi\sum_{j=1}^{n}\epsilon_j\left(\mu^{(j)}+\nu^{(j)}\right)^2\right)=\sum_{m\in\mathbb{Z}^n}\exp\left(-2\pi\sum_{j=1}^{n}\epsilon_j\left(\mu^{(j)}+\sum_{s=1}^{n}\alpha_s^{(j)}x_s\right)^2\right).$$

By the computation in Section 5.1, the asymptotic expansion is

$$\Theta_K(\omega+i\epsilon,\rho)\sim\frac{1}{\mathfrak{N}(\mathfrak{a})\sqrt{2^n|d_k||\epsilon|}}\sum_{\mu\in\mathcal{O}_K/\mathfrak{a}}e^{2\pi i\mathrm{Tr}(\omega\mu^2+\rho\mu)}. \tag{5.6}$$

The shape of the functional equation dictates that we compute another asymptotic expansion. Indeed, we write

$$-\frac{1}{4(\omega+i\epsilon)}=-\frac{1}{4\omega}+i\tau\quad\text{and so}\quad\frac{\rho}{2(\omega+i\epsilon)}-\frac{\rho}{2\omega}+2\rho i\tau$$

where $\mathrm{Re}(\epsilon)>0$ if and only if $\mathrm{Re}(\tau)>0$. It behooves us to consider $\Theta_K(\mathfrak{d},-1/4\omega+i\tau,-\rho/2\omega+2\rho i\tau)$. We choose an auxiliary ideal $\mathfrak{f}$ such that $\mathfrak{f}\mathfrak{d}=(\delta)$, where $\delta\in K$, and $(\mathfrak{f},2\mathfrak{b})=1$. The elements of $\mathfrak{d}^{-1}$ are then all of the form κ/δ, where κ runs over the elements of $\mathfrak{f}$. We also note that if we set $\mathfrak{b}_1$ to be the denominator of $\mathfrak{a}/4\mathfrak{b}$, then we may write elements κ of $\mathfrak{f}$ as $\kappa=\mu+\nu$, where μ runs through a complete residue system modulo $\mathfrak{b}_1$ such that $\mu=0\bmod\mathfrak{f}$, and ν runs through $\mathfrak{b}_1\mathfrak{f}$. Thus, we expand the theta function:

$$\begin{aligned}&\Theta_K\left(\mathfrak{d}^{-1},-\frac{1}{4\omega}+i\tau,-\frac{\rho}{2\omega}+2\rho i\tau\right)\\&=\sum_{\substack{\mu\bmod\mathfrak{b}_1\\ \mu=0\bmod\mathfrak{f}}}e^{2\pi i\mathrm{Tr}(\mu^2/4\omega\delta^2+2\rho\mu/4\omega\delta)}\sum_{\nu\in\mathfrak{b}_1\mathfrak{f}}\exp\left(-2\pi\sum_{j=1}^{n}\tau_j\left[\left(\frac{\mu^{(j)}+\nu^{(j)}}{\delta^{(j)}}\right)^2+2\rho^{(j)}\left(\frac{\mu^{(j)}+\nu^{(j)}}{\delta^{(j)}}\right)\right]\right),\end{aligned}$$

and the inner sum may be rewritten as

$$\sum_{m\in\mathbb{Z}^n}\exp\left(-2\pi\sum_{j=1}^{n}\tau_j\left[\left(\frac{\mu^{(j)}+\sum_{s=1}\gamma_s^{(j)}m_s}{\delta^{(j)}}\right)^2+2\rho^{(j)}\left(\frac{\mu^{(j)}+\sum_{s=1}\gamma_s^{(j)}m_s}{\delta^{(j)}}\right)\right]\right),$$

where $(\gamma_1,\dots,\gamma_n)$ is a $\mathbb{Z}$-basis for $\mathfrak{b}_1\mathfrak{f}$. By Euler–Maclaurin summation, this last sum is asymptotic to

$$\int_{\mathbb{R}^n}\exp\left(-2\pi\sum_{j=1}^{n}\tau_j\left[\left(\frac{\mu^{(j)}+\sum_{s=1}\gamma_s^{(j)}x_s}{\delta^{(j)}}\right)^2+2\rho^{(j)}\left(\frac{\mu^{(j)}+\sum_{s=1}\gamma_s^{(j)}x_s}{\delta^{(j)}}\right)\right]\right)dx+H.O.T(\tau) \tag{5.7}$$

as $|\tau|\to 0$. Upon making the linear change of variables $y_j=\sum_{s=1}\gamma_s^{(j)}x_s$ and Proposition 1.3.1.15, the integral at 5.7 is equal to

$$\begin{aligned}&\frac{1}{\mathfrak{N}(\mathfrak{b}_1\mathfrak{f})\sqrt{|d_K|}}\int_{\mathbb{R}^n}\exp\left(-2\pi\sum_{j=1}^{n}\tau_j\left[\left(\frac{\mu^{(j)}+y_j}{\delta^{(j)}}\right)^2+2\rho^{(j)}\left(\frac{\mu^{(j)}+y_j}{\delta^{(j)}}\right)\right]\right)dy\\&=\frac{\mathrm{N}(\delta)}{\mathfrak{N}(\mathfrak{b}_1\mathfrak{f})\sqrt{|d_K|}}\prod_{j=1}^{n}\int_{\mathbb{R}}\exp\left(-2\pi\tau_j\left(y_j^2+2\rho^j y_j\right)\right)dy_j\\&=\frac{\mathrm{N}(\delta)}{\mathfrak{N}(\mathfrak{b}_1\mathfrak{f})\sqrt{2^n|d_K||\tau|}}+H.O.T(\tau).\end{aligned} \tag{5.8}$$

Upon comparing 5.8 and 5.6 and using the functional equation 5.5, we obtain generalised Hecke reciprocity.

Proposition 5.2.2.1. *Let ω and ρ be nonzero elements of a totally real number field K such that*

$$\mathfrak{d}(\omega)=\frac{\mathfrak{b}}{\mathfrak{a}},\quad\mathfrak{d}(\rho)=\frac{\mathfrak{c}}{\mathfrak{a}},$$

where $\mathfrak{a}, \mathfrak{b}$ *and* $\mathfrak{c}$ *are integral ideals such that* $(\mathfrak{a}, \mathfrak{b}) = 1 = (\mathfrak{a}, \mathfrak{c})$. *Let* $\mathfrak{b}_1$ *stand for the least common multiple of the denominators of* $\mathfrak{a}/4\mathfrak{b}^{-1}$ *and* $\mathfrak{c}/2\mathfrak{b}^{-1}$, *and let* γ *be any element of* K *such that* $\mathfrak{d}(\gamma)$ *is integral and relatively prime to* $\mathfrak{b}_1$. *Define*

$$C(\omega, \rho) = \sum_{\mu \in \mathfrak{a}} e^{2\pi i \mathrm{Tr}(\mu^2\omega + \mu\rho)}.$$

Then

$$\frac{C(\omega, \rho)}{|\sqrt{\mathfrak{N}(\mathfrak{a})}|} = \left| \frac{\sqrt{\mathfrak{N}(2\mathfrak{b})}}{\mathfrak{N}(\mathfrak{b}_1)} \right| \exp\left(\frac{\pi i}{4}\left(\mathrm{Tr}\,\mathrm{sgn}(\omega)) - 2\mathrm{Tr}\left(\frac{\rho}{\omega}\right)\right)\right) C\left(-\frac{\gamma^2}{4\omega}, \frac{\rho\gamma}{2\omega}\right).$$

We conjecture that, with the extended definition of $\mathrm{Tr}\,\mathrm{sgn}(\omega)$, Proposition 5.2.2.1 actually holds for arbitrary number fields.

5.3 Twisted Hecke reciprocity

Recall the twisted Landsberg–Schaar relation from Proposition 4.6.0.1, valid for ac even a primitive characters χ of order m (and specialised to the "ungeneralised" case, for simplicity):

$$\frac{G(\overline{\chi})}{\sqrt{|c|m}} \sum_{n=0}^{|c|m-1} \chi(n) \exp\left(\frac{\pi i a n^2}{cm}\right) = \frac{\exp\left(\frac{\pi i \,\mathrm{sgn}\, a}{4}\right)}{|a|} \sum_{n=0}^{|a|m-1} \overline{\chi(n)} \exp\left(-\frac{\pi i c n^2}{am}\right).$$

In this section, we derive an analogous statement for characters

$$\chi : (\mathcal{O}_K/\mathfrak{m})^\times \to \mathbb{S}^1,$$

where $\mathfrak{m}$ is an ideal in the same ideal class as the inverse different; that is, there exists some $y \in K$ so that

$$\mathfrak{d}(y) = \mathfrak{m}^{-1}.$$

In fact, we will assume even more; namely, that y is *totally real* (this is the condition that $\sigma_j(y) > 0$ for all $j = 1, \ldots, n$). One expects that the restriction on $\mathfrak{m}$ is completely unnecessary, and that all the results of this section hold true in some suitably modified form for arbitrary *Hecke characters*; however, our assumption has the advantage that the theta functions with which we will deal are particularly simple.

In order to further limit the length of the calculations, we continue to assume that K is totally real. As above, we expect to be able to obtain analogues of all results over *arbitrary* number fields.

5.3.1 Analytic proof

Let K be a totally real number field, and suppose that $\chi : (\mathcal{O}_K/\mathfrak{m})^\times \to \mathbb{S}^1$ is a Dirichlet character for some integral ideal $\mathfrak{m}$ such that there exists some *totally real* $y \in K$ with $\mathfrak{d}(y) = \mathfrak{m}^{-1}$. We define, for $t = (t_j) \in \mathscr{H}^n$,

$$\Theta_K(t; \chi, y) = \sum_{\mu \in \mathcal{O}_K} \chi(\mu) \exp\left(2\pi i \sum_{j=1}^{n} t_j y^{(j)} \left(\mu^{(j)}\right)^2\right).$$

The results of the last subsection will suffice for us to deduce a functional equation for $\Theta_K(t; \chi, y)$ without recourse to Poisson summation. Indeed, we have

$$\begin{aligned}
\Theta_K(t; \chi, y) &= \sum_{\mu \in \mathcal{O}_K} \chi(\mu) \exp\left(2\pi i \sum_{j=1}^{n} t_j y^{(j)} \left(\mu^{(j)}\right)^2\right) \\
&= \sum_{\mu \in \mathcal{O}_K/\mathfrak{m}} \chi(\mu) \sum_{\nu \in \mathfrak{m}} \exp\left(2\pi i \sum_{j=1}^{n} t_j y^{(j)} \left(\nu^{(j)} + \mu^{(j)}\right)^2\right) \\
&= \sum_{\mu \in \mathcal{O}_K/\mathfrak{m}} \chi(\mu) \exp\left(2\pi i \sum_{j=1}^{n} t_j y^{(j)} \left(\mu^{(j)}\right)^2\right) \sum_{\nu \in \mathfrak{m}} \exp\left(2\pi i \sum_{j=1}^{n} \left[t_j y^{(j)} \left(\nu^{(j)}\right)^2 + 2t_j y^{(j)} \nu^{(j)}\right]\right) \\
&= \sum_{\mu \in \mathcal{O}_K/\mathfrak{m}} \chi(\mu) \exp\left(2\pi i \sum_{j=1}^{n} t_j y^{(j)} \left(\mu^{(j)}\right)^2\right) \Theta_K(\mathfrak{m}, ty, 2ty), \qquad (5.9)
\end{aligned}$$

where $ty = (t_1 y^{(1)}, \ldots, t_n y^{(n)})$.

By 5.5,

$$\Theta_K(\mathfrak{m}, ty, 2ty) = \frac{1}{\mathfrak{N}(\mathfrak{m})\sqrt{|d_K|(-2i)^n|t|}} \exp\left(-2\pi i \sum_{j=1}^{n} \left(\mu^{(j)}\right)^2\right) \Theta_K\left(\mathfrak{m}^{-1}\mathfrak{d}^{-1}, -\frac{1}{4ty}, \mu\right), \tag{5.10}$$

where $-1/4ty = (-1/4t_j y^{(j)})$ and $\mu = (\mu^j)$. Substituting 5.10 into 5.9, we have

$$\Theta_K(t;\chi,y) = \frac{1}{\mathfrak{N}(\mathfrak{m})\sqrt{|d_K|(-2i)^n|t|}} \sum_{\mu\in\mathcal{O}_K/\mathfrak{m}} \chi(\mu) \sum_{\lambda\in\mathfrak{m}^{-1}\mathfrak{d}^{-1}} \exp\left(2\pi i \sum_{j=1}^{n}\left[-\frac{1}{4t_j y^{(j)}}\left(\lambda^{(j)}\right)^2 - \lambda^{(j)}\mu^{(j)}\right]\right)$$

$$= \frac{1}{\mathfrak{N}(\mathfrak{m})\sqrt{|d_K|(-2i)^n|t|}} \sum_{\lambda\in\mathfrak{m}^{-1}\mathfrak{d}^{-1}} \left(\sum_{\mu\in\mathcal{O}_K/\mathfrak{m}} \chi(\mu)\exp\left(-2\pi i\sum_{j=1}^{n}\lambda^{(j)}\mu^{(j)}\right)\right) \exp\left(-2\pi i \sum_{j=1}^{n}\frac{1}{4t_j y^{(j)}}\left(\lambda^{(j)}\right)^2\right).$$

By assumption, $\mathfrak{m}^{-1}\mathfrak{d}^{-1} = (y)$, so we may replace λ with νy, as ν runs over $\mathcal{O}_K$. With Definition 1.3.3.5,

$$\Theta_K(t;\chi,y) = \frac{1}{\mathfrak{N}(\mathfrak{m})\sqrt{|d_K|(-2i)^n|t|}} \sum_{\nu\in\mathcal{O}_K} G_y(\chi,\nu)\exp\left(-2\pi i\sum_{j=1}^{n}\frac{y^{(j)}}{4t_j}\left(\nu^{(j)}\right)^2\right)$$

and the result $G_y(\chi,\nu) = \overline{\chi(\nu)}G_y(\chi)$ (Lemma 1.3.3.6) allows us to conclude:

Proposition 5.3.1.1. *Let $t, y\chi$ and $\mathfrak{m}$ be as at the start of this subsection. Then*

$$\Theta_K(t;\chi,y) = \frac{G_y(\chi)}{\mathfrak{N}(\mathfrak{m})\sqrt{|d_K|(-2i)^n|t|}}\Theta_K\left(-\frac{1}{4t};\overline{\chi},y\right).$$

The next item is to compute the asymptotic expansion of $\Theta_K(t;\chi,y)$. Similarly to the proof of Proposition 5.3.1.1, we will be able to avoid the use of Euler–Maclaurin summation by employing results from the previous section.

Proposition 5.3.1.2. *Let $\omega \in K$ and write $\mathfrak{d}(\omega) = \mathfrak{b}\mathfrak{a}^{-1}$, where $\mathfrak{a}$ and $\mathfrak{b}$ are coprime integral ideals. Let y be a totally real element of K such that $\mathfrak{d}(y) = \mathfrak{m}^{-1}$, and let χ be a character of $\mathfrak{m}$. Set $\mathfrak{a}_1$ to be the lowest common multiple of $\mathfrak{m}$ and the denominator of $\mathfrak{d}(\omega y)$, and $\mathfrak{b}_1$ to be the lowest common multiple of $\mathfrak{m}$ and the denominator of $1/\mathfrak{m}(\omega)$. Then, with the notation*

$$C_{\chi,y}(\omega) = \sum_{\mu\in\mathcal{O}_{\mathfrak{a}_1}} e^{2\pi i \mathrm{Tr} y\omega\mu^2},$$

we have

$$\frac{1}{\mathfrak{N}(\mathfrak{a}_1)}C_{\chi,y}(\omega) = \frac{|\sqrt{\mathfrak{N}(2\omega)}|}{\mathfrak{N}(\mathfrak{m})\sqrt{|d_K|}}\frac{G_y(\chi)\exp\left(\frac{\pi i}{4}\mathrm{Tr}\,\mathrm{sgn}\,\omega\right)}{\mathfrak{N}(\mathfrak{b}_1)}C_{\overline{\chi},y}(-1/4\omega).$$

Proof. We have the decomposition

$$\Theta_K(w+i\epsilon;\chi,y) = \sum_{\mu\in\mathcal{O}_K/\mathfrak{a}_1}\chi(\mu)\exp\left(2\pi i\sum_{j=1}^{n}\omega^{(j)}y^{(j)}\left(\mu^{(j)}\right)^2\right)\sum_{\nu\in\mathfrak{a}_1}\exp\left(-2\pi\sum_{j=1}^{n}\epsilon_j y^{(j)}\left(\mu^{(j)}+\nu^{(j)}\right)^2\right)$$

$$= \sum_{\mu\in\mathcal{O}_K/\mathfrak{a}_1}\chi(\mu)e^{2\pi i\mathrm{Tr}(\omega y\mu^2)}\exp\left(-2\pi\sum_{j=1}^{n}\epsilon_j y^{(j)}\left(\mu^{(j)}\right)^2\right)\sum_{\nu\in\mathfrak{a}_1}\exp\left(-2\pi\sum_{j=1}^{n}\epsilon_j y^{(j)}\left(\left(\nu^{(j)}\right)^2+2\mu^{(j)}\nu^{(j)}\right)\right),$$

and by the calculations of Subsection 5.1.1 the inner sum is asymptotic to

$$\frac{1}{\mathfrak{N}(\mathfrak{a}_1)\sqrt{2^n|d_K|\mathrm{N}(y)|\epsilon|}} + H.O.T(\epsilon).$$

It follows that

$$\Theta_K(\omega+i\epsilon;\chi,y) = \frac{1}{\mathfrak{N}(\mathfrak{a}_1)\sqrt{2^n|d_K|\mathrm{N}(y)|\epsilon|}}\sum_{\mu\in\mathcal{O}_K/\mathfrak{a}_1}\chi(\mu)e^{2\pi i\mathrm{Tr}(\omega y\mu^2)} + H.O.T(\epsilon). \tag{5.11}$$

On the other hand,

$$\Theta_K\left(-\frac{1}{4\omega}+i\tau;\overline{\chi},y\right)=\frac{1}{\mathfrak{N}(\mathfrak{b}_1)\sqrt{2^n|d_K|\mathrm{N}(y)|\tau|}}\sum_{\mu\in\mathcal{O}_K/\mathfrak{b}_1}\overline{\chi(\mu)}e^{2\pi i\mathrm{Tr}(-y\mu^2/4\omega)}+H.O.T(\tau). \tag{5.12}$$

Comparing 5.11 and 5.12 and using the functional equation 5.3.1.1, it transpires that

$$\frac{1}{\mathfrak{N}(\mathfrak{a}_1)}\sum_{\mu\in\mathcal{O}_K/\mathfrak{a}_1}\chi(\mu)e^{2\pi i\mathrm{Tr}(\omega y\mu^2)}=\frac{|\sqrt{\mathfrak{N}(2\omega)}|}{\mathfrak{N}(\mathfrak{m})\sqrt{|d_K|}}\frac{G_y(\chi)\exp\left(\frac{\pi i}{4}\mathrm{Tr}\,\mathrm{sgn}\,\omega\right)}{\mathfrak{N}(\mathfrak{b}_1)}\sum_{\mu\in\mathcal{O}_K/\mathfrak{b}_1}\overline{\chi(\mu)}e^{2\pi i\mathrm{Tr}(-y\mu^2/4\omega)},$$

as advertised. □

5.3.2 Towards local quartic Hecke reciprocity

The aim of this subsection is to obtain a version of Proposition 4.8.3.1 valid over totally real number fields. This aim goes unachieved, for it appears that the twisted version of Hecke reciprocity obtained in the previous subsection is insufficiently robust to produce a local quartic version without some fairly restrictive hypotheses. Since we have not yet given an elementary proof of Proposition 5.3.1.2 (for example, by reducing both sides to the usual Hecke sums along the lines of Subsection 4.6.2), we would not have been able to give an elementary proof in any case. One can see from the results we have proven for Hecke sums of degree two and four that an explicit quartic formula requires the evaluation of the quadratic Hecke sums. This provides further impetus for an elementary proof of Hecke reciprocity obtained by evaluating the Hecke sums, as was done over $\mathbb{Q}$ in Section 4.3, and over *quadratic* number fields by Boylan and Skoruppa [BS10].

For $\omega\in K$ such that

$$\mathfrak{d}(\omega)=\frac{\mathfrak{b}}{\mathfrak{a}},$$

and $(\mathfrak{a},\mathfrak{b})=1$, for $d\geq 1$ we define

$$C_d(\omega)=\sum_{\mu\in\mathcal{O}_K/\mathfrak{a}}e^{2\pi i\mathrm{Tr}(\mu^d\omega)}.$$

Proposition 5.3.2.1 (Product rule for Hecke sums)**.** *Let $\omega\in K$ such that*

$$\mathfrak{d}(\omega)=\frac{\mathfrak{b}}{\mathfrak{a}},$$

where $\mathfrak{a}=\mathfrak{a}_1\mathfrak{a}_2$ and $(\mathfrak{a}_1,\mathfrak{a}_2,\mathfrak{b})=1$. By Lemma 1.3.1.12 we may choose auxiliary ideals $\mathfrak{c}_1$ and $\mathfrak{c}_2$ such that

$$\mathfrak{a}_1\mathfrak{c}_1=(\alpha_1),\quad \mathfrak{a}_2\mathfrak{c}_2=(\alpha_2),$$

where α_1 and α_2 are integers of K and $(\mathfrak{a},\mathfrak{c}_1\mathfrak{c}_2)$=1, and therefore

$$\omega=\frac{\beta}{\alpha_1\alpha_2},\quad \textit{where } (\beta)=\frac{\mathfrak{b}\mathfrak{c}_1\mathfrak{c}_2}{\mathfrak{d}}.$$

Then we have

$$C_d(\omega)=C_d\left(\frac{\beta}{\alpha_1\alpha_2}\right)=C_d\left(\frac{\alpha_1^{d-1}\beta}{\alpha_2}\right)C_d\left(\frac{\alpha_2^{d-1}\beta}{\alpha_1}\right).$$

Proof. As μ and ν run over complete residue classes of $\mathcal{O}/\mathfrak{a}_1$ and $\mathcal{O}/\mathfrak{a}_2$ respectively, $\rho=\mu\alpha_2+\nu\alpha_1$ run over a complete residue class of $\mathcal{O}/\mathfrak{a}$. Therefore

$$C_d\left(\frac{\beta}{\alpha_1\alpha_2}\right)=\sum_{\rho\in\mathcal{O}_K/\mathfrak{a}}e^{2\pi i\mathrm{Tr}(\rho^d\beta/\alpha_1\alpha_2)}.$$

Since $\mathfrak{a}_1$ and $\mathfrak{a}_2$ are coprime, $\mathcal{O}_K/\mathfrak{a}=\mathcal{O}_K/\mathfrak{a}_1\times\mathcal{O}_K/\mathfrak{a}_2$, so

$$\sum_{\mu\in\mathcal{O}_K/\mathfrak{a}_1}\sum_{\nu\in\mathcal{O}_K/\mathfrak{a}_2}e^{2\pi i\mathrm{Tr}((\mu\alpha_2+\nu\alpha_1)^d\beta/\alpha_1\alpha_2)}=\sum_{\mu\in\mathcal{O}_K/\mathfrak{a}_1}e^{2\pi i\mathrm{Tr}(\mu^d\alpha_2^{d-1}/\alpha_1)}\sum_{\mu\in\mathcal{O}_K/\mathfrak{a}_1}e^{2\pi i\mathrm{Tr}(\mu^d\alpha_1^{d-1}/\alpha_2)},$$

and the claim follows. □

By induction, Proposition 5.3.2.1 extends to cover any number of coprime ideals appearing in the denominator of ω. Indeed, if we have an element $\omega \in K$ such that

$$\mathfrak{d}(\omega) = \frac{\mathfrak{b}}{\prod_i \mathfrak{a}_i},$$

where $(\mathfrak{a}_1, \dots, \mathfrak{a}_n, \mathfrak{b}) = 1$, then by Lemma 1.3.1.12 we may choose auxiliary ideals $\mathfrak{c}_1, \dots, \mathfrak{c}_n$ such that $\mathfrak{a}_i\mathfrak{c}_i = (\alpha_i)$ for some integer α of $\mathcal{O}_k$, and $(\mathfrak{a}_i, \mathfrak{c}_i) = 1$ for $i = 1, \dots, n$. It follows that $\mathfrak{b}\mathfrak{c}_1 \dots \mathfrak{c}_n/\mathfrak{d} = (\beta)$ for some integer $\beta \in \mathcal{O}_K$, and we have

$$C_d(\omega) = C_d\left(\frac{\beta}{\alpha_1 \dots \alpha_n}\right) = \prod_{i=1}^{n} C_d\left(\frac{\beta \prod_{j\neq i} \alpha_j^{d-1}}{\alpha_i}\right).$$

Proposition 5.3.2.2 (Reduction of degree for quartic Hecke sums)**.** *Let $\omega \in K$ such that*

$$\mathfrak{d}(\omega) = \frac{\mathfrak{b}}{\mathfrak{p}},$$

where $(\mathfrak{a}, \mathfrak{p}) = 1$ and $\mathfrak{p}$ is a prime ideal. Then if $\left(\frac{-1}{\mathfrak{p}}\right) = -1$, we have

$$C_4(\omega) = C_2(\omega).$$

Proof. By a modification of the argument in Proposition 1.3.3.4, we have

$$C_4(\omega) = \sum_{\mu\in\mathfrak{p}} e^{2\pi i \mathrm{Tr}(\mu^4\omega)} = \sum_{\mu\in\mathfrak{p}} \left(1 + \left(\frac{\mu}{\mathfrak{p}}\right)\right) e^{2\pi i \mathrm{Tr}(\mu^2\omega)},$$

and the sum over the residue symbol vanishes, as we see upon substituting $-\mu$ for μ. □

Let K be an algebraic number field. Let ω be an element of K such that

$$\mathfrak{d}(\omega) = \frac{\mathfrak{b}}{\mathfrak{p}\prod_i \mathfrak{q}_i},$$

where $\mathfrak{p}$ is an odd prime of K such that $\left(\frac{-1}{\mathfrak{p}}\right) = 1$; $\mathfrak{q}_i$ denotes a finite collection of odd primes of K such that $\left(\frac{-1}{\mathfrak{q}_i}\right) = -1$ for all $i = 1, \dots, m$. Suppose that $(\mathfrak{p}, \mathfrak{q}_1, \dots, \mathfrak{q}_m, \mathfrak{b}) = 1$. After picking auxiliary ideals $\mathfrak{c}_0, \mathfrak{c}_1, \dots \mathfrak{c}_m$ such that $\mathfrak{p}\mathfrak{c}_0 = (p)$, $\mathfrak{q}_i\mathfrak{c}_i = (q_i)$ and $(\mathfrak{p}\prod_i \mathfrak{q}_i, \mathfrak{c}_0 \prod_i \mathfrak{c}_i, (2)) = 1$, we may write $\omega = \beta/p\left(\prod_i q_i\right)$.

By Propositions 5.3.2.1 and 5.3.2.2,

$$C_4(\omega) = C_4\left(\frac{\beta\prod_i q_i^3}{p}\right)\prod_i C_4\left(\frac{\beta p^3\prod_{j\neq i} q_j^3}{q_i}\right) = C_4\left(\frac{\beta\prod_i q_i^3}{p}\right)\prod_i C_2\left(\frac{\beta p\prod_{j\neq i} q_j}{q_i}\right),$$

where we used the coprimality conditions to reduce the cubes. Then we have

$$\begin{aligned}
&C_4\left(\frac{\beta\prod_i q_i^3}{p}\right)\prod_i C_2\left(\frac{\beta p\prod_{j\neq i} q_j}{q_i}\right) \\
&\qquad = \left(C_2\left(\frac{\beta\prod_i q_i}{p}\right) + \sum_{\mu\in\mathcal{O}_K/\mathfrak{p}} \left(\frac{\mu}{\mathfrak{p}}\right) \exp\left(2\pi i \mathrm{Tr}\left(\frac{\mu^2\beta\prod_i q_i^3}{p}\right)\right)\right)\prod_i C_2\left(\frac{\beta p\prod_{j\neq i} q_j}{q_i}\right) \\
&= C_2\left(\frac{\beta\prod_i \gamma_i^3}{p}\right) + \left(\frac{\prod_i q_i}{\mathfrak{p}}\right)\sum_{\mu\in\mathcal{O}_K/\mathfrak{p}}\left(\frac{\mu}{\mathfrak{p}}\right)\exp\left(2\pi i\mathrm{Tr}\left(\frac{\mu^2\beta\prod_i\gamma_i^3}{p}\right)\right)\prod_i\sum_{\nu_i\in\mathcal{O}_K/\mathfrak{q}_i}\exp\left(2\pi i\mathrm{Tr}\left(\frac{\nu_i^2\beta p\prod_{j\neq i}\gamma_j}{q_i}\right)\right).
\end{aligned}$$

As μ runs through a complete residue class modulo $\mathfrak{p}$ and the ν_i run through complete residue classes modulo $\mathfrak{q}_i$,

$$\lambda = \mu\prod_i \gamma_i + \sum_i \nu_i p \prod_{j\neq i}\gamma_j$$

runs through a complete residue class modulo $\mathfrak{p}\prod_i \mathfrak{q}_i$. Therefore

$$C_4(\omega) = C_2(\omega) + \sum_{\rho \in \mathcal{O}_K/\mathfrak{p}\prod_i \mathfrak{q}_i} \left(\frac{\lambda}{\mathfrak{p}}\right) \exp\left(2\pi i \mathrm{Tr}\left(\frac{\lambda^2\beta}{p\prod_i q_i}\right)\right).$$

If ω satisfies all the conditions as above, then

$$C_4\left(\frac{\omega}{4}\right) = C_4\left(\frac{\beta}{4p\prod_i q_i}\right) = C_4\left(\frac{\beta p^3 \prod_i q_i^3}{4}\right) C_4\left(\frac{4^3\beta}{p\prod_i q_i}\right),$$

so using the proof of Proposition 5.3.2.2 together with Hecke's lemma that Gauss sums with denominator 4 are nonzero [Hec81, §59, Lemma (b)], we obtain

$$C_4\left(\frac{\omega}{4}\right) = \frac{C_4\left(\frac{\beta p^3\prod_i q_i^3}{4}\right)}{C_2\left(\frac{\beta p\prod_i q_i}{4}\right)} C_2\left(\frac{\beta p\prod_i q_i}{4}\right)\left(C_2(4\omega) + \left(\frac{4}{\mathfrak{p}}\right)\sum_{\lambda \in \mathcal{O}_K/\mathfrak{p}\prod_i \mathfrak{q}_i} \left(\frac{\lambda}{\mathfrak{p}}\right) \exp\left(2\pi i \mathrm{Tr}\left(\frac{\lambda^2 4\beta}{p\prod_i q_i}\right)\right)\right).$$

But $\left(\frac{4}{\mathfrak{p}}\right) = 1$, so we may rearrange to obtain an analogue of 4.63:

$$C_4\left(\frac{\omega}{4}\right) = \frac{C_4\left(\frac{\beta p^3\prod_i q_i^3}{4}\right)}{C_2\left(\frac{\beta p\prod_i q_i}{4}\right)}\left(C_2\left(\frac{\omega}{4}\right) + \sum_{\lambda \in \mathcal{O}_K/\mathfrak{p}\prod_i \mathfrak{q}_i} \left(\frac{\lambda}{\mathfrak{p}}\right) \exp\left(2\pi i \mathrm{Tr}\left(\frac{\lambda^2\beta}{4p\prod_i q_i}\right)\right)\right).$$

The reader will now appreciate that, in order to proceed to a statement of local quartic Hecke reciprocity, we must place additional restrictions on the different, and the prime $\mathfrak{p}$ in order to be able to apply 5.3.1.2. This we do not actually carry out, as it seems to involve an excess of computation for little gain (we would, for example, obtain a statement for real quadratic fields).

Instead, we defer to future investigations the enjoyable task of computing the asymptotic expansions of Hecke theta functions [Neu99, Chapter VII, Section 7], which the author believes will yield the "correct" twisted Hecke reciprocity laws, which should in turn bring about a smooth proof of a local quartic version of Hecke's reciprocity.

Chapter 6

Asymptotic expansions of metaplectic Eisenstein series

In the final paragraph of his treatise *Vorlesungen über die Theorie der algebraischen Zahlen*, Hecke makes the following observation [Hec23]:

> *Man hat bisher noch nicht solche transzendenten Funktionen entdeckt, welche, wie die Thetafunktionen unserer Theorie, eine Reziprozitätsbeziehung zwischen den Summen ergeben, die für höhere Potenzreste an Stelle der Gaußschen Summen treten.*[10]

Hecke's words deserve some attention. In particular, it is unlikely that by "the sums which occur for higher power residues in place of the Gauss sums", he means to refer to sums of the shape

$$S_{K,n} = \sum_{\mu \in \mathcal{O}_K/\mathfrak{p}} \left(\frac{\mu}{\mathfrak{p}}\right)_n e^{2\pi i \mathrm{Tr}(\mu x)},$$

where $\mathfrak{d}(x) = \mathfrak{p}^{-1}$ and $\left(\frac{\cdot}{\mathfrak{p}}\right)_n$ is the nth power residue symbol in $K \supseteq \mathbb{Q}(\mu_n)$, as it has been suspected since at least the time of Kummer [Kum42] that the behaviour of $S_{K,n}$ is too erratic to permit for S_k any kind of reciprocity law comparable to the Landsberg–Schaar relation. It is clear from the final part of Hecke's quote that he is instead interested in obtaining sums which *do* admit a reciprocity law akin to the Landsberg–Schaar relation, whilst also maintaining a connection to Hilbert reciprocity. In this chapter, we propose a process by which, from the nth power residue symbol in any number field containing the nth roots of unity, a reciprocity relation might be obtained in such a manner that when $n = 2$, one recovers Hecke reciprocity (and for $K = \mathbb{Q}$, the Landsberg–Schaar relation).

The point of this chapter is to make absolutely explicit how the quadratic Gauss sums, together with their reciprocity law, arise from quadratic reciprocity. The process is mediated by theta functions. In Section 6.2, we mention that, beginning with the law of quadratic reciprocity, one may construct a degree two cover of $\Gamma_0(4)$, on which there exists a theory of automorphic forms. There exists a canonical family of Eisenstein series on this cover, parametrised by a complex variable s, such that their Fourier coefficients may be worked out explicitly and turn out to be finite sums of L-functions. The Eisenstein series have a meromorphic continuation into the whole s-plane, with a single pole at $s = \frac{1}{2}$, and the residue at $s = \frac{1}{2}$ is a constant multiple of Jacobi's theta function. Thus, we build from the quadratic reciprocity law a distinguished automorphic form, which comes with the guarantee of a transformation law between asymptotic expansion at different representatives of the cusps.

In order to complete the calculation, we advocate the method of Mellin transforms to compute the asymptotic expansion of the Eisenstein series. This procedure is well-suited to the computation of asymptotic expansions of automorphic forms, as it reduces the problem to the calculation of the location and residues of the singularities of a family of twists of L-functions by Dirichlet characters. In the Section 6.1, we carry out this process in full for Jacobi's theta function, and the reader will note that the difficulty in recognising the final result as being the sum

[10] No transcendental functions have yet been discovered which, like the theta-functions of our theory, yield a reciprocity relation between the sums which occur for higher power residues in place of the Gauss sums. Translation by George U. Brauer and Jay R. Goldman, with the assistance of R. Kotzen [Hec81].

occurring in the Landsberg–Schaar relation increases with the complexity of the factors of the denominator of the rational number representing the chosen cusp.

In Subsection 6.2.3, we indicate how one might go about obtaining the asymptotic expansion of the Eisenstein series for any fixed s. No new techniques are required, but we require much more information concerning the locations of the poles and zeros of Dirichlet L-functions (of quadratic characters). The purpose of the calculation is simply to provide evidence for the assumption that no more interesting sums than the Landsberg–Schaar relation can possibly arise from the asymptotic expansion of the Eisenstein series.

In the final section, we indicate how theta functions on metaplectic covers of $GL(n)$ arise canonically from Hilbert reciprocity over number fields, in order to lend credence to our tentative hypothesis that metaplectic theta functions are the "transcendental functions" sought by Hecke. For a totally complex number field, we sketch, using $\mathbb{Q}(i)$ as a case study, that one obtains the familiar theta function for such a number field, and we outline the salient points of Bump and Hoffstein's construction [BH86] of a cubic theta function over $\mathbb{Q}(\omega)$, where ω is a primitive cube root of unity.

Lastly, we outline the analysis, due to Kazhdan and Patterson [KP84], of theta functions on n-fold metaplectic covers of $GL(n)$ determined by the n-th power Hilbert symbol, and we note that a generalisation of the relation between Fourier coefficients observed in Subsection 6.2.2 holds, thus suggesting that the asymptotic expansions of such functions may give rise to the kind of sums envisaged by Hecke.

6.1 The Landsberg–Schaar relation and the Riemann zeta function

In this section, we will acquaint the reader with the important analytical result mentioned in the introduction to this chapter, namely that the asymptotic expansion of a function may be determined by the locations and residues of the poles of its Mellin transform. We subsequently apply this theorem to Jacobi's theta function, obtaining another proof of the Landsberg–Schaar relation which makes use of the fact that the Mellin transform of the theta function is an L-function, as opposed to relying on the special form of its Fourier coefficients.

6.1.1 Asymptotic expansions and Mellin transforms

The results of this section rest on the next theorem, relating the asymptotic expansion of a sufficiently well-behaved function to the singularities of its Mellin transform.

Theorem 6.1.1.1 ([Doe55, 6. Kapitel, § 3, pp. 115]). *Suppose that*

1. *$F(z)$ is analytic in a left half plane $\mathrm{Re}(z) \le c$, except for poles or essential singularities at the points $\lambda_0, \lambda_1, \lambda_2, \dots$, where $c > \mathrm{Re}(\lambda_0) > \mathrm{Re}(\lambda_1) > \mathrm{Re}(\lambda_2) > \dots$.*
2. *The principal part of the Laurent expansion of $F(z)$ at $z = \lambda_\nu$ is*
$$b_1^{(\nu)}(z-\lambda_\nu)^{-1} + b_2^{(\nu)}(z-\lambda_\nu)^{-2} + \dots + b_{r_\nu}^{(\nu)}(z-\lambda_\nu)^{r_\nu} + \dots$$
3. *In every strip of finite width $c_0 \le \mathrm{Re}(z) \le c$, $F(z) \to 0$ uniformly in $\mathrm{Re}(z)$ as $|\mathrm{Im}(z)| \to \infty$.*
4. *Between every pair of consecutive singularities λ_ν and $\lambda_{\nu+1}$ there exists some real β_ν such that $\mathrm{Re}(\lambda_{\nu+1}) < \beta_\nu < \mathrm{Re}(\lambda_\nu)$, such that the integral*
$$\int_{-\infty}^{\infty} x^{-i\tau} F(\beta_\nu + i\tau) d\tau$$
converges uniformly for $0 < x \le x_\nu$.

Then
$$f(x) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} x^{-z} F(z) dz$$
converges for $0 < x \le x_0$ and
$$f(x) = \sum_{\nu=0}^{n}\left(b_1^{(\nu)} + \frac{b_2^{(\nu)}}{1!}(-\log(x)) + \dots + \frac{b_{r_\nu}^{(\nu)}}{(r_\nu-1)!}(-\log(x))^{r_\nu-1} + \dots\right)x^{-\lambda_\nu} + \frac{1}{2\pi i}\int_{\beta_n-i\infty}^{\beta_n+i\infty} x^{-z}F(z)dz,$$
where the last term, $(2\pi i)^{-1}\int_{\beta_n-i\infty}^{\beta_n+i\infty} x^{-z}F(z)dz$, is $O(x^{-\beta_n})$ as $x \to 0$.

We will soon be dealing with functions which are not holomorphic in $z = x + iy$, but merely eigenfunctions of certain systems of linear elliptic differential operators. In order to be able to use transformation laws satisfied by such functions to prove identities for expressions arising from asymptotic expansions, it is important to be able to compare the asymptotic expansion of functions $f(x_0 + i\epsilon)$ as $\epsilon \to 0^+$ with $f(x_0 + i\gamma(\epsilon))$ where γ is a path into $\mathbb{C}$. As in Chapters 4 and 5, we may prove that the leading terms are equal simply by noting that each term in the asymptotic expansion of $f(x_0, \epsilon)$ is, at least locally, the restriction to real ϵ of a holomorphic function *of* ϵ. Since a holomorphic function which vanishes along a segment of the imaginary axis is zero, the asymptotic expansion holds for all complex $\epsilon \in \mathscr{H}$ for which the holomorphic extensions are defined, and so for ϵ sufficiently close to zero, we may replace ϵ by $\gamma(\epsilon)$ in the asymptotic expansion.

6.1.2 Dirichlet L-functions and yet another proof of the Landsberg–Schaar relation

Suppose that $f(z)$ is a 1-periodic function on the upper half plane $\mathscr{H}$, possessing a representation as an absolutely convergent Fourier series:

$$f(z) = \sum_{n=1}^{\infty} c_n e^{2\pi i n z},$$

and an associated Dirichlet series

$$D(f, s) = \sum_{n=1}^{\infty} c_n n^{-s}.$$

Let a and b be integers, with a positive and $\mathrm{Re}(\epsilon) > 0$. Then

$$\begin{aligned} f\left(\frac{b}{a} + i\epsilon\right) &= \sum_{n=0}^{a-1} \exp\left(\frac{2\pi i n b}{a}\right) \sum_{\substack{m=1 \\ m=n \bmod a}}^{\infty} c_m e^{-2\pi m\epsilon} \\ &= \sum_{d|a} \sum_{\substack{n=0 \\ (n,a)=d}}^{a-1} \exp\left(\frac{2\pi i n b}{a}\right) \sum_{\substack{m=1 \\ d|m \\ m/d=n/d \bmod a/d}}^{\infty} c_m e^{-2\pi m\epsilon} \\ &= \sum_{d|(n,a)} \frac{1}{\phi(a/d)} \sum_{\substack{n=0 \\ (n,a)=d}}^{a-1} \exp\left(\frac{2\pi i n b}{a}\right) \sum_{|\chi|=a/d} \overline{\chi(n/d)} \sum_{m=1}^{\infty} c_{dm}\chi(m) e^{-2\pi d m\epsilon}. \end{aligned} \tag{6.1}$$

Now set

$$D_{\chi,d}(f, s) = \sum_{n=1}^{\infty} c_{dn}\chi(n)(dn)^{-s},$$

and assume that for all Dirichlet characters χ, $D_\chi(f, s)$ satisfies the conditions of Theorem 6.1.1.1. Then by analytic continuation, the asymptotic behaviour of $f\left(\frac{b}{a} + i\epsilon\right)$ as $\epsilon \to 0$ is governed by the residues of the Mellin transform of the inner sum at 6.1,

$$(2\pi)^{-s}\,\Gamma(s) D_{\chi,d}(f, s).$$

We specialise to

$$f(z) = \frac{1}{2}\left(\theta(z) - 1\right) = \sum_{n=1}^{\infty} e^{2\pi i n^2 z}. \tag{6.2}$$

Let d' denote the squarefree part of d (so that if d is a square, $d' = 1$), and observe that c_{dm} vanishes unless m is of the form $d'n^2$ for some positive integer n, in which case it is equal to unity. Therefore the Dirichlet series we are dealing with are of the form

$$D_{\chi,d}(f, s) = \sum_{n=1}^{\infty} \chi(d'n^2)(dd'n^2)^{-s} = \chi(d')(dd')^{-s} \sum_{n=1}^{\infty} \chi^2(n) n^{-2s} = \chi(d')(dd')^{-s} L(\chi^2, 2s),$$

where the final term is a Dirichlet L-series, defined for characters χ by

$$L(\chi, s) = \sum_{n=1}^{\infty} \chi(n) n^{-s}.$$

All information about residues of Dirichlet L-functions is contained in the next result, first stated as part of Proposition 1.3.4.2.

Theorem 6.1.2.1. *Let χ be a primitive even character modulo m. Then $L(\chi, s)$ is entire unless χ is the principal character, in which case $L(\chi, s)$ has a simple pole at $s = 1$ with residue $\phi(m)/m$.*

Applying Theorem 6.1.2.1 to $(2\pi)^{-s}\Gamma(s)D_{\chi,d}(f, s)$ as defined above, we have

$$\begin{aligned}\mathrm{Res}_{s=\frac{1}{2}}(2\pi)^{-s}\Gamma(s)D_{\chi,d}(f,s) &= \chi(d')\mathrm{Res}_{s=\frac{1}{2}}(2\pi d d')^{-s}\Gamma(s)L_{\chi^2,d}(f,2s)\\ &= \begin{cases}\dfrac{2\phi(a/d)}{a\sqrt{2}}\sqrt{\dfrac{d}{d'}}\chi(d') & \chi^2 = 1\\ 0 & \chi^2 \neq 1.\end{cases}\end{aligned}$$

Then by Theorem 6.1.1.1, the asymptotic expansion for $2f(b/a + i\epsilon)$ as $\epsilon \to 0$ takes the form

$$2f(b/a+i\epsilon) = \left(\frac{1}{a\sqrt{2}}\sum_{d|a} d'' \sum_{\substack{n=0\\(n,a)=d}}^{a-1} \exp\left(\frac{2\pi i n b}{a}\right) \sum_{\substack{|\chi|=a/d\\ \chi^2=1}} \overline{\chi(n/d)}\chi(d')\right)\epsilon^{-1/2} + O(1), \tag{6.3}$$

where d'' denotes the square part of d: $d = (d'')^2 d'$ where d' is squarefree.

Remark 6.1.2.2. *The completed L-function*

$$\Lambda(\chi, s) = (2\pi)^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)L(\chi, s)$$

has a simple pole at $s = 0$, caused by a pole of the gamma function, and depending on the parity of χ, it may also have poles for $s = -2, -4, -6, \ldots$.

The pole at $s = 0$ causes the constant term $-1/2$, which we expect to appear in the asymptotic expansion of f by its very definition (6.2), and one may check using similar techniques to those below that the sums arising from residues at $s = -2, -4, -6, \ldots$ vanish. So the only interesting term in the asymptotic expansion of $f(b/a + i\epsilon)$ arises from the residue at $s = 1$, as expected.

We wish to prove that

$$\sum_{d|a} d'' \sum_{\substack{n=0\\(n,a)=d}}^{a-1} \exp\left(\frac{2\pi i n b}{a}\right) \sum_{\substack{|\chi|=a/d\\ \chi^2=1}} \overline{\chi(n/d)}\chi(d') = \sum_{n=1}^{a-1}\exp\left(\frac{2\pi i n^2 b}{a}\right). \tag{6.4}$$

Note that we may drop the conjugate line over $\chi(n/d)$ since characters of order 2 are real. We set $a = da'$ and $n = dn'$, and using this notation the right hand side of 6.4 may be rewritten as

$$\sum_{\substack{a'|a\\(a',d')=1}} d'' \sum_{\substack{n'=0\\(n',a')=1}}^{a'-1} \exp\left(\frac{2\pi i n' b}{a'}\right) \sum_{\substack{|\chi|=a'\\ \chi^2=1}} \chi(n'd').$$

Since $(a', d') = 1$, we can make the change of variables $n' \mapsto d'n$ in the middle sum of 6.5 to obtain

$$\sum_{\substack{a'|a\\(a',d')=1}} d'' \sum_{\substack{n'=0\\(n',a')=1}}^{a'-1} \exp\left(\frac{2\pi i n' d' b}{a'}\right) \sum_{\substack{|\chi|=a'\\ \chi^2=1}} \chi(n'). \tag{6.5}$$

We now drop the prime adorning n. The next step is to simplify the innermost sum, which we may rewrite as

$$\sum_{\substack{|\chi|=a'\\ \chi^2=1}} \chi(n) = \sum_{\chi\in(\mathbb{Z}/a'\mathbb{Z})^\times[2]} \chi(n),$$

where we define

$$G[m] = \{g \in G \mid g^m = e\}$$

for groups G and positive integers m. One easily verifies that the map

$$\phi(\chi)(g + G^m) = \chi(g)$$

defines an isomorphism between $\hat{G}[m]$ and $\widehat{G/G^m}$, and so using 1.3, we see that

$$\sum_{\substack{|\chi|=a' \\ \chi^2=1}} \chi(n) = \begin{cases} |\widehat{(\mathbb{Z}/a'\mathbb{Z})^\times}[2]| & n \text{ is a square modulo } a' \\ 0 & \text{otherwise.} \end{cases} \tag{6.6}$$

Note that

$$|\widehat{(\mathbb{Z}/a'\mathbb{Z})^\times}[2]| = |(\mathbb{Z}/a'\mathbb{Z})^\times[2]| = \#\{x \in (\mathbb{Z}/a'\mathbb{Z}) \mid x^2 = 1\}.$$

Consequently, substituting 6.6 into the expression 6.5, we obtain

$$\begin{aligned}
\sum_{\substack{a'|a \\ (a',d')=1}} d'' \left|(\mathbb{Z}/a'\mathbb{Z})^\times[2]\right| \sum_{\substack{n=0 \\ (n,a')=1 \\ n=\square \bmod a'}}^{a'-1} \exp\left(\frac{2\pi i n d' b}{a'}\right) &= \sum_{\substack{a'|a \\ (a',d')=1}} d'' \sum_{\substack{n=0 \\ (n,a')=1}}^{a'-1} \exp\left(\frac{2\pi i n^2 d' b}{a'}\right) \\
&= \sum_{\substack{a'|a \\ (a',d')=1}} \sum_{\substack{n=0 \\ (n,a')=1}}^{a'd''-1} \exp\left(\frac{2\pi i n^2 d' b}{a'} \frac{(d'')^2 d'}{(d'')^2 d'}\right) \\
&= \sum_{\substack{a'|a \\ (a',d')=1}} \sum_{\substack{n=0 \\ (n,a')=1}}^{a'd''-1} \exp\left(\frac{2\pi i (d''d'n)^2 d' b}{a'd}\right) \\
&= \sum_{\substack{a'|a \\ (a',d')=1}} \sum_{\substack{n=0 \\ (d''n,a)=d}}^{a-1} \exp\left(\frac{2\pi i n^2 b}{a}\right).
\end{aligned}$$

Therefore, 6.4 will be proved once we know that

$$\sum_{\substack{a'|a \\ (a',d')=1}} \sum_{\substack{n=0 \\ (d''n,a)=d}}^{a-1} \exp\left(\frac{2\pi i n^2 b}{a}\right) = \sum_{n=0}^{a-1} \exp\left(\frac{2\pi i n^2 b}{a}\right),$$

for which it suffices to show that

$$\bigcup_{\substack{a'|a \\ (a',d')=1}} S_{a'} = \{0, 1, \ldots, a-1\} \text{ and } S_{a'_1} \cap S_{a'_2} = \varnothing \text{ if } a'_1 \neq a'_2, \tag{6.7}$$

where

$$S_{a'} = \{n \in \{0, 1, \ldots, a-1\} \mid (d''n, a) = d\}.$$

In other words, we must prove the following result:

Proposition 6.1.2.3. *For a given integer n between 0 and $a-1$, the system of equations*

$$(d''n, a) = d, \qquad (a, dd') = d, \tag{6.8}$$

has a unique solution d.

In addition to the notation already defined in this Subsection, we let p denote a prime and define $\nu_p(x)$ to be the highest power of p which divides x.

Proof. We begin by proving that a solution exists. We set $d = (a, (a, n)^2)$ and show that d satisfies 6.8. Our strategy is to show that $\nu_p((d''n, a)) = \nu_p(d)$ and $\nu_p((a, dd')) = \nu_p(d)$ for each prime p. Since

$$\nu_p(d) = \min(\nu_p(a), 2\min(\nu_p(a), \nu_p(n))),$$

we proceed case-by-case.

Case 1: $\nu_p(a) \leq 2\min(\nu_p(a), \nu_p(n))$.

This assumption means that $\nu_p(d) = \nu_p(a)$. We first verify that d satisfies the leftmost equation at 6.8. Aiming for a contradiction, we suppose that $\nu_p(d) = \nu_p(d''n) < \nu_p(a)$. Then $\nu_p(n) < \nu_p(a)$, so $v_p(a) \leq 2\nu_p(n)$ by our first assumption. Suppose that $\nu_p(d)$ is even. Then $\nu_p(d'') = \frac{1}{2}\nu_p(d)$, and therefore

$$\nu_p(n) < \nu_p(d) - \nu_p(d'') = \frac{1}{2}\nu_p(d) = \frac{1}{2}\nu_p(a).$$

So $\nu_p(a) \leq 2\nu_p(n) < \nu_p(a)$: contradiction. Now suppose that $\nu_p(d)$ is odd. Then $\nu_p(n) < \frac{1}{2}(\nu_p(a) + 1)$ and since $\nu_p(a)$ is odd, $\nu_p(a) \leq 2\nu_p(n) - 1$. So $\nu_p(a) \leq 2\nu_p(n) - 1 < \nu_p(a)$: contradiction again. All told,

$$\nu_p((d''n, a)) = \nu_p(a) = \nu_p(d).$$

Now we verify the rightmost equation at 6.8. This is straightforward:

$$\nu_p((a, dd')) = \min(\nu_p(a), \nu_p(d) + \nu_p(d')) = \nu_p(a) = \nu_p(d).$$

Case 2: $\nu_p(a) > 2\min((\nu_p(a), \nu_p(n))$.

Note that $\nu_p(a) > \nu_p(n)$, so we have the stronger inequality $\nu_p(a) > 2\nu_p(n)$. Then $\nu_p(d) = 2\nu_p(n)$, so $\nu_p(d'') = \nu_p(n)$, and it follows that

$$\nu_p((d''n, a)) = \min(2\nu_p(n), \nu_p(a)) = 2\nu_p(n) = \nu_p(d),$$

and the leftmost equation is proved. On the other hand, since $\nu_p(d) = 2\nu_p(n)$ is even, $\nu_p(d') = 0$, so

$$\nu_p((a, dd')) = \min(\nu_p(a), 2\nu_p(n)) = 2\nu_p(n) = \nu_p(d),$$

and thus the rightmost equation is proved.

Now we prove that the solution is unique. We require a lemma:

Lemma 6.1.2.4. *Suppose $(a, dd') = d$. If $\nu_p(a)$ is even, then $\nu_p(d)$ is even. If $\nu_p(a)$ and $\nu_p(d)$ are both odd, then $\nu_p(a) = \nu_p(d)$.*

Proof. We have

$$\nu_p(d) = \min(\nu_p(a), \nu_p(d) + \nu_p(d')),$$

so the first part of the claim is clear if $\nu_p(a) \leq \nu_p(d) + \nu_p(d')$. If $\nu_p(a) > \nu_p(d) + \nu_p(d')$, then $\nu_p(d) = \nu_p(d) + \nu_p(d')$, so $\nu_p(d') = 0$. Now suppose that $\nu_p(d) = 2k + 1$, $\nu_p(a) = 2l + 1$. Then

$$2k + 1 = \min(2l + 1, 2k + 2),$$

so $k = l$. □

Returning to the proof of uniqueness, suppose that there exist divisors d_1 and d_2 of a such that

$$(d_1''n, a) = d_1 = (a, d_1 d_1') \text{ and } (d_2''n, a) = d_2 = (a, d_2 d_2');$$

that is, d_1 and d_2 both satisfy 6.8. Our strategy is to prove that $\nu_p(d_1) = \nu_p(d_2)$ for all primes p, and once again we proceed case-by-case:

Case 1: $\nu_p(d_1) = 2\mu$ is even.

Since $\nu_p(d_1) = \min(\nu_p(d_1''n), \nu_p(a))$, we divide our attention between two sub-cases.

Case 1.a): $\nu_p(a) < \nu_p(d_1''n)$.

Then $\nu_p(a) = \nu_p(d_1)$ is even, and by Lemma 6.1.2.4, $\nu_p(d_2)$ is even. We also have $\nu_p(d_1) = \mu$, so the inequality for this sub-case implies that $\nu_p(n) > \mu$. Set $s = \nu_p(d_2)$ and $t = \nu_p(n)$. Since d_2 satisfies 6.8, we have $s = \min\left(\frac{1}{2}s + t, 2\mu\right)$. If $s = \frac{1}{2}s + t$, then $s = 2t$, so $\nu_p(d_2) > 2\mu = \nu_p(a)$: contradiction. So

$$\nu_p(d_2) = s = 2\mu = \nu_p(d_1)$$

.

Case 1.b): $\nu_p(a) \geq \nu_p(d_1''n)$.

Then $\nu_p(d_1) = \nu_p(d_1'') + \nu_p(n)$, so $\nu_p(n) = \mu$. Set $s = \nu_p(d_2)$ and $t = \nu_p(a)$.

If s is even, $s = \min\left(\frac{1}{2}s + \mu, t\right)$. If $s = t$, then $s = t \leq \frac{1}{2}s + \mu$, and so $s \leq 2\mu \leq \nu_p(a) = t$. Since $s = t$, we have equality throughout and $\nu_p(d_2) = 2\mu = \nu_p(d_1)$. If $s = \frac{1}{2}s + \mu$, we arrive at $\nu_p(d_1) = \nu_p(d_2)$ as well.

If s is odd, then $s = \min\left(\frac{1}{2}(s-1) + \mu, t\right)$. If $s = t$, then $s \leq 2\mu - 1 \leq \nu_p(a) - 1 = t - 1$: a contradiction. So $s = \frac{1}{2}(s-1) + \mu$, which means that $\nu_p(d_2)$ is odd. By the first part of Lemma 6.1.2.4, $\nu_p(a)$ cannot be even, so the second part of Lemma 6.1.2.4 implies that $2\mu - 1 = \nu_p(d_2) = \nu_p(a)$. But $2\mu = \nu_p(d_1''n) \leq \nu_p(a)$, so we have another contradiction.

Case 2: $\nu_p(d_1) = 2\mu + 1$ is odd.

Again, we have $\nu_p(d_1) = \min\left(\nu_p(d_1''n), \nu_p(a)\right)$, so there are two sub-cases to consider.

Case 2.a): $\nu_p(d_1''n) \geq \nu_p(a)$.

Then $\nu_p(a) = 2\mu + 1$ and $\nu_p(n) \geq \mu + 1$. If $\nu_p(d_2)$ is odd then $\nu_p(d_2) = \nu_p(a) = \nu_p(d_1)$ by Lemma 6.1.2.4, so suppose that $s = \nu_p(d_2)$ is even. Since $s = \min\left(\frac{1}{2}s + \nu_p(n), 2\mu + 1\right)$, we must have $s = \frac{1}{2}s + \nu_p(n)$, so $s = 2\nu_p(n) \geq 2\mu + 2 > nu_p(a)$: contradiction.

Case 2.b): $\nu_p(d_1''n) < \nu_p(a)$.

By Lemma 6.1.2.4, $\nu_p(a)$ is not even, so $\nu_p(a) = \nu_p(d_1) = \nu_p(d_1''n)$: contradiction. □

So 6.7 is proved, which implies 6.4. It follows that the asymptotic expansion of Jacobi's theta function yields the Landsberg–Schaar relation, as claimed.

6.2 The metaplectic group and the Landsberg–Schaar relation over $\mathbb{Q}$

We use the following notation, some of which will be familiar from Subsection 2.2.3:

$$\Gamma_0(4) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{Z}) \text{ such that } c = 0 \bmod 4 \right\};$$

$$\Gamma_\infty = \left\{ \pm \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \text{ such that } m \in \mathbb{Z} \right\};$$

$$j(\gamma, z) = \begin{cases} \left(\dfrac{c}{d}\right) \epsilon_d^{-1} (cz + d)^{1/2} \text{ where } \left(\dfrac{c}{d}\right) \text{ and } \epsilon_d \text{ are as usual and } |\arg{(cz+d)^{1/2}}| < \dfrac{\pi}{2}, & d > 0, \\ j(-\gamma, z) & d < 0. \end{cases}$$

Consequently, $j(\gamma, z)$ satisfies the cocycle condition

$$j(\gamma_1\gamma_2, z) = j(\gamma_1, \gamma_2 z) j(\gamma_2, z), \tag{6.9}$$

for γ_1, γ_2 in $\Gamma_0(4)$ and $z \in \mathscr{H}$. We proved in Proposition 2.2.3.1 that

$$\theta(\gamma z) = j(\gamma, z)\theta(z),$$

which, together with the fact that θ is holomorphic at each of the three cusps of $\Gamma_0(4)$, shows that θ is a modular form of half-integer weight.

A more up-to-date perspective on modular forms is to view them as functions on quotients of $SL(2,\mathbb{R})$ by the appropriate discrete subgroups [Gel75, §2]. This perspective permits certain technical and conceptual simplifications, and is a natural starting point for the generalisation of the theory of modular forms in which $SL(2,\mathbb{R})$ is replaced by any reductive algebraic Lie group. The objects which replace modular forms are called *automorphic forms.*

However, when we attempt to associate to Jacobi's theta function to a function on a quotient of $SL(2,\mathbb{R})$, we find ourselves blocked by the nasty residue symbol appearing in the factor of automorphy. Weil observed [Wei64] that one can get around this problem by instead viewing theta functions as functions on quotients of a certain double cover of $SL(2,\mathbb{R})$, which he constructed using the quadratic residue symbol. He dubbed this double cover the *metaplectic group*, in analogy with symplectic groups, to which they are closely related.

Once one has a group on which automorphic forms might conceivably exist, the first item on the agenda is to construct Eisenstein series, via the "averaging" procedure of Subsection 2.1.2. For the metaplectic cover of $SL(2,\mathbb{R})$, the Eisenstein series have *two* variables, rather than one: a z-variable, which keeps track of *automorphic* behaviour; and an s-variable, in which the Eisenstein series behaves rather like a zeta function. In the s-variable, there is a pole at $s = 1/2$, and upon taking the residue there, we recover Jacobi's theta function. Another remarkable aspect of this process is that, by virtue of the fact that the Eisenstein series has been specifically manufactured to transform like a modular form of half-integral weight, Jacobi's theta function has appeared, out of thin air, already equipped with the transformation law Proposition 2.2.3.1!

In the interest of keeping things concrete, we dispense with the metaplectic group and tell the story using only the simpler group-theoretic machinery already in place from earlier chapters. We mention it only to impress upon the reader that the theta function may be constructed purely from the quadratic residue symbol, with no need for guesswork.

6.2.1 Construction of the Eisenstein series

We now construct some Eisenstein series associated canonically to $\Gamma_0(4)$ (and the quadratic residue symbol).

Definition 6.2.1.1. *Let $z = x + iy$. Then the Eisenstein series associated to $\Gamma_0(4)$ at the cusp at infinity is*

$$E_\infty(z,s) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(4)} j(\gamma,z)^{-1}\mathrm{Im}(\gamma z)^s.$$

It is easily checked that the sum defining $E_\infty(z,s)$ converges absolutely and uniformly if $\mathrm{Re}(s) > \frac{3}{4}$. As a consequence of the cocycle condition 6.9, for any such s, $E_\infty(z,s)$ is an *automorphic*[11] form of weight $\frac{1}{2}$ in the z-variable:

$$E_\infty(\gamma z,s) = j(\gamma,z)E_\infty(z,s) \text{ for all } \gamma\in\Gamma_0(4).$$

For the purposes of this chapter, it is easier to work with the Eisenstein series at the cusp at 0.

Definition 6.2.1.2. *The Eisenstein series associated to $\Gamma_0(4)$ at the cusp at 0 is*

$$E_0(z,s) = z^{-1/2}E_\infty(-1/4z,s).$$

Of course, for $\mathrm{Re}(s) > \frac{3}{4}$ we have

$$E_0(\gamma z,s) = j(\gamma,z)E_0(z,s) \text{ for all } \gamma\in\Gamma_0(4)$$

as well.

It is possible, using methods similar to those mentioned immediately after Proposition 1.3.4.1, to determine that $E_0(z,s)$, which converges only for $\mathrm{Re}(s)$ sufficiently large in Definition 6.2.1.1, admits a meromorphic extension in s to all of $\mathbb{C}$ and satisfies a functional equation. However, this will become clear after the next subsection, in which we compute the Fourier coefficients of $E_0(z,s)$, so we defer the proofs at this point.

[11] These functions are not, strictly speaking, modular forms in z for most fixed s. The condition in Definition 2.1.1.3 which is violated is (1), for the Eisenstein series of this chapter are merely analytic, as opposed to being holomorphic.

6.2.2 The Fourier expansion of $E_0(z,s)$

In this subsection we follow the treatment of Goldfeld and Hoffstein [GH85] in computing the Fourier expansion of the Eisenstein series of half-integral weight. They in turn follow the treatment of Shimura [Shi75; Shi72], who ultimately follows Maaß [Maa38]. As we do not require results at the level of generality of either Goldfeld and Hoffstein or Shimura in this particular subsection, we give our own proofs rather than referring to theirs. Since we are not aiming for maximal generality, the proofs below are somewhat more self-contained. The reader should keep in mind that Propositions 6.2.2.1 and 6.2.2.2 are very similar to the contentions of Subsection 2.1.2, and the proofs proceed along the same lines.

Proposition 6.2.2.1.

$$E_0(z,s) = \left(\frac{y}{4}\right)^s \sum_{\substack{(u,2v)=1\\ u>0}} \frac{\left(\frac{-v}{u}\right)\epsilon_u}{|v+uz|^{2s}(v+uz)^{1/2}}. \tag{6.10}$$

Proof. Each coset of $\Gamma_\infty\backslash\Gamma_0(4)$ may be represented uniquely by a matrix $\left(\begin{smallmatrix} a & b\\ c & d\end{smallmatrix}\right)$ in $SL(2,\mathbb{Z})$ with the property that $4 \mid c$, $d>0$ and $(c,d)=1$. Note that for all $\gamma \in SL(2,\mathbb{Z})$,

$$\mathrm{Im}(\gamma z)^s = |cz+d|^{-2s}\mathrm{Im}(z)^s.$$

Then we have

$$\begin{aligned} E_\infty\left(-\frac{1}{4z}, s\right) &= \sum_{\substack{c\in\mathbb{Z}, d>0\\ (4c,d)=1}} \left(\frac{4c}{d}\right)\epsilon_d^{-1}\left(-\frac{4c}{4z}+d\right)^{-\frac{1}{2}} \left|-\frac{4c}{4z}+d\right|^{-2s} \mathrm{Im}\left(-\frac{1}{4z}\right)^s\\ &= z^{-\frac{1}{2}} \sum_{\substack{c\in\mathbb{Z}, d>0\\ (2c,d)=1}} \left(\frac{-c}{d}\right)\epsilon_d^{-1}\left(c+dz\right)^{-\frac{1}{2}} |c+dz|^{-2s}|z|^{2s}\left(\frac{\mathrm{Im}(z)}{4|z|^2}\right)^s\\ &= z^{-\frac{1}{2}}\left(\frac{y}{4}\right)^s \sum_{\substack{c\in\mathbb{Z}, d>0\\ (2c,d)=1}} \frac{\left(\frac{-c}{d}\right)\epsilon_d^{-1}}{(c+dz)^{\frac{1}{2}}|c+dz|^{2s}}. \end{aligned}$$

□

We are now able to compute the Fourier expansion of $E_0(z,s)$.

Proposition 6.2.2.2.

$$E_0(z,s) = \sum_{m=-\infty}^{\infty} a_m(s,y)e^{2\pi imx},$$

where

$$\begin{aligned} a_m(s,y) &= \left(\frac{y}{4}\right)^s A_m(2s)K_m(s,y),\\ A_m(s) &= \sum_{\substack{n=1\\ (n,2)=1}}^{\infty} \frac{\epsilon_n g(-m,n)}{n^{s+\frac{1}{2}}} = \prod_{\substack{p \text{ prime}\\ p\neq 2}} \sum_{l=0}^{\infty} \frac{\epsilon_{p^l} g(-m,p^l)}{p^{l(s+1/2)}}\\ g(m,n) &= \sum_{a \bmod n} \left(\frac{a}{n}\right)\exp\left(\frac{2\pi iam}{n}\right),\\ K_m(s,y) &= \int_{-\infty}^{\infty} \frac{e^{-2\pi imx}}{(x^2+y^2)^s(x+iy)^{1/2}}dx. \end{aligned} \tag{6.11}$$

Note that $K_m(s,y)$ is holomorphic in s for $\mathrm{Re}(s) > \frac{1}{4}$ and $y \neq 0$, since the integrand is absolutely integrable by the estimate

$$\left|\frac{e^{-2\pi imx}}{(x^2+y^2)^s(x+iy)^{1/2}}\right| \leq \frac{\left(x^2+y^2\right)^{-|s|}}{|x+iy|^{1/2}} = \left(x^2+y^2\right)^{-|s|-1/4}.$$

Proof. By the Fourier inversion formula,

$$\begin{aligned}
a_m(s,y) &= \left(\frac{y}{4}\right)^s \int_0^1 \sum_{\substack{(u,2v)=1\\ u>0}} \frac{\left(\frac{-v}{u}\right)\epsilon_u}{|v+uz|^{2s}(v+uz)^{1/2}}dx \\
&= \left(\frac{y}{4}\right)^s \int_0^1 \sum_{\substack{u=1\\(u,2)=1}}^{\infty} \sum_{\substack{v=-\infty\\ v=k \bmod u}}^{\infty} \sum_{k=0}^{u-1} \frac{\left(\frac{-v}{u}\right)\epsilon_u e^{-2\pi i m x}}{u^{2s+1/2}\left|\frac{v}{u}+x+iy\right|^{2s}\left(\frac{v}{u}+x+iy\right)^{1/2}}dx \\
&= \left(\frac{y}{4}\right)^s \sum_{\substack{u=1\\(u,2)=1}}^{\infty} \frac{\epsilon_u}{u^{2s+1/2}} \sum_{k=0}^{u-1}\left(\frac{-k}{u}\right)e^{2\pi i m k/u}\int_0^1 \sum_{n=-\infty}^{\infty} \frac{e^{-2\pi i m(k/u+n+x)}}{\left|\frac{k}{u}+n+x+iy\right|^{2s}\left(\frac{k}{u}+n+x+iy\right)^{1/2}}dx \\
&= \left(\frac{y}{4}\right)^s \sum_{\substack{u=1\\(u,2)=1}}^{\infty} \frac{\epsilon_u g(-m,u)}{u^{2s+1/2}} \int_{-\infty}^{\infty} \frac{e^{-2\pi i m w}}{|w+iy|^{2s}(w+iy)^{1/2}}dw.
\end{aligned}$$

Thus most of Proposition 6.2.2.2 is proved; the second equality at 6.11 follows from the fact that the sequence $(\epsilon_n g(-m,n))_n$, where n runs over odd positive integers and m is any integer, is multiplicative. That is, if $(n_1,n_2)=1$, then

$$\epsilon_{n_1n_2}g(-m,n_1n_2) = \epsilon_{n_1n_2}\left(\frac{n_1}{n_2}\right)\left(\frac{n_2}{n_1}\right)g(-m,n_1)g(-m,n_2) = \epsilon_{n_1}g(-m,n_1)\epsilon_{n_2}g(-m,n_2), \tag{6.12}$$

where we have used the main law of quadratic reciprocity (Theorem 1.3.2.1). We refer the reader elsewhere [Kna92, Chapter VII, Section 2] for the straightforward proof that 6.12 implies that $A_m(s)$ may be represented as an Euler product. □

Proposition 6.2.2.3. *We have*

$$A_0(s) = \frac{\zeta(2s-1)}{\zeta(2s)}\frac{(1-2^{1-2s})}{(1-2^{-2s})}, \tag{6.13}$$

and for m squarefree,

$$A_m(s) = \frac{L(s,\chi_m)}{\zeta(2s)}\frac{(1-\chi_m(2)2^{-s})}{(1-2^{-2s})}, \tag{6.14}$$

where χ_m is defined by

$$\chi_m(n) = \begin{cases} \left(\frac{m}{n}\right) & m = 1 \bmod 4, \\ \left(\frac{4m}{n}\right) & m = 2,3 \bmod 4. \end{cases}$$

Note that χ_m is the real primitive Dirichlet character associated to the field $\mathbb{Q}(\sqrt{m})$: the zeta function $\zeta_{\mathbb{Q}(\sqrt{m})}(s)$ factors as $\zeta(s)L(\chi_m,s)$.

Now write $m = m_0n^2$, where m_0 squarefree. Then

$$A_m(s) = A_{m_0}(s)\sum_{\substack{d_2d_3|n\\(2,d_2d_3)=1}} \chi_{m_0}(d_3)\mu(d_3)d_2^{1-2s}d_3^{-s}, \tag{6.15}$$

where μ is the Möbius function.

The proof of Proposition 6.2.2.3 is rather long; however, it is difficult to find it in full in the literature. The earliest version is probably due to Maaß [Maa38], and Koblitz [Kob84, pp. 188–191, 195–200] presents a version for half-integral weight modular modular forms (for which the weight is at least 5/2) for $s=0$.

Proof. We begin by treating $A_0(s)$. It is easily checked that

$$g(0,n) = \begin{cases} \phi(n) & n \text{ is a square,} \\ 0 & \text{otherwise,} \end{cases}$$

where ϕ is Euler's totient function. Since $\phi(n^2) = n\phi(n)$, the Dirichlet series for $A_0(s)$ simplifies to

$$A_0(s) = \sum_{\substack{n=1 \\ (n,2)=1}}^{\infty} \frac{\phi(n)}{n^{2s}}.$$

But ϕ is a multiplicative function, so the Dirichlet series has an Euler product:

$$\sum_{\substack{n=1 \\ (n,2)=1}}^{\infty} \frac{\phi(n)}{n^{2s}} = \prod_{p\neq 2}\sum_{l=0}^{\infty} \frac{\phi(p^l)}{p^{2ls}} = \prod_{p}\sum_{l=0}^{\infty} \frac{\phi(p^l)}{p^{2ls}} \Big/ \sum_{l=0}^{\infty} \frac{\phi(2^l)}{4^{ls}} = \sum_{n=1}^{\infty} \frac{\phi(n)}{n^{2s}} \Big/ \sum_{l=0}^{\infty} \frac{\phi(2^l)}{4^{ls}}.$$

A basic property of the totient function is that for all primes and $l \geq 1$, $\phi(p^l) = p^{l-1}(p-1)$. Consequently, the final sum is a geometric series, and evaluates to

$$\frac{1-2^{-2s}}{1-2^{1-2s}}.$$

The Dirichlet series has a well-known expression in terms of the Riemann zeta function:

$$\sum_{n=1}^{\infty} \frac{\phi(n)}{n^{2s}} = \frac{\zeta(2s-1)}{\zeta(2s)},$$

so 6.13 is proved.

Before we go any further, we need to evaluate $g(-m, p^l)$. To start with,

$$g(m,p^l) = \begin{cases} 1 & l = 0, \\ \displaystyle\sum_{\substack{n \bmod p^l \\ (n,p)=1}} \exp\left(\frac{2\pi i m n}{p^l}\right) & l \text{ even}, \ l \geq 2, \\ \displaystyle\sum_{\substack{n \bmod p^l \\ (n,p)=1}} \left(\frac{n}{p}\right) \exp\left(\frac{2\pi i m n}{p^l}\right) & l \text{ odd}. \end{cases} = \begin{cases} 1 & l = 0 \\ c_{p^l}(m) & l \text{ even}, \ l \geq 2, \\ 0 & l \text{ odd}, \ \gcd(m,p) \neq p^{l-1}, \\ \epsilon_{p^l} p^{l-1/2} \left(\frac{m}{p}\right) & l \text{ odd}, \ \gcd(m,p) = p^{l-1}. \end{cases} \tag{6.16}$$

where the reader may recognise $c_{p^l}(m)$ as a Ramanujan sum, which can be evaluated as follows:

$$c_{p^l}(m) = \begin{cases} 0 & p^{l-1} \nmid m, \\ -p^{l-1} & p^{l-1} \mid m, p^l \nmid m, \\ \phi(p^l) & p^l \mid m. \end{cases} \tag{6.17}$$

The only assertion in 6.16 which is not immediately obvious is the statement that

$$g(m,p^l) = \begin{cases} 0 & l \text{ odd}, \ \gcd(m,p) \neq p^{l-1}, \\ \epsilon_{p^l} p^{l-1/2} \left(\frac{m}{p}\right) & l \text{ odd}, \ \gcd(m,p) = p^{l-1}. \end{cases} \tag{6.18}$$

Since l is odd, we may write

$$g(m,p^l) = \sum_{n \bmod p^l} \left(\frac{n}{p}\right) \exp\left(\frac{2\pi i m n}{p^l}\right) = p^{l-1} \sum_{n \bmod p} \left(\frac{n}{p}\right) \exp\left(\frac{2\pi i m/p^{l-1} n}{p}\right),$$

and upon evaluating the last expression (a classical Gauss sum) using Lemma 4.3.2.3, we find that

$$g(m,p^l) = \epsilon_{p^l} p^{l-1+1/2} \left(\frac{m/p^{l-1}}{p}\right),$$

and 6.18 is immediate.

To prove 6.14, we work with the Euler product for $A_m(s)$. Indeed, for convenience, set

$$A_{m,p}(s) = \sum_{l=0}^{\infty} \frac{\epsilon_{p^l} g(-m, p^l)}{p^{l(s+1/2)}}.$$

First, suppose that $p \nmid m$. By 6.16,

$$g(-m, p^l) = \begin{cases} 1 & l = 0, \\ \left(\frac{-m}{p}\right)\epsilon_p p^{l/2} & l = 1 \\ 0 & l \geq 2. \end{cases}$$

and $\epsilon_{p^l}^2 = \left(\frac{-1}{p^l}\right)$. It follows that

$$A_{m,p}(s) = 1 + \left(\frac{m}{p}\right) p^{-s}.$$

Now suppose that p divides m exactly once. Then by 6.16 and 6.17,

$$A_{m,p}(s) = 1 - p^{-2s},$$

so if m is squarefree,

$$\begin{aligned} A_m(s) &= \prod_{\substack{p \neq 2 \\ p \nmid m}} \left(1 + \left(\frac{m}{p}\right) p^{-s}\right) \prod_{\substack{p \neq 2 \\ p \mid m}} \left(1 - p^{-2s}\right) \\ &= \prod_{\substack{p \neq 2 \\ p \nmid m}} \left(1 - \left(\frac{m}{p}\right) p^{-s}\right)^{-1} \prod_{\substack{p \neq 2 \\ p \nmid m}} \left(1 - p^{-2s}\right) \prod_{\substack{p \neq 2 \\ p \mid m}} \left(1 - p^{-2s}\right) \\ &= \prod_{p \neq 2} \left(1 - \left(\frac{m}{p}\right) p^{-s}\right)^{-1} \Big/ \prod_{p \neq 2} \left(1 - p^{-2s}\right)^{-1} \end{aligned}$$

Note that for $p \neq 2$, $\left(\frac{m}{p}\right) = \chi_m(p)$. It follows that

$$\prod_{p \neq 2} \left(1 - \left(\frac{m}{p}\right) p^{-s}\right)^{-1} = L(\chi_m, s)\left(1 - \chi_m(2)2^{-s}\right)$$

and

$$\prod_{p \neq 2} \left(1 - p^{-2s}\right)^{-1} = \zeta(2s)\left(1 - 2^{-2s}\right),$$

so the proof of 6.14 is complete.

The expression 6.15 follows from the next identity, valid for m squarefree, $\mathrm{Re}(s) \geq \frac{1}{2}$ and $\mathrm{Re}(z)$ sufficiently large:

$$\sum_{n=1}^{\infty} A_{mn^2}(s) n^{-z} = A_m(s) \frac{\zeta(z)\zeta(z + 2s - 1)\left(1 - 2^{-z-2s+1}\right)}{L(s + z, \chi_m)\left(1 - \chi_m(2)2^{-s-z}\right)}. \tag{6.19}$$

Indeed, if 6.19 is true, we may expand the right hand side as a Dirichlet series in z and equate coefficients. The Euler product for $\zeta(z + 2s - 1)\left(1 - 2^{-z-2s+1}\right)$ is the Euler product for $\zeta(z + 2s - 1)$ with the factor for $p = 2$ removed, so

$$\zeta(z + 2s - 1)\left(1 - 2^{-z-2s+1}\right) = \sum_{\substack{n=1 \\ (n,2)=1}} n^{1-2s} n^{-z}.$$

Similarly, the Dirichlet series for $L(s + z, \chi_m)\left(1 - \chi_m(2)2^{-s-z}\right)$ is

$$L(s + z, \chi_m)\left(1 - \chi_m(2)2^{-s-z}\right) = \sum_{\substack{n=1 \\ (n,2)=1}} \chi_m(n) n^{-s} n^{-z},$$

and upon using the Möbius inversion formula,

$$\left(L(s+z,\chi_m)\left(1-\chi_m(2)2^{-s-z}\right)\right)^{-1} = \sum_{\substack{n=1\\(n,2)=1}} \chi_m(n)\mu(n)n^{-s}n^{-z}.$$

We employ the Dirichlet convolution formula to write

$$\sum_{n_1=1}^{\infty} n_1^{-z} \sum_{\substack{n_2=1\\(n_2,2)=1}}^{\infty} n_2^{1-2s}n_2^{-z} \sum_{\substack{n_3=1\\(n_3,2)=1}}^{\infty} \mu(n_3)\chi_m(n_3)n_3^{-s}n_3^{-z}$$

as a single Dirichlet series, and 6.15 follows.

Our last task is to prove 6.19. It suffices to prove, for all odd primes p and m such that $p^2 \nmid m$,

$$\sum_{\nu=0}^{\infty} A_{mp^{2\nu}}(s)p^{-\mu z} = A_m(s)\frac{1-\chi_m(p)p^{-s-z}}{(1-p^{-z})\left(1-p^{1-z-2s}\right)} \tag{6.20}$$

and for $p=2$ with no restriction on m,

$$\sum_{\nu=0}^{\infty} A_{m4^\nu}(s)2^{-\mu z} = \frac{A_m(s)}{1-2^{-z}}. \tag{6.21}$$

Indeed, if we set $f_p = (1-\chi_m(p)p^{-s-z})\left(1-p^{-z}\right)^{-1}\left(1-p^{1-z-2s}\right)^{-1}$ for p odd and $f_2 = 1-2^{-z^{-1}}$, and write $n = p_1^{\nu_1}\cdots p_r^{\nu_r}$, then 6.19 is equivalent to the statement that $A_{mn^2}(s)$ is equal to the $A_m(s)$ times the coefficient of n^{-z} in the product $f_{p_1}\cdots f_{p_r}$. We prove this by induction. If $r=1$, then this is true by 6.20 or 6.21. If $r>1$, set $k = p_1^{\nu_1}\cdots p_{r-1}^{\nu_{r-1}}$ and assume the statement for $r-1$. Then if we replace m with mk^2 in 6.20 or 6.21 and use the induction assumption,

$$\begin{aligned} A_{mk^2p_r^{2\nu_r}}(s) &= A_{mk^2}(s)\cdot\ \text{coefficient of } k^{-z} \text{ in } f_{p_1}\cdots f_{p_r},\\ &= A_m(s)\cdot\ \text{coefficient of } p_r^{-z\nu_r} \text{ in } f_{p_r}\cdot\ \text{coefficient of } p_r^{-z\nu_r} \text{ in } f_{p_r},\\ &= A_m(s)\cdot\ \text{coefficient of } n^{-z} \text{ in } f_{p_1}\cdots f_{p_r}. \end{aligned}$$

We will prove 6.21 first. Note that if n is odd, then $g(-4^\nu m,n) = g(-m,n)$. It follows from the definition of $A_m(s)$ that $A_{m4^\nu}(s) = A_m(s)$, so

$$\sum_{\nu=0}^{\infty} A_{m4^\nu}(s)2^{-\mu z} = A_m(s)\sum_{\nu=0}^{\infty} 2^{-\mu z} = \frac{A_m(s)}{1-2^{-z}}.$$

Now we prove 6.20. We may assume that $p^2 \nmid m$, so we split the problem into two cases, depending on whether p divides m or not. The strategy in each case is very direct: we use 6.16 and 6.17 to compute $A_{mp^{2\nu},q}(s)$ for all odd primes q so that we have an expression for $A_{mp^{2\nu}}(s)$, and then the Dirichlet series at 6.19 is easily calculated.

Case 1: $p \mid m$.

Write $m = m'p$. Since m is squarefree, $(m',p)=1$. We begin by calculating $A_{m'p^{2\nu+1},p}(s)$, for which we require nothing more than 6.16 and 6.17.

$$\begin{aligned} A_{m'p^{2\nu+1},p}(s) &= 1 + \sum_{2\le l \text{ even}} \frac{\epsilon_{p^l}c_{p^l}(m'p^{2\nu+1})}{p^{l(s+1/2)}} + \sum_{l \text{ odd}} \frac{\epsilon_{p^l}g(-m'p^{2\nu+1},p^l)}{p^{l(s+1/2)}}\\ &= 1 + \sum_{l=1}^{\nu} \frac{\phi(p^{2l})}{p^2l(s+1/2)} - \frac{p^{2\nu+1}}{p^{(\nu+1)(2s+1)}}\\ &= 1 + \frac{p-1}{p}\sum_{l=1}^{\nu} p^{l(1-2s)} - \frac{1}{p}p^{(\nu+1)(1-2s)}. \end{aligned}$$

Note that the sum over the odd l vanishes since the criterion $\gcd(m'p^{2\nu+1}, p) = p^{l-1}$ is never satisfied for $g(-m'p^{2\nu+1}, p^l)$. The sum occurring in the final line is a geometric series, and so

$$A_{m'p^{2\nu+1},p}(s) = 1 + \frac{p-1}{p}\left(\frac{1-p^{(1-2s)(\nu+1)}}{1-p^{1-2s}} - 1\right) - \frac{1}{p}p^{(1-2s)(\nu+1)}.$$

Next, we calculate $A_{m'p^{2\nu+1},q}(s)$ for primes q dividing m'. As for $q = p$, the sum over odd l vanishes (this time because m' is squarefree), and we use 6.17 to simplify the remaining terms.

$$\begin{aligned} A_{m'p^{2\nu+1},q}(s) &= 1 + \sum_{2\le l \text{ even}} \frac{\epsilon_{q^l} c_{q^l}(m'p^{2\nu+1})}{q^{l(s+1/2)}} + \sum_{l \text{ odd}} \frac{\epsilon_{q^l} g(-m'p^{2\nu+1}, q^l)}{q^{l(s+1/2)}} \\ &= 1 + q^{-2s}. \end{aligned}$$

Lastly, we deal with $A_{m'p^{2\nu+1},q}(s)$ for primes $q \nmid m'p$. This time, the sum over even l vanishes, and only the $l = 1$ term of the sum over odd l can possibly be nonzero:

$$\begin{aligned} A_{m'p^{2\nu+1},q}(s) &= 1 + \sum_{l \text{ odd}} \frac{\epsilon_{q^l} g(-m'p^{2\nu+1}, q^l)}{q^{l(s+1/2)}} \\ &= 1 + \left(\frac{m'p}{q}\right) q^{-s}. \end{aligned}$$

To arrive at an expression for $A_{m'p^{2\nu+1}}(s)$, we take the product over odd primes q of the $A_{m'p^{2\nu+1},q}(s)$. We observe that most of these terms come from $A_m(s)$:

$$\begin{aligned} \prod_{q|m'} A_{m'p^{2\nu+1},q}(s) \prod_{q\nmid m'} A_{m'p^{2\nu+1},q}(s) &= \prod_{q|m'} (1-q^{-2s}) \prod_{q\nmid m'} \left(1 + \left(\frac{m'p}{q}\right) q^{-s}\right) \\ &= \prod_{q|m} (1-q^{-2s}) \prod_{q\nmid m} \left(1 + \left(\frac{m}{q}\right) q^{-s}\right) \Big/ (1-p^{-2s}). \end{aligned}$$

It follows that

$$A_{mp^{2\nu}}(s) = \frac{A_m(s)}{1-p^{-2s}}\left(1 + \frac{p-1}{p}\left(\frac{1-p^{(1-2s)(\nu+1)}}{1-p^{1-2s}} - 1\right) - \frac{1}{p}p^{(1-2s)(\nu+1)}\right),$$

so

$$\begin{aligned} \sum_{\nu=0}^{\infty} A_{mp^{2\nu}}(s) p^{-\mu z} &= \frac{A_m(s)}{1-p^{-2s}}\left[\frac{1}{1-p^{-z}} + \frac{p-1}{p}\left(\frac{1}{1-p^{1-2s}}\frac{1}{1-p^{-z}} - \frac{p^{1-2s}}{1-p^{1-2s}}\frac{1}{1-p^{1-z-2s}} - \frac{1}{1-p^{-z}}\right)\right. \\ &\qquad \left. -p^{-2s}\frac{1}{1-p^{1-z-2s}}\right] \\ &= \frac{1}{(1-p^{-z})(1-p^{1-z-2s})}, \end{aligned}$$

and thus 6.19 is proved for $p \mid m$.

Case 2: $p \nmid m$.

First we compute $A_{mp^{2\nu},p}(s)$. In contrast to the situation in Case 1, the sum over odd l does not disappear.

$$\begin{aligned} A_{mp^{2\nu},p}(s) &= 1 + \sum_{2\le l \text{ even}} \frac{\epsilon_{p^l} c_{p^l}(m'p^{2\nu+1})}{p^{l(s+1/2)}} + \sum_{l \text{ odd}} \frac{\epsilon_{p^l} g(-m'p^{2\nu+1}, p^l)}{p^{l(s+1/2)}} \\ &= 1 + \sum_{l=1}^{\nu} \frac{\phi(p^{2l})}{p^2 l(s+1/2)} + \frac{\epsilon_{p^{2\nu+1}} \epsilon_p p^{2\nu+1/2}\left(\frac{-m}{p}\right)}{p^{(2\nu+1)(s+1/2)}} \\ &= 1 + \frac{p-1}{p}\sum_{l=1}^{\nu} p^{l(1-2s)} + \left(\frac{m}{p}\right) p^{-s} p^{\nu(1-2s)}. \end{aligned}$$

We proceed to the task of evaluating $A_{mp^{2\nu},q}(s)$ for odd $q \mid m$. Since the values of the Ramanujan symbols $c_{q^l}(mp^i)$ are independent of i when $(q,p)=1$, and the sum over odd l vanishes again as m is squarefree, the evaluation is no different to the corresponding result from Case 1:

$$A_{mp^{2\nu},q}(s) = 1 - q^{-2s}.$$

Similarly, if $q \nmid mp^{2\nu}$, then for the evaluation of $A_{mp^{2\nu},q}(s)$, the sum over even l vanishes and only the first term of the sum over the odd l is nonzero:

$$\begin{aligned} A_{m'p^{2\nu},q}(s) &= 1 + \sum_{l \text{ odd}} \frac{\epsilon_{q^l} g(-m'p^{2\nu}, q^l)}{q^{l(s+1/2)}} \\ &= 1 + \left(\frac{m}{q}\right) q^{-s}. \end{aligned}$$

As in Case 1, most of the terms of $A_{mp^{2\nu}}(s)$ are spoken for by $A_m(s)$, but this time we remove a different factor:

$$\begin{aligned} \prod_{q \mid m} A_{mp^{2\nu},q}(s) \prod_{q \nmid mp} A_{m'p^{2\nu},q}(s) &= \prod_{q \mid m} (1 - q^{-2s}) \prod_{q \nmid mp} \left(1 + \left(\frac{m'p}{q}\right) q^{-s}\right) \\ &= \prod_{q \nmid m} (1 - q^{-2s}) \prod_{q \nmid m} \left(1 + \left(\frac{m}{q}\right) q^{-s}\right) \Big/ \left(1 + \left(\frac{m}{p}\right) p^{-s}\right). \end{aligned}$$

It follows that

$$A_{mp^{2\nu}}(s) = \frac{A_m(s)}{1 + \left(\frac{m}{p}\right)p^{-s}} \left(1 + \frac{p-1}{p}\left(\frac{1 - p^{(1-2s)(\nu+1)}}{1 - p^{1-2s}} - 1\right) + \left(\frac{m}{p}\right) p^{-s} p^{\nu(1-2s)}\right),$$

so

$$\begin{aligned} \sum_{\nu=0}^{\infty} A_{mp^{2\nu}}(s) p^{-\mu z} &= \frac{A_m(s)}{1 + \left(\frac{m}{p}\right)p^{-s}} \left[\frac{1}{1 - p^{-z}} + \frac{p-1}{p}\left(\frac{1}{1 - p^{1-2s}}\frac{1}{1 - p^{-z}} - \frac{p^{1-2s}}{1 - p^{1-2s}}\frac{1}{1 - p^{1-z-2s}} - \frac{1}{1 - p^{-z}}\right)\right. \\ &\qquad \left. + p^{-s}\left(\frac{m}{p}\right)\frac{1}{1 - p^{1-z-2s}}\right] \\ &= \frac{1 - \left(\frac{m}{p}\right)p^{-s}}{(1 - p^{-z})(1 - p^{1-z-2s})}, \end{aligned}$$

and thus 6.19 is proved for $p \nmid m$. □

Theorem 6.2.2.4. *Let $\theta(z)$ stand for Jacobi's theta function from Section 2.2. We have*

$$\mathrm{Res}_{s=1/2} E_0(z,s) = \frac{1-i}{2\pi}\theta(z).$$

Proof. For each nonzero m, write $m = m_0 n^2$, where m_0 is not divisible by any odd square. If $m_0 \neq 1$, then χ_{m_0} is non-principal, so $L(s, \chi_{m_0})$ is entire. Therefore $\mathrm{Res}_{s=1} A_m(s)$ vanishes unless m is a perfect square. In this case, $m_0 = 1$ and $m = n^2$, so

$$\mathrm{Res}_{s=1} A_m(s) = \frac{2}{3\zeta(2)} \sum_{\substack{d_2 d_3 \mid n \\ (2, d_2 d_3) = 1}} \mu(d_3)(d_2 d_3)^{-1}. \tag{6.22}$$

The sum at 6.22 is equal to unity:

$$\sum_{\substack{d_1 d_2 d_3 = n \\ (2, d_2 d_3) = 1}} \mu(d_3)(d_2 d_3)^{-1} = \sum_{\substack{d_1 \mid n \\ (n/d_1, 2) = 1}} \frac{d_1}{n} \sum_{d_3 \mid n/d_1} \mu(d_3) = 1,$$

where we used the fact that

$$\sum_{d \mid N} \mu_d = \begin{cases} 1 & N = 1, \\ 0 & N > 1. \end{cases}$$

For $m = 0$, we easily calculate that

$$\mathrm{Res}_{s=1} A_0(s) = \frac{2}{3\zeta(2)} \mathrm{Res}_{s=1} \zeta(2s-1) = \frac{1}{3\zeta(2)}.$$

With the residues of the arithmetic parts of the Fourier coefficients explicated, we now show that the Whittaker functions $K_m(s, y)$ for $m \geq 0$ simplify at $s = 1/2$. Indeed, we prove that

$$K_m(1/2, y) = \sqrt{2}\pi e^{-\pi i/4} y^{-1/2} e^{-2\pi m y}. \tag{6.23}$$

Set

$$f(z) = \frac{e^{-2\pi i m z}}{(z^2+y^2)^s (z+iy)^{1/2}},$$

and note that as a function of the complex variable z, f is holomorphic in the lower half plane, except for a singularity at $z = -iy$, and if m is any positive real number, $f(z)$ is rapidly decreasing as $|z| \to \infty$. By the residue theorem,

$$K_m(s, y) = \int_{\mathbb{R}} f(z) = 2\pi i \ \mathrm{Res}_{z=-iy} f(z) = 2\pi i \frac{e^{-2\pi m y}}{(-2iy)^s} \ \mathrm{Res}_{z=-iy}(x+iy)^{-(s+1/2)},$$

and the residue is easily evaluated after setting $s = \frac{1}{2}$, which proves 6.23 for $m \neq 0$. The case $m = 0$ is then immediate upon taking the limit.

It follows that

$$\begin{aligned} \mathrm{Res}_{s=1/2} E_0(z, s) &= \frac{1}{2}\left(\frac{y}{4}\right)^{\frac{1}{2}} \sum_{m=-\infty}^{\infty} \mathrm{Res}_{s=1}\left(A_m(s)\right) K_m\left(\frac{1}{2}, y\right) \\ &= \frac{\sqrt{2}\pi e^{-\pi i/4}}{4}\left(\frac{1}{3\zeta(2)} + \frac{2}{3\zeta(2)} \sum_{m=1}^{\infty} e^{2\pi i m^2 (x+iy)}\right) \\ &= \frac{1-i}{2\pi} \sum_{m=-\infty}^{\infty} e^{2\pi i m^2 z}. \end{aligned}$$

□

6.2.3 The asymptotic expansion of $E_0(z, s)$

In this subsection, we explain how Proposition 6.2.2.3, in which we determined the Fourier coefficients of $E_0(z, s)$, may be used to calculate the asymptotic expansion of $E_0(z, s)$ towards the cusps. We do not give full details, as it is will be clear that the results obtained are no more interesting than the Landsberg–Schaar relation, preferring to truncate our calculations when it becomes clear that the expressions involved do simplify.

We make liberal use of the results appearing in the work of Goldfeld and Hoffstein [GH85] mentioned in the last subsection. For simplicity, we will only consider asymptotic expansions towards cusps with prime denominator; that is, we set $z = b/p + i\epsilon$, where p is prime, and let ϵ tend to zero.

We begin, as in Subsection 6.1.2, by isolating the variable in which we wish to compute the asymptotic expansion:

$$\begin{aligned} E_0\left(\frac{b}{p} + i\epsilon, s\right) &= a_0(s, i\epsilon) + \sum_{n=0}^{p-1} \exp\left(\frac{2\pi i n b}{p}\right) \sum_{\substack{m=-\infty \\ m \neq 0 \\ m = n \bmod p}}^{\infty} a_m(s, \epsilon) \\ &= \sqrt{2\pi} e^{\pi i/4} \epsilon^{1/2} (4\epsilon)^{-s} \frac{\zeta(4s-1)(1-2^{-4s+1})\Gamma(2s-1/2)}{\zeta(4s)(1-2^{-4s})\Gamma(2s)} \\ &+ \left(\frac{\epsilon}{4}\right)^s \left(\frac{1}{\phi(p)} \sum_{n=1}^{p-1} \exp\left(\frac{2\pi i n b}{p}\right) \sum_{|\chi|=p} \overline{\chi(n)} \sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \chi(m) A_m(2s) K_m(s, \epsilon) + \sum_{\substack{m=-\infty \\ m \neq 0}}^{\infty} A_{pm}(2s) K_{pm}(s, \epsilon)\right). \end{aligned} \tag{6.24}$$

By Theorem 6.1.1.1, in order to investigate the asymptotic expansion of $E_0\left(\frac{b}{p} + i\epsilon, s\right)$ as $\epsilon \to 0$, we need to calculate the Mellin transform of non-constant terms in the expression above. This comes down to evaluating

$$\int_0^{\infty} K_m(s, t) t^{z+s-1} dt.$$

Fortunately, Goldfeld and Hoffstein [GH85, Proposition 2.1] have already done this for us.

Proposition 6.2.3.1. *Suppose $m \neq 0$. Then*

$$\int_0^\infty K_m(s,t)t^{z+s-1}dt = \frac{e^{\pi i/4}F_\pm(z,s)}{2^{2z-\frac{1}{2}}\pi^{w-s-\frac{1}{2}}}|m|^{-(z-s+\frac{1}{2})},$$

where

$$F_\pm(z,s) = \Gamma\left(z-s+\frac{1}{2}\right)\Gamma(z+s)\times\begin{cases}\dfrac{{}_2F_1\left(z+s,z-s+\frac{1}{2},z+\frac{1}{2};\frac{1}{2}\right)}{\Gamma\left(s+\frac{1}{2}\right)\Gamma\left(z+\frac{1}{2}\right)} & m>0\\ \dfrac{{}_2F_1\left(z+s,z-s+\frac{1}{2},z+1;\frac{1}{2}\right)}{\Gamma(s)\Gamma(z+1)} & m<0.\end{cases}$$

According to Theorem 6.1.1.1, we need to find the poles and residues of the following Dirichlet series, considered as functions of z:

$$\sum_{m=1}^\infty \chi(m)A_m(2s)m^{s-z-\frac{1}{2}}, \tag{6.25}$$

$$\sum_{m=1}^\infty \chi(m)A_{-m}(2s)m^{s-z-\frac{1}{2}}, \tag{6.26}$$

$$\sum_{m=1}^\infty A_{pm}(2s)(am)^{s-z-\frac{1}{2}}, \tag{6.27}$$

$$\sum_{m=1}^\infty A_{-pm}(2s)(am)^{s-z-\frac{1}{2}}. \tag{6.28}$$

The Dirichlet series at 6.27 and 6.28 may meromorphically continued into the whole z-plane. We set $\rho = 2s$ and $s-z-1/2 = 1-\delta$. For $\rho \neq 1/2$, the only singularities are simple poles at $\delta = 0, 1/2-\rho$ and -1. The residue at $\delta = -1$ is independent of p, and the residues at the other two singularities may be inferred from work of Goldfeld and Hoffstein [GH85, Proposition 4.1].

For $\rho = 1/2$, the singularities consist of a double pole at $\delta = 0$ and a simple pole at $\delta = 1$. The residues are similarly complicated and we refer the reader to Goldfeld and Hoffstein [GH85, Proposition 4.1].

Since the Dirichlet series 6.27 and 6.28 are independent of χ (comprising the constant term $n = 0$), we need not pursue them further.

The Dirichlet series at 6.25 and 6.26 depend on χ and so we should determine their residues. By Theorem 6.1.1.1, these Dirichlet series will only have poles in the case that χ is *effectively principal*: that is, $\chi(m) = 1$ whenever $A_m(2s) \neq 0$ and m is coprime to the (prime) modulus of χ. By 6.15, if we write $m = m_0n^2$, where m_0 is squarefree, then

$$A_m(s) = A_{m_0}(s)\sum_{\substack{d_2d_3|n\\(2,d_2d_3)=1}}\chi_{m_0}(d_3)\mu(d_3)d_2^{1-2s}d_3^{-s}.$$

We tacitly assume that $A_m(s)$ is nonzero if and only if $A_{m_0}(s) \neq 0$. For m_0 squarefree,

$$A_{m_0}(s) = \frac{L(s,\chi_{m_0})}{\zeta(2s)}\frac{(1-\chi_{m_0}(2)2^{-s})}{(1-2^{-2s})} = \frac{L\left(s,\left(\frac{\cdot}{m_0}\right)\right)}{\zeta(2s)}\frac{\left(1-\left(\frac{2}{m_0}\right)2^{-s}\right)}{(1-2^{-2s})},$$

where we have used that fact that $L(s,\chi_{m_0})$ agrees with $L\left(s,\left(\frac{\cdot}{m_0}\right)\right)$ away from the Euler factor at $p=2$.

For the rest of this subsection, we will only deal with the Dirichlet series 6.25: the calculations for 6.26 are similar.

Recall that, for *fixed* χ, we characterised the poles and zeros of the Dirichlet L-function $L(s,\chi)$ in Proposition 1.3.4.2, up to the generalised Riemann hypothesis (GRH). To recapitulate, if χ is primitive, the classification is as follows (recall that the principal character modulo m is primitive only when $m = 1$).

Zeros: Trivial zeros occur at $s = 0, -2, -4, -6, \ldots$ if χ is even and non-principal; at $s = -2, -4, -6, \ldots$ if $\chi = 1$, and at $s = -1, -3, -5, -7, \ldots$ if χ is odd. By the functional equation they are all simple. Nontrivial zeros occur for $0 \leq \mathrm{Re}(s) < 1$, and the GRH is the assertion that any nontrivial zero s has $\mathrm{Re}(s) = \frac{1}{2}$.

Poles: $L(s,\chi)$ is entire unless χ is the trivial character, in which case there is a simple pole at $s=1$ with residue 1.

We now analyse the zeros and poles of $A_{m_0}(\rho)$ (we remind the reader that m_0 is squarefree). We assume the GRH in order to ensure that the six cases are mutually exclusive.

1: $\rho_0 = 1$. There is a simple pole at ρ_0 if $m_0 = 1$; otherwise, $A_{m_0}(\rho)$ is holomorphic and nonzero in a neighbourhood of ρ_0.

Since $A_m(\rho)$ has a pole near ρ_0 for infinitely many m (the perfect squares), we attach the asymptotic expansion to $\mathrm{Res}_{s=1/2}E_0(z,s)$. Then the effectively trivial characters for $\mathrm{Res}_{\rho=1}\sum_{m=1}^{\infty}A_m(\rho)m^{-\rho}$ are those which are trivial on the perfect squares, using the results of Subsection 6.1.2, we recover the Landsberg–Schaar relation.

2: $\rho_0 = 0$. There is a simple zero if $\left(\frac{2}{m_0}\right) = 1$ and $m_0 \neq 1$, and $A_{m_0}(\rho)$ is holomorphic and nonzero in a neighbourhood of ρ_0 if $\left(\frac{2}{m_0}\right) = -1$ and $m_0 \neq 1$ or $\left(\frac{2}{m_0}\right) = 1$ and $m_0 = 1$.

The relevant Dirichlet series is

$$\sum_{m=1}^{\infty}\mathrm{Res}_{s=\rho_0/2}A_m(2s)m^{-w},$$

and the effectively principal characters are those that are trivial on p-smooth[12] m such that the odd squarefree part m_0 satisfies $\left(\frac{2}{m_0}\right) = -1$ or $m_0 = 1$. The latter implies that the character is quadratic; the former implies that the character is unity on all m_0 congruent to 3 or 5 modulo 8. By Dirichlet's theorem on primes in arithmetic progression the only such character is the principal character modulo p.

3: $\rho_0 = -1, -3, -5, \ldots$. In this case, $A_{m_0}(\rho)$ has a simple zero at ρ_0 if $\left(\frac{\cdot}{m_0}\right)$ is odd and is holomorphic and nonzero near the indicated points otherwise.

In other words, $A_m(\rho)$ is nonzero and only if the squarefree odd part of m is congruent to 1 modulo 4. Any effectively trivial character for $\sum_{m=1}^{\infty}A_m(\rho_0)m^{w}$ must be trivial, in particular, on the primes congruent to 1 modulo 4 (and unequal to p), and on all even numbers. It follows that only the principal character modulo p is effectively trivial.

4: $\rho_0 = -2, -4, -6, \ldots$. In this case, $A_{m_0}(\rho)$ has a simple pole at ρ_0 if $\left(\frac{\cdot}{m_0}\right)$ is odd and is holomorphic and nonzero near the indicated points otherwise.

The Dirichlet series we must contemplate is

$$\sum_{m=1}^{\infty}\mathrm{Res}_{s=\rho_0/2}A_m(2s)m^{-w}.$$

We seek characters modulo p which are trivial on p-smooth m such that the odd squarefree part of m is congruent to 3 modulo 4. In particular, such characters would have to be trivial on all primes congruent to 3 modulo 4 (and unequal to p) and on all even numbers, and must therefore be principal.

5: ρ_0 is a nontrivial zero of $L\left(\rho, \left(\frac{\cdot}{n_0}\right)\right)$ for some fixed n_0. Then $A_{m_0}(\rho)$ has a simple zero for $m_0 = n_0$ near ρ_0, and, if we assume the Linear Independence Conjecture [Win38, Chapter 13], is nonzero for all other values of m_0.

A character is effectively trivial for $\sum_{m=1}^{\infty}A_m(\rho_0)m^{w}$ if it is trivial on all p-smooth numbers with squarefree part not equal to n_0. Any such character must in fact be trivial; however, note that if we instead focus on

$$\mathrm{Res}_{s=\frac{\rho_0}{2}}\left(\frac{E_0(z,s)}{s-\frac{\rho_0}{2}}\right) - E_0\left(z,\frac{\rho_0}{2}\right),$$

so that the Dirichlet series in question is

$$\sum_{m=1}^{\infty}\left(\mathrm{Res}_{s=\rho_0/2}\left(\frac{1}{s-\rho_0}A_m(2s)m^{-s}\right) - A_m(\rho_0)\right)m^{w},$$

then the effectively trivial characters are those trivial on the p-smooth numbers with squarefree part *equal* to n_0. Any such character must be quadratic; for each p, the sum over characters modulo p at 6.24 therefore

[12] An integer n is p-smooth if it is not divisible by p. We temporarily use this terminology to ease clutter.

consists of either just the trivial character, in the case that $\left(\frac{n_0}{p}\right) = -1$, or both the trivial character and $\left(\frac{\cdot}{p}\right)$ if $\left(\frac{n_0}{p}\right) = 1$. Though we possess not a shred of evidence, we tentatively conjecture that, if the sums arising in this way are nontrivial, we recover the twisted Landsberg–Schaar relation with character $\left(\frac{\cdot}{n_0}\right)$.

6: ρ is a nontrivial zero of $\zeta(2\rho)$. Then $A_{m_0}(\rho)$ has a simple pole for all m_0.

The relevant Dirichlet series is

$$\sum_{m=1}^{\infty} \operatorname{Res}_{s=\rho_0/2} A_m(2s) m^{-w},$$

and $\operatorname{Res}_{s=\rho_0/2} A_m(2s)$ is never zero, so the only effectively trivial character is the principal character.

If $2s = \rho$ does not belong to any of the six classes listed above, then $A_{m_0}(s)$ has no poles or zeros for any squarefree m_0, and so we may reasonably expect that the asymptotic expansion of $E_0(z, s)$ is determined by the residues of the "constant term", corresponding to the second sum at 6.24.

To summarise, the most interesting (and least involved) asymptotic expansion is the one attached to the residue at $s = 1/2$; this gives rise to the Landsberg–Schaar relation. The only other special values of $E_0(z, s)$ which give potentially interesting asymptotic expansions are the s for which one of the Dirichlet L-functions has a nontrivial zero, but even in this case, we observe that only quadratic characters may appear in the character sum at 6.24. Indeed, the most important point of the discussion of this section is that the relation, valid for m_0 squarefree:

$$A_{n^2 m_0}(s) = A_{m_0}(s) \sum_{\substack{d_2 d_3 | n \\ (2, d_2 d_3) = 1}} \chi_{m_0}(d_3) \mu(d_3) d_2^{1-2s} d_3^{-s}. \tag{6.29}$$

serves to exclude characters of order greater than two from having any influence. In the next section, we will see that the analogue of 6.29 appears, as expected, when we sketch the theory of metaplectic Eisenstein series over $\mathbb{Q}(i)$, and remarkably, a version of 6.29 holds for a certain *cubic* theta function.

6.3 Metaplectic covers of $GL(n)$ and theta functions of higher degree

In this chapter, we outline the theory of higher theta functions, with particular emphasis on the analogies with the familiar quadratic theta functions.

The first construction and investigation of theta functions of higher degree is the book of Kubota [Kub69], in which he observes that Weil's construction of theta functions using quadratic residue symbols may be generalised to produce new "theta functions" from the Hilbert symbol over any *totally imaginary* field (containing the nth roots of unity). The key to his argument is that if K is a totally imaginary number field containing the nth roots of unity and $\Gamma_1(\lambda)$ is a principal congruence subgroup of $SL(2, \mathcal{O}_K)$ modulo some ideal λ (related to the different of K), then the reciprocity law for the Hilbert symbol $\left(\frac{c}{d}\right)_n$ is equivalent to the statement that the map

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{cases} \left(\frac{c}{d}\right)_n & c \neq 0, \\ 1 & c = 0, \end{cases}$$

is a homomorphism from $\Gamma_1(q)$ into the group of nth roots of unity in K. Kubota uses this observation to construct Eisenstein series on metaplectic covers of $SL(2, \mathcal{O}_K)$, and posits that the residues in s of these Eisenstein series are appropriate candidates for higher theta functions.

The details of the construction quickly become cumbersome; in Subsection 6.3.2 we outline the process for the quadratic residue symbol over $\mathbb{Q}(i)$. It will be clear by the end of our sketch that we recover, as a residue, the familiar theta function over $\mathbb{Q}(i)$, and that the shape of the Fourier coefficients of the Eisenstein series already indicate that we may prove Hecke reciprocity by studying the asymptotic expansion at the cusps.

One might think, at this point, that a completely analogous relation exists between Gauss sums associated to residue symbols and Kubota's theta functions. However, this is not the case. The determination of the Fourier coefficients of Kubota's theta functions is quite difficult, and Patterson's determination of the coefficients of the cubic theta function [Pat77a; Pat77b] was considered a breakthrough at the time.

On the other hand, it was through a careful analysis of Kubota's cubic theta functions that Heath-Brown and Patterson [HBP79] were finally able to disprove a tentative proposal of Kummer [Kum42] (irresponsibly referred to as Kummer conjecture during the intervening years) on the distribution of the arguments of rational cubic Gauss sums. They note that their proof may be extended, using Eisenstein series on metaplectic covers of degree $r \geq 4$, to prove an equidistribution statement for the arguments of Gauss sums formed with characters of arbitrary order.

It was then only a matter of time before Kubota's construction of metaplectic Eisenstein series was extended to $GL(n)$. Needless to say, the theory becomes ever more complicated as n increases. The extension of Kubota's theory to $GL(n)$ is due to Kazhdan and Patterson [KP84] in 1984, and shortly thereafter Proskurin [Pro85a; Pro85b] investigated the Eisenstein series attached to metaplectic covers of degree 3 on $GL(3)$. In 1986 Bump and Hoffstein [BH86] investigated the theta functions obtained as residues of such Eisenstein series, finding that they are a far better analogue of the classical quadratic theta functions than the cubic theta functions of Kubota, attached to $GL(2)$.

The superiority of cubic theta functions on $GL(3)$ over those on $GL(2)$ may be explained by a result of Deligne [Del80]: if we denote by $\theta_{n,r}$ the functions arising as a residue of the Eisenstein series on the metaplectic cover of $GL(n)$ formed with the degree r residue symbol[13], then $\theta_{n,r}$ has a unique *Whittaker model*[14] only if $n = r$ or $n = r - 1$ [BB05, pp. 117]. As a consequence, the Fourier coefficients of $\theta_{n,r}$ for $r > n + 1$ remain mysterious, even for $n = 2$; whereas we expect much stronger results when $n = r$.

This brings us to the main point of our discussion: we tentatively suggest that one may attempt to fill in the gap observed by Hecke by computing asymptotic expansions of the higher theta functions $\theta_{n,n}$, as these are the closest higher-degree analogues yet discovered of the classical theta functions. In Subsection 6.3.1, we provide an extremely lean overview of the theory of automorphic forms on $GL(n)$, our only aim being to establish the minimal terminology necessary to discuss higher theta functions. In Subsection 6.3.2, we outline the computation of the Fourier coefficients of the Kubota theta function on $\mathbb{Q}(i)$, where we obtain a classical theta function, as expected, followed by an overview of the most important features of the cubic theta function on $GL(3)$ investigated by Bump and Hoffstein. In the final subsection, we describe the discovery of Banks, Bump and Lieman [BBL03] that Fourier coefficients of $\theta_{n,n}$ may be expressed in terms of L-series formed with n-th power residue symbols, which generalises the $GL(3)$ result of Bump and Hoffstein [BH86].

6.3.1 Automorphic forms on $GL(n)$

In this subsection, we state the bare essentials of the theory of automorphic forms necessary to understand how one writes down higher theta functions on $GL(n)$ in terms of their Fourier coefficients. Our exposition follows the treatment of Goldfeld [Gol06].

Recall that, in Subsection 1.1.1, we briefly made us of the fact that, as a manifold, the upper half plane $\mathscr{H}$ may be expressed as

$$\mathscr{H} = SL(2,\mathbb{R})/O(2,\mathbb{R}),$$

where $O(2,\mathbb{R})$ is the stabiliser of i in $SL(2,\mathbb{R})$. Alternatively, we may write

$$\mathscr{H} = GL(2,\mathbb{R})/Z(2)O(2,\mathbb{R}),$$

where $Z(2)$ is the group of all nonzero multiples of the identity matrix. This is the *Iwasawa decomposition* for $GL(2,\mathbb{R})$, and it is the starting point for the generalisation of the theory of modular forms to $GL(n,\mathbb{R})$. Indeed, the Iwasawa decomposition for $GL(n)$ is the following statement:

Proposition 6.3.1.1. *Let $n \geq 2$. Then every $g \in GL(2,\mathbb{R})$ may be uniquely expressed as*

$$g = x \cdot k \cdot z,$$

where $z \in O(n,\mathbb{R})$, k is an element of the group of nonzero multiples of the $n \times n$ identity matrix (this group is often called the group of scalar matrices*) and x belongs to the following set:*

$$\mathscr{H}_n = \left\{ u \cdot v \text{ such that } u = \begin{pmatrix} 1 & x_{1,2} & x_{1,3} & \cdots & x_{1,n} \\ & 1 & x_{2,3} & & \\ & & \ddots & & \\ & & & 1 & x_{n-1,n} \\ & & & & 1 \end{pmatrix}, \begin{pmatrix} y_1 \cdots y_n & & & & \\ & y_1 \cdots y_{n-2} & & & \\ & & \ddots & & \\ & & & y_1 & \\ & & & & 1 \end{pmatrix}, \right\}$$

[13] We ignore the fact that there are *several* Eisenstein series that may be constructed on metaplectic covers.
[14] Automorphic-form parlance for Fourier coefficients.

where the $x_{i,j}$ are real and $y_i > 0$. Note that we use blank spaces to denote zero *entries in the matrices, not arbitrary entries.*

The set $\mathscr{H}_n$ is called the *generalised upper half plane*, and plays the same role in the theory of automorphic forms on $GL(n,\mathbb{R})$ as $\mathscr{H}$ does for the modular forms. Proposition 6.3.1.1 implies that $\mathscr{H}_n$ is a homogeneous space for $GL(n,\mathbb{R})$:

$$\mathscr{H}_n = GL(n,\mathbb{R})/Z(n)O(n,\mathbb{R}),$$

where $Z(n)$ now stands for the group of scalar matrices.

Remark 6.3.1.2. *The reader may easily check that $\mathscr{H}_3$ is a five-dimensional real manifold, and therefore it does not admit a complex structure. This is typical for $\mathscr{H}_n$: in place of holomorphy, one insists that automorphic forms are eigenfunctions of certain linear elliptic differential operators determined by $GL(n,\mathbb{R})$.*

As for $n = 2$, we wish to consider quotients of $\mathscr{H}_n$ by discrete subgroups. It is not so trivial to show that the subgroup $GL(n,\mathbb{Z})$ does in fact act discretely on $GL(n,\mathbb{R})$ for all n: we refer the reader to Goldfeld [Gol06, Chapter 1, Section 3] for a proof.

Next, we outline the analogue of holomorphy for $\mathscr{H}_n$, which we will require automorphic forms to satisfy. Recall that given any Lie group G, one may form its *Lie algebra* $\mathfrak{g}$, defined as the tangent space to G at the identity element, or, equivalently, the vector space of left-invariant vector fields on G. A Lie group comes equipped with a Lie bracket, which is an alternating bilinear map $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ satisfying the Jacobi identity:

$$[X,[Y,Z]] + [Z,[X,Y]] + [Y,[Z,X]] = 0$$

for all $X, Y, Z \in \mathfrak{g}$. The Lie algebra $\mathfrak{gl}(n,\mathbb{R})$ of $GL(n,\mathbb{R})$ is identified as the vector space of all $n \times n$ real matrices together with the *bracket* operation given by

$$[X,Y] = XY - YX, \text{ for } X, Y \in \mathfrak{gl}(n,\mathbb{R}).$$

By the definition of the Lie algebra, each $h \in \mathfrak{gl}(n,\mathbb{R})$ gives rise to a differential operator on G:

$$(D_h f)(g) = \frac{d}{dt} f(g + t(gh))\Big|_{t=0}, \tag{6.30}$$

where f is any smooth complex-valued function on $GL(n,\mathbb{R})$. One may easily check that, for $h \in \mathfrak{gl}(n,\mathbb{R})$, $g \in GL(n,\mathbb{R})$ and f_1, f_2 smooth functions on $GL(n,\mathbb{R})$, every differential operator D arising by 6.30 satisfies

$$D_h(f_1 f_2)(g) = D_h(f_1)(f_2(g)) + f_1(g)D_h(f_2)(g)$$
$$D_h(f_1 \circ f_2)(g) = D_h(f_1)(f_2(g)) \cdot D_h(f_2)(g),$$

and for $h_1, h_2 \in \mathfrak{gl}(n,\mathbb{R})$, we have the equalities of operators

$$D_{h_1+h_2} = D_{h_1} + D_{h_2}$$
$$D_{h_1} \circ D_{h_2} - D_{h_2} \circ D_{h_1} = D_{[h_1,h_2]},$$

where composition is defined by noting that a (left-invariant) differential operator on a Lie group G may be canonically realised as a function on G. To ease the notation, we write composition as multiplication of operators, so that $D^2 = D \circ D$. The relations given above imply that the set of all such differential operators generates an associative algebra $\mathcal{D}^n$ over $\mathbb{R}$, with multiplication defined by composition.

Given any Lie algebra $\mathfrak{g}$, one may form the *universal enveloping algebra* $\mathfrak{A}(\mathfrak{g})$:

$$\mathfrak{A}(\mathfrak{g}) = \left(\bigoplus_{k=0}^{\infty} \bigotimes^k \mathfrak{g} \right) / \langle X \otimes Y - Y \otimes X - [X,Y] \rangle$$

The universal enveloping algebra of a Lie group is isomorphic to the algebra $\mathcal{D}^n$ of *left-invariant* differential operators on G.

We now consider the centre $\mathscr{D}^n$ of $\mathcal{D}^n$ (or equivalently, the centre of the universal enveloping algebra of $\mathfrak{gl}(n,\mathbb{R})$), which consists of all differential operators $D \in \mathcal{D}^n$ such that $D \circ D' = D' \circ D$ for every $D' \in \mathcal{D}^n$. The import of focusing attention on the centre is that differential operators on the centre are well defined on quotients of the generalised upper half plane by congruence subgroups:

Proposition 6.3.1.3. *Let $D \in \mathscr{D}^n$. Then D is well defined on smooth complex-valued functions on $GL(n,\mathbb{Z})\backslash\mathscr{H}_n$.*

The center $\mathscr{D}^n$ of the universal enveloping algebra contains some distinguished elements which we will use to define analogues of the hyperbolic Laplacian. These are the *Casimir elements*, and they are constructed as follows:

Let $E_{(i,j)}$ denote the matrix in $\mathfrak{gl}(n,\mathbb{R})$ with a 1 in the (i,j) position and zeros elsewhere, and set $D_{(i,j)} = D_{E_{(i,j)}}$. Then for $2 \le m \le n$, the Casimir elements are defined to be the differential operators

$$\sum_{i_1,i_2,\ldots,i_m=1}^{n} D_{(i_1,i_2} \circ D_{(i_2,i_3)} \circ \cdots \circ D_{(i_m,i_1)}.$$

One checks that the Casimir elements (for each m) are indeed in $\mathscr{D}^n$. By a result of Capelli [Cap90], $\mathscr{D}^n$ is a polynomial algebra of rank $n-1$, and every differential operator on $\mathscr{D}^n$ may be expressed as a polynomial in the Casimir elements.

A direct computation reveals that the Casimir element (and this time there is only one) of $\mathfrak{gl}(2,\mathbb{R})$ is

$$2y^2\left(\frac{\partial}{\partial x^2} + \frac{\partial}{\partial y^2}\right). \tag{6.31}$$

Recall that, in Subsection 2.1.2 and Section 6.2, we defined Eisenstein series by averaging over the action of a congruence subgroup on the function $\mathrm{Im}(z)^{-s}$. The analogue of the imaginary-part function for automorphic forms a complex-valued function on $\mathscr{H}_n$, denoted I_s, is defined by

$$I_s(x) = \prod_{i=1}^{n}\prod_{j=1}^{n} y_i^{b_{(i,j)}s_j},$$

where $s = (s_1, \ldots, s_{n-1}) \in \mathbb{C}^{n-1}$, and $b_{(i,j)}$ is equal to ij if $i + j \le n$ and $(n-i)(n-j)$ if $i + j \ge n$. For each $s \in \mathbb{C}^{n-1}$, $I_s(x)$ satisfies

$$I_s(x \cdot k \cdot z) = I_s(x)$$

for every $x \in \mathscr{H}_n$, $k \in O(n,\mathbb{R})$ and $z \in Z(n)$. Then $I_s(x)$ is an eigenfunction of the Casimir operators, with eigenvalue given by

$$D^k_{(i,j)} I_s(x) = \begin{cases} s^k_{n-i} I_s(x) & i = j \\ 0 & \text{otherwise,} \end{cases}$$

for $n \ge 2$, $1 \le i, j \le n$ and $k \ge 1$. It follows from Capelli's result that $I_s(x)$ is an eigenfunction of every $D \in \mathscr{D}^n$, and the function $\lambda : D \to \mathbb{C}$ which associates to each D the eigenvalue $\lambda(D)$ of $I_s(x)$ under D is a character of $\mathscr{H}_n$, called the *Harish-Chandra character.*

We finally define automorphic forms, albeit only those of a special type.

Definition 6.3.1.4. *Let $n \ge 2$ and $\nu = (\nu_1, \ldots \nu_{n-1}) \in \mathbb{C}^{n-1}$. Then a* Maaß form of type ν for $SL(n,\mathbb{Z})$ *is a smooth function f in $L^2(SL(n,\mathbb{Z})\backslash\mathscr{H}_n)$ which satisfies:*

1. *$f(\gamma x) = f(x)$ for all $\gamma \in SL(n,\mathbb{Z})$ and $x \in \mathscr{H}_n$,*
2. *$Df(x) = \lambda_D f(x)$ for all $D \in \mathscr{D}^n$, where λ is the Harish-Chandra character,*
3. *$\int_{(SL(n,\mathbb{Z})\cap U)\backslash U} f(ux)du = 0$ for all groups U consisting of upper triangular elements of the form*

$$\begin{pmatrix} I_{r_1} & & * \\ & \ddots & \\ & & I_{r_s} \end{pmatrix},$$

where $r_1 + \cdots + r_s = n$ and the asterisk denotes arbitrary real entries.

The second condition is the analogue of holomorphy for modular forms (by 6.31 the zero eigenspace of the Casimir element for $\mathfrak{gl}(2,\mathbb{R})$ is the space of harmonic functions) and the third condition is the analogue of cuspidality: there exist interesting automorphic forms violating this condition, just as there exist interesting modular forms which are not cusp forms.

We now turn to our final task: the problem of describing, in as little detail as necessary, the Fourier coefficients of Maaß forms.

For $n \geq 2$, we define

$$U_n(\mathbb{R}) = \left\{ u = \begin{pmatrix} 1 & u_{1,2} & u_{1,3} & \dots & u_{1,n} \\ & 1 & u_{2,3} & \dots & u_{2,n} \\ & & \ddots & & \vdots \\ & & & 1 & u_{n-1,n} \\ & & & & 1 \end{pmatrix} \text{ such that } u_{i,j} \in \mathbb{R} \right\},$$

and by $U_n(\mathbb{Z})$ we mean the subgroup of $U_n(\mathbb{R})$ with $u_{i,j} \in \mathbb{Z}$. One may check that for $m = (m_1, \dots, m_{n-1}) \in \mathbb{Z}^{n-1}$, the function $\psi_m : U_n(\mathbb{R}) \to \mathbb{C}^\times$ defined by

$$\psi_m(u) = e^{2\pi i(m_1 u_{1,2} + m_2 u_{2,3} + \dots + m_{n-1} u_{n-1,n})}$$

is a character, and in fact all characters of $U_n(\mathbb{R})$ are of the form ψ_m for some such m. We may now define the special functions that will appear instead of exponentials in our Fourier expansion.

Definition 6.3.1.5. *Let $n \geq 2$ and $\nu = (\nu_1, \dots, \nu_{n-1}) \in \mathbb{C}^{n-1}$, with other notation as in Definition 6.3.1.4. An $SL(n,\mathbb{Z})$* Whittaker function *of type ν associated to a character ψ_m of $U_n(\mathbb{R})$ is a smooth function $\Phi_m^\nu : \mathscr{H}_n \to \mathbb{C}$ such that*

1. *$\Phi_m^\nu(u \cdot z) = \psi_m(u)\Phi_m^\nu(z)$ for all $u \in U_n(\mathbb{R}), z \in \mathscr{H}_n$,*
2. *$D\Phi_m^\nu(z) = \lambda_D(z)$ for all $D \in \mathscr{D}^n, z \in \mathscr{H}_n$,*
3. *$\int_S |\Phi_m^\nu|^2 dz < \infty$,*

where S is a substitute for a fundamental domain, known as a Siegel set [Gol06, Definition 1.3.1] and dz is the unique-up-to-scale left-invariant $GL(n,\mathbb{R})$-measure on $\mathscr{H}_n$.

The following explicit example of a Whittaker function is due to Jacquet [Jac67]:

$$W_m^\nu(z) = \int_{U_n(\mathbb{R})} I_\nu \left(\begin{pmatrix} & & & (-1)^{\lfloor \frac{n}{2} \rfloor} \\ & & 1 & \\ & 1 & & \\ & \iddots & & \\ 1 & & & \end{pmatrix} \cdot u \cdot z \right) \overline{\phi_m(u)} du,$$

where du is the normalised left-invariant $GL(n,\mathbb{R})$-measure on $U_n(\mathbb{R})$. Jacquet's Whittaker function also satisfies the following growth condition: for sufficiently large $s_1, \dots s_{n-1}$, the integral

$$\int_0^\infty \cdots \int_0^\infty |\Phi_m^\nu(y)| \, y_1^{s_1} \dots y_{n-1}^{s_{n-1}} \frac{dy_1}{y_1} \cdots \frac{dy_{n-1}}{y_{n-1}} \tag{6.32}$$

must converge. A result of Shalika [Sha74] states that the dimension of the vector space of functions with the properties (1), (2) and (3) from Definition 6.3.1.5 which also satisfy 6.32 is exactly one, so any Whittaker function is a constant multiple of W_m^ν.

Now observe that if f is a Maaß form of type $\nu = (\nu_1, \dots \nu_{n-1})$ for $SL(n,\mathbb{Z})$ then the mth Fourier coefficient, initially defined by

$$\bar{f}_m(x) = \int_0^1 \cdots \int_0^1 f(ux)\overline{\psi_m(u)} \prod_{1 \leq i < j \leq n} du_{i,j},$$

where $m = (m_1, \dots, m_{n-1}) \in \mathbb{Z}^{n-1}$, is a Whittaker function according to Definition 6.3.1.5, and it satisfies the growth condition 6.32. We can therefore write

$$\bar{f}_m(x) = \frac{a_{m_1,\dots,m_{n-1}}}{\prod_{k=1}^{n-1} |m_k|^{(k(n-k)/2}} W \left(\begin{pmatrix} m_1 \cdot \dots \cdot m_{n-2}|m_{n-1}| & & & \\ & \ddots & & \\ & & m_1 m_2 & \\ & & & m_1 \end{pmatrix} x \right)$$

for some array $(a_{m_1,\ldots,m_{n-1}})$ of complex numbers. From this, one may associate Fourier expansions to automorphic forms.

Proposition 6.3.1.6. *Let $n \geq 2$ and suppose that f is a Maaß form of type $\nu = (\nu_1, \ldots \nu_{n-1})$ for $SL(n, \mathbb{Z})$. Then the Fourier expansion of f is*

$$f(x) = \sum_{\gamma \in U_{n-1}(\mathbb{Z})\backslash SL(n-1,\mathbb{Z})} \sum_{m_1=1}^{\infty} \cdots \sum_{m_{n-2}=1}^{\infty} \sum_{m_{n-1}\neq 0} \frac{a_{m_1,\ldots,m_{n-1}}(f)}{\prod_{k=1}^{n-1} |m_k|^{k(n-k)/2}} W^{\nu}_{(1,\ldots,1,\mathrm{sgn}\,(m_{n-1})}\left(\begin{pmatrix} m_1 \cdot \cdots \cdot m_{n-2}|m_{n-1}| & & & \\ & \ddots & & \\ & & m_1 m_2 & \\ & & & m_1 \end{pmatrix} \begin{pmatrix} \gamma & \\ & 1 \end{pmatrix} x\right),$$

where $(a_{m_1,\ldots,m_{n-1}}(f))$ is the infinite array of Fourier–Whittaker coefficients *of f. If f is not cuspidal (that is, f violates the third condition of Definition 6.3.1.4), then we include*

$$m_1, m_2, \ldots, m_{n-1} = 0$$

in the sums.

This state of affairs is very roughly the one encountered in the next subsection, in which we consider metaplectic forms.

6.3.2 Metaplectic theta functions of degree two and three

We will now sketch the process by which one can recover the classical theta function of $\mathbb{Q}(i)$ via Kubota's construction. Our reference for this material is an article by Hoffstein [Hof91].

Let $\lambda = 1 + i$. The principal congruence subgroup is defined as

$$\Gamma_1(\lambda^3) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{Z}[i]) \text{ such that } \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \bmod \lambda^3 \right\},$$

and the Kubota symbol is

$$\kappa(\gamma) = \begin{cases} \left(\frac{c}{d}\right) & c \neq 0 \\ 1 & c = 0, \end{cases}$$

where $\left(\frac{c}{d}\right)$ now stands for the quadratic residue symbol in $\mathbb{Q}(i)$. The law of quadratic reciprocity implies that the Kubota symbol is a homomorphism from $\Gamma_1(\lambda^3)$ into the group of units of $\mathbb{Z}[i]$.

We now define the Eisenstein series. Following the pattern of Subsection 6.3.1, we apply the Iwasawa decomposition to $GL(2, \mathbb{C})$. Indeed, let denote by Z the set of *complex* scalar matrices: then each coset of $GL(2, \mathbb{C})/ZU(2)$ can be represented uniquely by a matrix of the form $z = \left(\begin{smallmatrix} y & x \\ 0 & 1 \end{smallmatrix}\right)$, where $x \in \mathbb{C}$ and $y \in \mathbb{R}_{>0}$. For $s \in \mathbb{C}$ and $z \cdot w \cdot k \in GL(2, \mathbb{C})$, where $w \in Z$ and $k \in U(2)$, we define $I_s(zwk) = y^{2s}$ and the Eisenstein series at the cusp at infinity:

$$E_\infty(z, s) = \sum_{\Gamma_\infty \cap \Gamma_1(\lambda^3)\backslash\Gamma_1(\lambda^3)} \kappa(\gamma) I_s(\gamma z),$$

where $\Gamma_\infty \cap \Gamma_1(\lambda^3)$ consists of the matrices in $\Gamma_1(\lambda^3)$ with $c = 0$. The Eisenstein series at the cusp 0 is then defined to be

$$E_0(z, s) = E_\infty\left(-\frac{1}{z}, s\right),$$

where the appearance of the transformation $-1/z$ instead of $-1/4z$ is related to the fact that the discriminant of $\mathbb{Q}(i)$ is 4, whereas the discriminant of $\mathbb{Q}$ is 1. By construction, $E_0(z, s)$ is an automorphic form on the double cover of $GL(2, \mathbb{C})$ determined by the Kubota symbol:

$$E_0(\gamma z, s) = \kappa(\gamma) E_0(z, s).$$

Using the expression

$$E_0(z,s) = \sum_{\substack{(c,d)=1 \\ c=1 \bmod (\lambda^3) \\ d=1 \bmod (\lambda^3)}} \frac{\left(\frac{c}{d}\right) y^{2s}}{\left(|dx-c|^2+|d|^2y^2\right)^{2s}},$$

we can determine the Fourier coefficients of $E_0(z,s)$. The different $\mathfrak{d}$ of $\mathbb{Q}(i)$ is (2), so we may choose a nonzero $k \in \mathfrak{d}^{-1}(\lambda^3)^{-1}$ so that $m \mapsto e^{2\pi i \mathrm{Tr}(mk)}$ is an additive character of $\mathbb{Z}[i]$ which is trivial on (λ^3). We may write

$$E_0(z,s) = \sum_{m \in \mathbb{Z}[i]} a_m(s,y) e^{2\pi i \mathrm{Tr}(mk)},$$

where

$$a_m(s,y) = \frac{1}{\mu(\mathbb{C}\backslash(\lambda^3))} \int_{\mathbb{C}\backslash(\lambda^3)} E_0(z,s) e^{-2\pi i \mathrm{Tr}(mk)} dx = \frac{y^{2-2s}}{\mu(\mathbb{C}\backslash(\lambda^3))} A_m(s) \int_{\mathbb{C}} \frac{e^{-2\pi i \mathrm{Tr}(mkxy)}}{(|x|^2+1)^{2s}} dx;$$

$$A_m(s) = \sum_{d=1 \bmod (\lambda^3)} \frac{g(m,d)}{(N(d))^{2s}} = \prod_{\substack{p \text{ prime} \\ p \neq \lambda}} A_{m,p}(s); \tag{6.33}$$

$$A_{m,p}(s) = \sum_{l=0}^{\infty} \frac{g(m,p^l)}{(N(p))^{2sl}},$$

$$g(m,d) = \sum_{r \in \mathbb{Z}[i]/(d)} \left(\frac{r\lambda^3}{d}\right) e^{2\pi i \mathrm{Tr}(mr\lambda^3/d)}.$$

Some further computation reveals that the constant term is

$$a_0(s,y) = 2\pi y^{2s-2} \frac{\zeta^{(\lambda)}_{\mathbb{Q}(i)}(4s-2)}{\zeta^{(\lambda)}_{\mathbb{Q}(i)}(4s-1)} \frac{\Gamma(4s-2)}{\Gamma(4s-1)},$$

where the exponent adorning the zeta functions indicates that the factor in the Euler product corresponding to the prime λ is omitted, and for m_0 squarefree,

$$A_{m_0}(s) = \frac{L^{(\lambda)}_{\mathbb{Q}(i)}\left(2s-\frac{1}{2}, \left(\frac{m_0}{\cdot}\right)\right)}{\zeta^{(\lambda)}_{\mathbb{Q}(i)}(4s-1)}.$$

Proceeding along the same lines as Proposition 6.2.2.3, one checks that if m_0 is squarefree,

$$A_{n^2m_0}(s) = A_{m_0}(s) \sum_{\substack{d_1d_2d_3=n \\ (d_2d_3,\lambda)=1}} (N(d_2))^{2-4s} \left(\frac{m_0}{d_3}\right) \mu(d_3) \left(N(d_3)\right)^{1/2-2s}$$

where μ is the Möbius function.

For $m \neq 0$,

$$\frac{y^{2-2s}}{\mu(\mathbb{C}\backslash(\lambda^3))} \int_{\mathbb{C}} \frac{e^{-2\pi i \mathrm{Tr}(mkxy)}}{(|x|^2+1)^{2s}} dx = \frac{y(2\pi)^{2s}(N(m))^{s-\frac{1}{2}} K_{2s-1}(4\pi y|m|)}{\Gamma(2s)},$$

where $K_\nu(w)$ is the modified Bessel function of the second kind. So for $m \neq 0$,

$$a_m(s,y) = \frac{A_m(s) y (2\pi)^{2s} (N(m))^{s-\frac{1}{2}} K_{2s-1}(4\pi y|m|)}{\Gamma(2s)}.$$

By Proposition 1.3.4.2, $A_m(s)$ is entire unless m is a perfect square, in which case it has a simple pole at $s = 3/4$. One computes that in this case, $\mathrm{Res}_{s=3/4} A_{m^2}(s)$ is a constant, independent of m. It follows that if $m \neq 0$, the mth

Fourier coefficient of $\mathrm{Res}_{s=3/4}E_0(z,s)$ vanishes unless m is a square, and has a particularly simple form if m is a square:

$$\mathrm{Res}_{s=3/4}E_0(z,s) = C_1 y^{1/2} + C_2 \sum_{m\in\mathbb{Z}[i]\setminus\{0\}} \sqrt{N(m)}yK_{1/2}\left(4\pi y|m|^2\right)e^{-2\pi i \mathrm{Tr}(mk)}.$$

We have $K_{1/2}(y) = \sqrt{\frac{\pi}{2y}}e^{-y}$, so we have recovered the theta function for $\mathbb{Q}(i)$.

So much for theta functions on $GL(2)$. One can prove that the same procedure produces a theta function over any totally imaginary field, but there is not (yet) a complete, unified treatment for all number fields, owing to the need to account for the *epsilon factors* (already present in the calculation over $\mathbb{Q}$ in Subsection 6.2.3) from the very start. By uniqueness of Whittaker models when $r = n+1$, it is possible to determine the Fourier coefficients for the theta function constructed from the cubic residue symbol over $GL(2)$ (over appropriate fields). This is much more complicated; we refer the interested reader to Patterson's works [Pat77a; Pat77b].

We now turn to Bump and Hoffstein's investigation of a cubic theta function on $GL(3)$. Their analysis is much more complicated than anything else presented so far, so we take an impressionistic approach. We only wish to make plain the close relationship with the quadratic theta functions on $GL(2)$, and to highlight (in the next subsection) that the main ingredients we used to compute asymptotic expansions in Subsection 6.2.3 are also present in this higher degree case.

We begin with the usual explication of the generalised upper half plane. This time, we are apply the Iwasawa decomposition to $GL(3,\mathbb{C})$, using Z to denote the group of *complex* scalar matrices, to obtain that every coset $GL(3,\mathbb{C})/ZU(3)$ may be represented uniquely by

$$\alpha = \begin{pmatrix} 1 & x_2 & x_3 \\ & 1 & x_1 \\ & & 1 \end{pmatrix}\begin{pmatrix} y_1y_2 & & \\ & y_1 & \\ & & 1 \end{pmatrix},$$

where $x_1, x_2, x_3 \in \mathbb{C}$ and $y_1, y_2 \in \mathbb{R}_{>0}$. Some of the formulae are simplified with the introduction of an auxiliary coordinate x_4, satisfying $x_1x_2 = x_3 + x_4$. For $(\nu_1,\nu_2)\in\mathbb{C}^2$, we define

$$I_{(\nu_1,\nu_2)}(\alpha) = y_1^{2\nu_1+\nu_2}y_2^{\nu_1+2\nu_2}.$$

The congruence subgroup we will be dealing with is defined as the subgroup of all matrices in $SL(3,\mathcal{O}_{\mathbb{Q}(\omega)})$ congruent to the identity modulo 3. We denote this group by $\Gamma(3)$. The cubic residue symbol $\left(\frac{b}{a}\right)$ takes values in $\mu_3\cup\{0\}$, and is defined on integers a and b in $\mathcal{O}_{\mathbb{Q}(\omega)}$, where we stipulate that a is not divisible by $\lambda = \sqrt{-3}$, the generator of the different $\mathfrak{d}$ of $\mathbb{Q}(\omega)$.

Now we need to define the Kubota symbol, which is to be a homomorphism from $\Gamma(3)$ into the group of cube roots of unity, effected by the cubic residue symbol. The determination of the Kubota symbol is quite complicated over $GL(n)$ when $n>2$, and was first achieved by Bass, Milnor and Serre [BMS67].

For $GL(3)$, explicit formulae for the Kubota symbol were apparently first given by Bump, Friedberg and Hoffstein, but the article never appeared, so we follow Bump and Hoffstein's "summary" of their results [BH86, pp. 485]. Note that the permutation matrix

$$w = \begin{pmatrix} & & 1 \\ & 1 & \\ 1 & & \end{pmatrix}$$

effects an involution of $GL(3,\mathbb{C})$ by $\iota : g\mapsto wg^{-t}w$. For $g\in GL(3,\mathbb{C})$, collect the bottom rows of g and ιg as $[A_1,B_1,C_1]$ and $[A_2,B_2,C_2]$ respectively. They are called the *invariants* of $GL(3,\mathbb{C})$ as they only depend on the orbit of $GL(3,\mathbb{C})$ in $GL(3,\mathbb{C})_\infty\backslash GL(3,\mathbb{C})$, where $GL(3,\mathbb{C})_\infty$ is the group of upper triangular unipotent matrices in $GL(3,\mathbb{C})$. If $g\in SL(3,\mathcal{O}_{\mathbb{Q}(\omega)})$, then the invariants satisfy

$$(A_1,B_1,C_1) = (A_2,B_2,C_2) = 1 \quad \text{and} \quad A_1C_2 + B_1B_2 + C_1A_2 = 0,$$

and if $g\in\Gamma(3)$,

$$A_1 = A_2 = B_1 = B_2 = 0 \bmod (3), \quad C_1 = C_2 = 1 \bmod (3).$$

Furthermore, the invariants may be factored as

$$B_1 = r_1B_1', \quad B_2 = r_2B_2', \quad C_1 = r_1r_2C_1', \quad C_2 = r_1r_2C_2', \tag{6.34}$$

where $r_1 = r_2 = 1 \bmod (3)$ and $(C_1^{'}, C_2^{'}) = 1$. Finally, we can define the Kubota map:

$$\kappa(g) = \left(\frac{B_1^{'}}{C_1^{'}}\right)\left(\frac{B_2^{'}}{C_2^{'}}\right)\left(\frac{C_1^{'}}{C_2^{'}}\right)^{-1}\left(\frac{A_1}{r_1}\right)\left(\frac{A_2}{r_2}\right),$$

which turns out to be independent of the factorisation 6.34 and is indeed a character on $\Gamma(3)$.

Set $\Gamma_\infty(3) = GL(3,\mathbb{C})_\infty \cap \Gamma(3)$. We can now define the Eisenstein series:

$$E_{(\nu_1,\nu_2)}(\alpha) = \sum_{\Gamma_\infty(3)\backslash\Gamma(3)} \kappa(\gamma) I_{(\nu_1,\nu_2)}(\gamma\alpha).$$

The series defining $E_{(\nu_1,\nu_2)}(\alpha)$ converges absolutely for $\mathrm{Re}(\nu_1), \mathrm{Re}(\nu_2) > \frac{4}{3}$. It has a meromorphic continuation to all of $\mathbb{C}^2$ and satisfies some functional equations in ν_1 and ν_2. It is polar on $\left\{(\nu_1,\nu_2) \mid \nu_1 \text{ or } \nu_2 = \frac{8}{9}\right\}$ and the residue at $\left(\frac{8}{9}, \frac{8}{9}\right)$ is defined to be the cubic theta function:

$$\Theta(\alpha) = \mathrm{Res}_{\nu_1=\frac{8}{9}} \mathrm{Res}_{\nu_2=\frac{8}{9}} E_{(\nu_1,\nu_2)}(\alpha). \tag{6.35}$$

Recall from Subsection 6.3.1 that the Fourier coefficients of automorphic forms on $GL(n)$ are specified by $n-1$-dimensional arrays of indices. For $GL(3)$, this means that the Fourier coefficients are described by pairs (n_1, n_2). Following some intricate manoeuvres involving the maximal parabolic Eisenstein series associated to the cubic metaplectic cover of $GL(2,\mathbb{C})$, Bump and Hoffstein determine the *nondegenerate* (that is, n_1 and n_2 are nonzero) Fourier coefficients of Θ as

$$\begin{aligned} &|n_1n_2|^2 \int_{\mathbb{C}/(3)}\int_{\mathbb{C}/(3)}\int_{\mathbb{C}/(3)} \Theta\left(\begin{pmatrix} 1 & \xi_2 & \xi_3 \\ & 1 & \xi_1 \\ & & 1 \end{pmatrix}\begin{pmatrix} y_1y_2 & & \\ & y_1 & \\ & & 1 \end{pmatrix}\right) \exp\left(4\pi i \mathrm{Re}(-n_1\xi_1 - n_2\xi_2)\right) d\xi_1 d\xi_2 d\xi_3 \\ &= 2\cdot 3^4\pi^2 \mathrm{Res}_{s=\frac{4}{3}} D(s;n_1,n_2) W_{\left(\frac{8}{9},\frac{8}{9}\right)}\left(\begin{pmatrix} n_1n_2 & & \\ & n_1 & \\ & & 1 \end{pmatrix}\begin{pmatrix} y_1y_2 & & \\ & y_1 & \\ & & 1 \end{pmatrix}\right), \end{aligned}$$

where the Whittaker functions are defined for $\mathrm{Re}(\nu_1), \mathrm{Re}(\nu_2) > \frac{2}{3}$ by

$$\begin{aligned} W_{(\nu_1,\nu_2)}(\alpha) = (2\pi)^{-3\nu_1-3\nu_2+1}\Gamma\left(\frac{3\nu_1}{2}\right)\Gamma\left(\frac{3\nu_2}{2}\right)\Gamma\left(\frac{3\nu_1+3\nu_2-2}{2}\right) \\ \int_{\mathbb{C}}\int_{\mathbb{C}}\int_{\mathbb{C}} I_{(\nu_1,\nu_2)}\left(w\begin{pmatrix} 1 & \xi_1 & \xi_3 \\ & 1 & \xi_1 \\ & & 1 \end{pmatrix}\alpha\right) \exp\left(4\pi i \mathrm{Re}(-\xi_1 - \xi_2)\right) d\xi_1 d\xi_2 d\xi_3, \end{aligned}$$

and may be analytically continued to all $(\nu_1,\nu_2) \in \mathbb{C}^2$ by a very general theorem of Jacquet [Jac67], whilst $D(s;n_1,n_2)$ is defined for $\mathrm{Re}(s) > \frac{4}{3}$ by

$$D(s;n_1,n_2) = \prod_{\substack{p\nmid\lambda^6 n_1n_2 \\ p=1 \bmod 3}} \left(1 - \left(\frac{n_1n_2^2}{p}\right)N(p)^{1-\frac{[}{3}2s}\right)^{-1}\left(1 - N(p)^{3-\frac{9}{2}s}\right) \cdot \sum_{\substack{C=1 \bmod 3 \\ p|C \implies p|\lambda^6 n_1n_2}} |C|^{1-3s} \sum_{\substack{C=cd \\ c=d=1 \bmod (3) \\ d|\lambda^3 n_1 c}} \tau\left(n_1cd^{-1}\right) T(c,d), \tag{6.36}$$

and $\tau(k)$ and $T(c,d)$ eventually boil down to expressions involving nothing worse than cubic Gauss sums formed with cubic residue symbols:

$$g(\mu,d) = \sum_{\substack{c \bmod d \\ 3|c}} \left(\frac{c}{d}\right)\exp\left(4\pi \mathrm{Re}\left(\frac{\mu c}{d}\right)\right). \quad \left(\mu \in \mathfrak{d}^{-3},\ d \in \mathcal{O}_{\mathbb{Q}(\omega)},\ d = 1 \bmod (3)\right)$$

It follows that the (n_1,n_2)th Fourier coefficient (the expression at 6.36) is zero unless n_1 and n_2 have the form $n_i = \pm\omega^a\lambda^{b_i} f h_i^3$ for $f = h_i = 1 \bmod (3)$, f squarefree, $b_1 = b_2 \bmod (3)$ and $\tau(\omega^a\lambda^{b_i}) \neq 0$. If this is the case, then the integral is equal to

$$\frac{4\cdot 3^{\frac{7}{2}}\pi^3 3^{\frac{b_1+2b_2}{6}}}{13\zeta_{\mathbb{Q}(\omega)}(3)}\left(\frac{\omega^a\lambda^{b_1}}{f}\right)\tau(\omega^a\lambda^{b_1})|h_1h_2|^2\overline{g(1,f)}W_{\left(\frac{8}{9},\frac{8}{9}\right)}\left(\begin{pmatrix} n_1n_2 & & \\ & n_1 & \\ & & 1 \end{pmatrix}\begin{pmatrix} y_1y_2 & & \\ & y_1 & \\ & & 1 \end{pmatrix}\right).$$

Most remarkably, as a beautiful counterpart to the vanishing of the Fourier coefficients of Jacobi's theta function off the perfect rational squares, the discussion above implies that the $(n_1, 1)$ and $(1, n_2)$ Fourier coefficients vanish unless n_1 and n_2 are perfect cubes [Hof91, Section 4, pp. 89].

6.3.3 Metaplectic theta functions on n-fold covers of $GL(n)$

In this final sitting, we collect together our observations on metaplectic forms and highlight the aspects which we believe bode well for the prospect of obtaining interesting asymptotic expansions.

First, we re-iterate that the *only* theta functions for which one can reasonably expect to be able to determine Fourier coefficients explicitly are those arising from n-fold or $n+1$-fold covers of $GL(n)$. More precisely, these are the theta functions arising from Eisenstein series with unique Whittaker models [Del80; KP84]. In these cases, one may consult [Hof91] for a determination of the structure of the Fourier–Whittaker coefficients using a Hecke operator; very little is known about theta functions on r-fold covers of $GL(n)$ for which $r \neq n$ or $n+1$.

Secondly, although for both the n-fold and the $n+1$-fold theta functions on $GL(n)$, the Fourier coefficients are built out of Gauss sums formed from n-th order residue symbols, only the n-fold metaplectic theta function are associated to a Dirichlet series admitting an Euler product. This is indeed a separate issue to the uniqueness of Whittaker models, as Patterson [Pat77a; Pat77b] was able to determine the Fourier coefficients of the cubic metaplectic theta function on $GL(2)$.

To illustrate this second point, we display the relevant Dirichlet series for the Fourier coefficients of each of the Eisenstein series we have investigated:

1. The metaplectic Eisenstein series over $\mathbb{Q}$ on $GL(2)$ (Proposition 6.2.2.2):

$$A_m(s) = \sum_{\substack{n=1 \\ (n,2)=1}}^{\infty} \frac{\epsilon_n g(-m,n)}{n^{s+\frac{1}{2}}} = \prod_{\substack{p \text{ prime} \\ p \neq 2}} \sum_{l=0}^{\infty} \frac{\epsilon_{p^l} g(-m, p^l)}{p^{l(s+1/2)}}.$$

2. The metaplectic Eisenstein series over $\mathbb{Q}(i)$ on $GL(2)$ at 6.33:

$$A_m(s) = \sum_{d=1 \bmod (\lambda^3)} \frac{g(m,d)}{(N(d))^{2s}} = \prod_{\substack{p \text{ prime} \\ p \neq \lambda}} \sum_{l=0}^{\infty} \frac{g(m,p^l)}{(N(p))^{2sl}}.$$

3. The metaplectic Eisenstein series over $K = \mathbb{Q}\left(-\frac{1}{2} + \frac{\sqrt{3}}{2}\right)$ on $GL(3)$ (Proposition 6.36)[15]:

$$D(s; n_1, n_2) = \prod_{\substack{p \nmid \lambda^6 n_1 n_2 \\ p=1 \bmod 3}} \left(1 - \left(\frac{n_1 n_2^2}{p}\right) N(p)^{1-\frac{[}{3}2s}\right)^{-1} \left(1 - N(p)^{3-\frac{9}{2}s}\right) \cdot \sum_{\substack{C=1 \bmod 3 \\ p|C \implies p|\lambda^6 n_1 n_2}} |C|^{1-3s} \sum_{\substack{C=cd \\ c=d=1 \bmod (3) \\ d|\lambda^3 n_1 c}} \tau\left(n_1 c d^{-1}\right) T(c,d),$$

valid for $\mathrm{Re}(s) > \frac{4}{3}$. In particular, if n_1 and n_2 are cubefree coprime (algebraic) integers congruent to 1 modulo (3) and $n_1 = r^2 t$, where $r = t = 1$ modulo 3 and r and t are coprime squarefree integers, we have

$$D(s, n_1, n_2) = 27|t|^{-1}|r|^{2-3s}\overline{g(1,t)}g(n_2,t)L_K(\chi_{n_1 n_2^2}, 3s/2-1)L_K(\chi_1, 9s/2-3)^{-1}.$$

We also note that the methods used in Subsection 6.3.2 to obtain a cubic metaplectic form on $GL(3)$ have been generalised to $GL(n)$ by Banks, Bump and Lieman [BBL03]. They construct a theta function on the n-fold metaplectic cover of $GL(n)$, for every field K containing the nth roots of unity, such that the Fourier-Whittaker coefficients are formed from L-functions over K twisted by characters of order n.

In Subsection 6.2.3, we showed that with some very basic finite Fourier analysis, one can compute the asymptotic expansion of the quadratic Eisenstein series over $\mathbb{Q}$, and it is clear from the proof that the appearance of quadratic characters in the asymptotic expansion is directly caused by the appearance of L-functions of quadratic characters

[15] See also [BH86, Section 5] for the proofs, which, despite being much more computationally involved, bear striking similarities to the quadratic case.

in the Fourier coefficients. Similarly, the Fourier coefficients of the cubic theta function 6.35 are formed from L-functions twisted by cubic characters, so one suspects that the asymptotic expansion should involve cubic characters.

By a careful treatment of epsilon factors, we intend to verify that an extension of the methods of Section 6.2.3 to the asymptotic expansions of the quadratic Eisenstein series over *every* number field (not just those containing all fourth roots of unity) suffices to recover Hecke reciprocity[16] (Theorem 5.1.2.1), as preparation for the more difficult case of computing the asymptotic expansion of the Bump–Hoffstein theta function on $GL(3)$.

[16]It seems [Brubaker, personal communication] that researchers have avoided the problem of recovering *all* quadratic theta functions as residues of Eisenstein series, preferring to make assumptions on the number of roots of unity contained in the field in order to eliminate ad-hoc constructions when defining the Eisenstein series.

References

[AB50] L. Ahlfors and A. Beurling. "Conformal invariants and function-theoretic null-sets". In: *Acta Math.* 83 (1950), pp. 101–129.

[Apé79] R. Apéry. "Irrationalité de $\zeta(2)$ et $\zeta(3)$". In: *Astérisque* 61 (1979), pp. 11–13.

[Apo73] T. M. Apostol. "Another elementary proof of Euler's formula for $\zeta(2n)$". In: *Amer. Math. Monthly* 80 (1973), pp. 425–431.

[BB05] B. Brubaker and D. Bump. "Residues of Weyl group multiple Dirichlet series associated to $\widetilde{GL}_{n+1}$". In: *Multiple Dirichlet Series, Automorphic Forms and Analytic Number Theory.* Ed. by S. Friedberg et al. Vol. 75. pp. 115–134. Proc. Symp. Pure Math., 2005.

[BBL03] W. Banks, D. Bump, and D. Lieman. "Whittaker–Fourier Coefficients of Metaplectic Eisenstein Series". In: *Compos. Math.* 135 (2003), pp. 153–178.

[BE81] B. C. Berndt and R. J. Evans. "The determination of Gauss sums". In: *Bull. Amer. Math. Soc* 5.2 (1981), pp. 107–129.

[BE82] B. C. Berndt and R. J. Evans. "Corrigendum to "The determination of Gauss sums"". In: *Bull. Amer. Math. Soc. (N.S.)* 7.2 (1982), p. 441.

[Ber13] J. Bernoulli. *Ars Conjectandi.* Reprinted by Culture et Civilization, Bruxelles, 1968. Thurnisiorum, Fratrum, Basileæ, 1713.

[Ber73] B. C. Berndt. "On Gaussian sums and other exponential sums with periodic coefficients". In: *Duke Math. J.* 40.1 (1973), pp. 145–156.

[BEW98] B. C. Berndt, R. J. Evans, and K. S. Williams. *Gauss and Jacobi Sums.* John Wiley & Sons, Inc., 1998.

[BH86] D. Bump and J. Hoffstein. "Cubic metalectic forms on $GL(3)$". In: *Invent. math.* 84 (1986), pp. 481–505.

[BMS67] H. Bass, J. Milnor, and J.-P. Serre. "Solution of the congruence subgroup problem for SL_n $(n \geq 3)$ and Sp_{2n} $(n \geq 2)$". In: *Publ. Math. IHÉS* 33 (1967), pp. 59–137.

[BS10] H. Boylan and N.-P. Skoruppa. "Explicit formulas for Hecke Gauss sums in quadratic number fields". In: *NP. Abh. Math. Semin. Univ. Hambg.* 80 (2010), pp. 213–226.

[BS13] H. Boylan and N.-P. Skoruppa. "A quick proof of reciprocity for Hecke Gauss sums". In: *J. Number Theory* 133 (2013), pp. 110–114.

[BZ10] D. M. Bradley and X. Zhou. "On Mordell-Tornheim sums and multiple zeta values". In: *Ann. Sci. Math. Québec* 34.1 (2010), pp. 15–23.

[Cap90] A. Capelli. "Sur les opérations dans la théorie des formes algébriques". In: *Math. Ann.* 37 (1890), pp. 1–27.

[Cas69] J. W. S. Cassels. "On Kummer sums". In: *Arch. Math. (Brno)* 5 (1969), pp. 79–84.

[Cau40] A.-L. Cauchy. "Méthode simple et nouvelle pour la détermination complète des sommes alternées formées avec les racines primitives des equationes binômes". In: *C. R. Acad. Sci. Paris* 10 (1840), pp. 560–572.

[Che55] S. S. Chern. "An Elementary Proof of the Existence of Isothermal Parameters on a Surface". In: *Proceedings of the American Mathematical Society* 6.5 (1955), pp. 771–782.

[Con] K. Conrad. "Characters of finite Abelian groups". Unpublished. May be found online at: `https://kconrad.math.uconn.edu/blurbs/grouptheory/charthy.pdf`.

[Cox89] D. A. Cox. *Primes of the form $x^2 + ny^2$.* John Wiley & Sons, Inc, 1989.

[CS17] H. Cohen and F. Strömberg. *Modular Forms.* Graduate Studies in Mathematics 179. American Mathematical Society, 2017.

[Del71] P. Deligne. "Formes modulaires et représentations ℓ-adiques". In: *Séminaire N. Bourbaki* (1971). Exposé no. 355, pp. 139–172.

[Del80] P. Deligne. "Sommes de Gauss cubiques et revêtements de $SL(2)$". In: *Séminaire Bourbaki* 21 (1980), pp. 244–277.

[Dic57] L. E. Dickson. *Introduction to the Theory of Numbers*. Dover Publications, 1957.

[Dio21] Diophantus. *Diophanti Alexandrini Arithmeticorum libri sex et de numeris multangulis liber unus*. Ed. by C. G. B. de Méziriac. Bachet's remark on sums of four squares may be found online at: https://archive.org/details/bub_gb_62jsEVHjkfEC/page/n275/mode/2up. 1621.

[Doe55] G. Doetsch. *Handbuch der Laplace-Transformation*. Vol. 15. Lehrbücher und Monographen aus dem Gebiete der Exakten Wissenschaften. Birkhäuser Verlag Basel und Stuttgart, 1955.

[DS05] F. Diamond and J. Shurman. *A First Course in Modular Forms*. Graduate Texts in Mathematics 228. Springer Science and Business Media LLC, 2005.

[Eas08] M. Eastwood. "Symmetries and Overdetermined Systems of Partial Differential Equations". In: vol. 144. The IMA Volumes in Mathematics and its Applications. Spinger, 2008. Chap. Notes on Projective Differential Geometry.

[Eic77] M. Eichler. "On theta functions of real algebraic number fields". In: *Acta Arithmetica* 33.3 (1977), pp. 269–292.

[Eis44] G. Eisenstein. "Geometrischer Beweis des Fundamentaltheorems für die quadratischen Reste". In: *J. Reine Angew. Math.* 28 (1844), pp. 246–248.

[Est45] T. Estermann. "On the sign of the Gaussian sum". In: *J. London Math. Soc* 20 (1945), pp. 66–67.

[Est47] T. Estermann. "Elementary evaluation of $\zeta(2k)$". In: *J. London Math. Soc.* 22 (1947), pp. 10–13.

[Eul38] L. Euler. "Methodus Generalis summandi progressiones". In: *Novi Comment. Acad. Sci. Imp. Petropol.* 6 (1738), pp. 68–97.

[Eul48] L. Euler. *Introduction in analysum infinitorum*. Translated from the Latin by J. D. Blantan. Reprinted Springer-Verlag, New York, 1988. Marcum–Michaelem Bosquet, Lausanne, 1748.

[FB09] E. Freitag and R. Busam. *Complex Analysis*. Universitext. Springer-Verlag Berlin Heidelberg, 2009.

[For29] L. R. Ford. *Automorphic Functions*. AMS Chelsea Publishing, 1929.

[Gau01a] C. F. Gauss. "Disquisitiones Arithmeticae". In: Leipzig: Gerh. Fleischer, 1801. Chap. Art. 125–145, pp. 94–145.

[Gau01b] C. F. Gauss. "Disquisitiones Arithmeticae". In: Leipzig: Gerh. Fleischer, 1801. Chap. Art. 262, p. 262.

[Gau08] C. F. Gauss. "Theorematis arithmetici demonstratio nova". In: *Comment. Soc. Reg. Sci. Gottingensis* 69 (1808).

[Gau11] C. F. Gauss. "Summatio quarandum serierium singularium". In: *Comment. Soc. Reg. Sci. Gottingensis* (1811). May be found online at: https://gdz.sub.uni-goettingen.de/id/PPN602151724?tify={"view":"export"}.

[GE18] A. R. Gover and M. Eastwood. "Volume growth and puncture repair in conformal geometry". In: *J. Geom. Phys.* 127 (2018), pp. 128–132.

[Gel75] S. S. Gelbart. *Automorphic Forms on Adele Groups*. Vol. 83. Annals of Mathematics Studies. Princeton University Press, 1975.

[Gen52] A. Genocchi. "Sulla formula sommatoria di Eulero, e sulla teorica de'residui quadratici". In: *Ann. Sci. Mat. Fis. (Roma)* 3 (1852), pp. 406–436.

[Gen53] A. Genocchi. "Note sur la théorie des résidus quadratiques". In: *Ann. Sci. Mat. Fis. (Roma)* 25 (1853), p. 54.

[GH85] D. Goldfeld and J. Hoffstein. "Eisenstein series of $\frac{1}{2}$ integral weight and the mean value of real Dirichlet L-series". In: *Invent. Math.* 80 (1985), pp. 185–208.

[Gla84] J. W. L. Glaisher. "On the function $\chi(n)$". In: *Quarterly Journal of Pure and Applied Mathematics* 20 (1884), pp. 97–167.

[Gol06] D. Goldfeld. *Automorphic Forms and L-functions for the group* $GL(n, \mathbb{R})$. Cambridge Studies in Advanced Mathematics 99. Cambridge University Press, 2006.

[GS93] M. Gromov and M. A. Shubin. "The Riemann–Roch Theorem for Elliptic Operators". In: *Adv. in Soviet Math.* 16.1 (1993), pp. 211–241.

[Gui45] A. P. Guinand. "Gauss sums and primitive characters". In: *The Quarterly Journal of Mathematics* 16.1 (1945), pp. 59–63.

[Gun66] R. C. Gunning. *Lectures on Riemann Surfaces*. Princeton Mathematical Notes: Princeton University Press, 1966.

[Har40] G. H. Hardy. *Ramanujan*. Cambridge, 1940.

[Har77] R. Hartshorne. *Algebraic Geometry*. Vol. 52. Graduate Texts in Mathematics. Springer, 1977.

[HBP79] R. Heath-Brown and S. J. Patterson. “The distribution of Kummer sums at prime arguments”. In: *J. Reine Angew. Math.* 310 (1979), pp. 111–130.

[Hec23] E. Hecke. *Vorlesungen über die Theorie der algebraischen Zahlen.* Translated from the German by George U. Brauer and Jay R. Goldman, with the assistance of R. Kotzen. Akademische Verlagsgesellschaft, 1923.

[Hec36] E. Hecke. “Über die Bestimmung Dirichletscher Reihen durch ihre Funktionalgleichung”. In: *Math. Ann.* 112.1 (1936), pp. 664–699.

[Hec81] E. Hecke. *Lectures on the Theory of Algebraic Numbers.* Vol. 77. Graduate Texts in Mathematics. Springer-Verlag, 1981.

[HM73] D. Husemoller and J. Milnor. *Symmetric Bilinear Forms.* Ergeb. Math. Grenzeb., 1973.

[Hof91] J. Hoffstein. “Eisenstein Series and Theta Functions on the Metaplectic Group”. In: *Theta Functions From the Classical to the Modern.* Ed. by M. Ram Murty. Vol. 1. CRM Proceedings & Lecture Notes. American Mathematical Society, 1991.

[Hör73] L. Hörmander. *An Introduction to Complex Analysis in Several Variables.* North-Holland, 1973.

[Hur91] A. Hurwitz. “Über Riemann’sche Fläche mit gegebenen Verzweigungspunkten”. In: *Math. Ann.* 39 (1891), pp. 1–61.

[Hur93] A. Hurwitz. “Über algebraische Gebilde mit eindeutigen Transformationen in sich”. In: *Math. Ann.* 41 (1893), pp. 403–442.

[Huy05] D. Huybrechts. *Complex Geometry: An Introduction.* Universitext. Springer, 2005.

[IR90] K. Ireland and M. Rosen. *A Classical Introduction to Modern Number Theory.* Vol. 84. Graduate Texts in Mathematics. Springer-Verlag New York, 1990.

[Jac29] C. G. J. Jacobi. *Fundamenta nova theoriae functionarum ellipticarum.* Reprinted in Gessamelte Werke, vol. 1, 49–239. Bornträger, Königsberg, 1829.

[Jac46] C. G. J. Jacobi. “Über die Kreisteilung und ihre Anwedung auf die Zahlentheorie”. In: *J. Reine. Angew. Math.* 30 (1846), pp. 166–182.

[Jac67] H. Jacquet. “Fonctions de Whittaker associées aux groupes de Chevalley”. In: *Bull. Soc. Math. Fr.* 95 (1967), pp. 243–309.

[Kna92] A. Knapp. *Elliptic Curves.* Mathematical Notes 40. Princeton University Press, 1992.

[Kob84] N. Koblitz. *Introduction to Elliptic Curves and Modular Forms.* Graduate Texts in Mathematics 97. Springer-Verlag, 1984.

[Kob95] S. Kobayashi. *Transformation Groups in Differential Geometry.* Classics in Mathematics. Springer-Verlag Berlin Heidelberg, 1995.

[Kod86] K. Kodaira. *Complex Manifolds and Deformation of Complex Structures.* Classics in Mathematics. Originally published as Vol. 283 in *Grundlehren der mathematischen Wissenschaften.* Translated from the Japanese by Kazuo Akao. Springer-Verlag, 1986.

[Kor14] A. Korn. “Zwei Anwendungen der Methode der sukzessiven Annäherungen”. In: *Schwarz Abhandlungen* (1914), pp. 215–299.

[KP84] D. A. Kazhdan and S. J. Patterson. “Metaplectic forms”. In: *Publ. Math. IHÉS* 59 (1984), pp. 35–142.

[Kra72] I. Kra. *Automorphic Functions and Kleinian Groups.* Benjamin, Reading, MA, 1972.

[Kub69] T. Kubota. *On Automorphic Functions and the Reciprocity Law in a Number Field.* Vol. 2. Lectures in Mathematics. Department of Mathematics, Kyoto University. Kinokuniya Book-Store Co. Ltd. Tokyo, Japan, 1969.

[Kum42] E. E. Kummer. “Eine Aufgabe, betreffend die Theorie der kubischen Reste”. In: *J. Reine Angew. Math.* 23 (1842), pp. 285–286.

[Lag72] J. L. Lagrange. “Démonstration d’un théorème d’arithmétique”. In: *Nouveaux memoires de l’académie royale des sciences et belles-lettres de Berlin* (1772), pp. 123–133.

[Lan70] R. P. Langlands. “Lectures in modern analysis and applications III”. In: ed. by C. T. Taam. Vol. 170. Springer Lecture Notes in Mathematics. Springer, 1970. Chap. Problems in the Theory of Automorphic Forms, pp. 18–61.

[Lan93] G Landsberg. “Zur Theorie der Gaussschen Summen und der linearen Transformation der Theta functionen”. In: *J. Reine Angew Math.* 111 (1893). May be found online at: `https://archive.org/details/journalfrdierei97crelgoog/page/n241`, pp. 234–253.

[Lee12] J. Lee. *Introduction to Smooth Manifolds.* Vol. 218. Graduate Texts in Mathematics. Springer-Verlag New York, 2012.

[Leh47] D. H. Lehmer. “The vanishing of Ramanujan’s function $\tau(n)$”. In: *Duke Math. J.* 14.2 (1947), pp. 429–433.

[Leh64] J. Lehner. *Discontinuous Groups and Automorphic Functions*. Mathematical Surveys 8. AMS, Providence, RI, 1964.

[Lic16] L. Lichtenstein. "Zur Theorie der konformen Abbildung, Konforme Abbildung nichtanalytischer, singularitätenfreier Flächenstücke auf eben Gebiete". In: *Bull. Int. Acad. Sci. Cracovie* Ser. A (1916), pp. 192–217.

[Lin47] E. L. Lindelöf. *Le calcul des résidus*. Chelsea, New York, 1947.

[Lio58] J. Liouville. "Sur quelques formules gènérales qui peuvant ètre utiles dans la théorie des nombres". In: *J. Math. Pures Appl.* 3.2 (1858), 143–152, 193–200, 201–208, 241–250, 273–288 and 325–336.

[Lox76] J. H. Loxton. "On the determination of Gauss sums". In: *Sem. Delange-Pisot-Poitou* 18.12 (1976/1977).

[Lox78] J. H. Loxton. "Some conjectures concerning Gauss sums". In: *J. Reine Angew. Math.* 297 (1978), pp. 153–158.

[Maa38] H. Maaß. "Konstruktion ganzer Modulformen halbzahliger Dimension mit ϑ-Multiplikatoren in einer und zwei Variabeln". In: *Math. Z.* 43 (1938), pp. 709–738.

[Mac42] C. Maclaurin. *A Treatise of Fluxions*. Vol. 2. T. W. and T. Ruddimans, Edinburgh, 1742.

[Mat79a] C. R. Matthews. "Gauss sums and elliptic functions. I. The Kummer sum". In: *Invent. Math.* 52 (1979), pp. 163–185.

[Mat79b] C. R. Matthews. "Gauss sums and elliptic functions. II. The quartic sum". In: *Invent. Math.* 54 (1979), pp. 23–52.

[Mil72] J. Milnor. *Introduction to Algebraic K-Theory*. Vol. 72. Annals of Mathematical Studies. Princeton Universiy Press, 1972.

[MM97] H. McKean and V. Moll. *Elliptic Curves*. Cambridge University Press, 1997.

[Moo20] B. Moore. "A proof of the Landsberg–Schaar relation by finite methods". In: *Ramanujan J.* (2020). DOI: 10.1007/s11139-019-00195-4.

[Mor33] L. Mordell. "The definite integral $\int_{-\infty}^{\infty} e^{ax^2+bx}/\left(e^{cx}+d\right)$ and the analytic theory of numbers". In: *Acta Math.* 6 (1933), pp. 177–180.

[Mor58] L. J. Mordell. "On the Evaluation of Some Multiple Series". In: *J. London Math. Soc.* 33.3 (1958), pp. 368–371.

[Mor62] L. Mordell. "The sign of the Gaussian sum". In: *Illinois J. Math.* 6 (1962), pp. 177–180.

[Mov12] H. Movasati. "Quasi-modular forms attached to elliptic curves, I". In: *Ann. Math. Blaise Pascal* 19.2 (2012), pp. 307–377.

[MP05] M. Ram Murty and A. Pacelli. "Quadratic reciprocity via theta functions". In: vol. 1. Ramanujan Math. Society Lecture Notes, 2005, pp. 107–116.

[MP17] M. Ram Murty and S. Pathak. "Evaluation of the quadratic Gauss sum". In: *Mathematics Student* 86 (2017), pp. 139–150.

[Mur08] M. Ram Murty. *Problems in Analytic Number Theory*. Vol. 206. Graduate Texts in Mathematics. Springer, 2008.

[Mur82] M. Ram Murty. "Some Omega results for Ramanujan's tau function". In: *Proceedings of the Number Theory Conference at Mysore*. Vol. 938. Springer Lecture Notes. 1982, pp. 123–137.

[Nek91] J. Nekovář. "Beilinson's conjectures". In: *Motives*. Vol. 55. Proc. Sympos. Pure Math. Part 1. Amer. Math. Soc., Providence, RI, 1991, pp. 537–570.

[Neu99] J. Neukirch. *Algebraic Number Theory*. Vol. 322. Grundlehren der mathematischen Wissenschaften. Translated from the German by Norbert Schappacher. Springer-Verlag Berlin Heidelberg New York, 1999.

[NN57] A. Newlander and L. Nirenberg. "Complex analytic coordinates in almost complex manifolds". In: *Ann. Math.* 65 (1957), pp. 391–404.

[OT05] V. Ovsienko and S. Tabachnikov. *Projective Differential Geometry Old and New. From the Schwarzian Derivative to the Cohomology of Diffeomorphism Groups*. Vol. 165. Cambridge Tracts in Mathematics. Cambridge University Press, 2005.

[OT09] V. Ovsienko and S. Tabachnikov. "What is... the Schwartzian Derivative?" In: *Notices of the AMS* 56.1 (2009), pp. 34–36.

[Pat77a] S. J. Patterson. "A cubic analogue of the theta series". In: *J. Reine Angew. Math* 296 (1977), pp. 125–161.

[Pat77b] S. J. Patterson. "A cubic analogue of the theta series. II." In: *J. Reine Angew. Math* 296 (1977), pp. 217–220.

[Poi93] S. D. Poisson. “Mémoire sur le calcul mumérique des integrales défines”. In: *Mem. Acad. Roy. Sci. Paris* (1893), pp. 570–602.

[Pól27] G. Pólya. “Elementaren Beweis einer Thetaformel”. In: *Sitz. der Phys. Math. Klasse.* Reprinted in *Collected Papers*, vol. 1, 303–306. MIT Press, Cambridge, MA, 1974–1984. Preuss. Akad. der Wiss., Berlin, 1927.

[Pro85a] N. V. Proskurin. “Automorphic functions and the Bass–Milnor–Serre homomorphism, I”. In: *J. Sov. Math.* 29 (1985), pp. 1160–1191.

[Pro85b] N. V. Proskurin. “Automorphic functions and the Bass–Milnor–Serre homomorphism, II”. In: *J. Sov. Math.* 29 (1985), pp. 1192–1219.

[Ram16] S. Ramanujan. “On certain arithmetical functions”. In: *Trans. Cambridge Philos. Soc.* 22 (1916), pp. 159–184.

[Ram57] S. Ramanujan. *Notebooks (2 volumes).* Tata Institute of Fundamental Research, Bombay, 1957.

[Ran76] R. A. Rankin. “Elementary proofs of relationships between Eisenstein series”. In: *Proc. Roy. Soc. Edinburgh. Sect. A* (1976), pp. 107–117.

[Sch50] M Schaar. “Mémoire sur la théorie des résidus quadratiques”. In: *Acad. Roy. Sci. Lettres Beaux Arts Belgique* 24 (1850). May be found online at: `https://archive.org/details/mmoiresdelacad24acad/page/n465/`.

[Ser81] J.-P. Serre. “Quelques applications du théorème de densité de Chebotarev”. In: *Publ. Math. IHÉS* 54 (1981), pp. 123–201.

[Sha51] D. Shanks. “A short proof of an identity of Euler”. In: *Proc. AMS* 2 (1951), pp. 747–749.

[Sha74] J. A. Shalika. “The multiplicity one theorem for $GL(n)$”. In: *Ann. Math.* 100 (1974), pp. 171–193.

[Shi71] G. Shimura. *Introduction to the Arithmetic Theory of Automorphic Functions.* Princeton University Press, 1971.

[Shi72] G. Shimura. “On modular forms of half-integral weight”. In: *Modular Functions of One Variable I.* Lecture Notes in Mathematics 320. Springer, 1972, pp. 57–74.

[Shi75] G. Shimura. “On the holomorphy of certain Dirichlet series”. In: *Proc. London Math. Soc.* 31.3 (1975), pp. 79–98.

[Sko93] N.-P. Skoruppa. “A quick combinatorial proof of Eisenstein series identities”. In: *J. Number Theory* 43.1 (1993), pp. 68–73.

[Sny97] J. P. Snyder. *Flattening the Earth: Two Thousand Years of Map Projections.* Chicago University Press, 1997.

[ST87] I. N. Stewart and D. O. Tall. *Algebraic Number Theory.* Chapman & Hall / CRC, 1987.

[Syl84] J. J. Sylvester. “Question 7382”. In: *Mathematical Questions with their solutions, from the “Educational Times”.* Vol. 41. May be found online at: `https://archive.org/details/mathematicalque10millgoog/page/n5/mode/2up`. Hodgson, 1884, p. 21.

[Tor50] L. Tornheim. “Harmonic Double Series”. In: *Am. J. Math.* 72.2 (1950), pp. 303–314.

[vS09] A. Čap and J. Slovák. *Parabolic Geometries.* Vol. 154. Mathematical Surveys and Monographs. American Mathematical Society, 2009.

[Wal64] C. T. C. Wall. “Quadratic forms on finite groups, and related topics”. In: *Topology* 2 (1964), pp. 281–298.

[Wei64] A. Weil. “Sur certains groupes d’opérateurs unitaires”. In: *Acta Math.* 111 (1964), pp. 143–211.

[Wei67] A. Weil. “Über die Bestimmung Dirichletster Riehen durch Funktionalgleichungen”. In: *Math. Ann.* 168 (1967), pp. 149–156.

[Wel86] R. O. Wells. *Differential Analysis on Complex Manifolds.* Vol. 65. Graduate Texts in Mathematics. Springer, 1986.

[Win38] A. Wintner. *Asymptotic distributions and infinite convolutions.* Princeton University Press, 1938.

[Zag08] D. Zagier. “Elliptic Modular Forms and Their Applications”. In: *The 1-2-3 of Modular Forms. Lectures at a Summer School in Nordfjordeid, Norway.* Springer-Verlag Berlin Heidelberg, 2008.

[Zeu71] H. G. Zeuthen. “Nouvelle démonstration des théorèmes sur les séries de points correspondants sur deux courbes.” In: *Math. Ann.* 3 (1871). May be found online at: `http://reader.digitale-sammlungen.de/de/fs1/object/display/bsb11043309_00160.html`, pp. 150–156.

[Zyg88] A. Zygmund. *Trigonometric Series.* (Volume I and II combined). Cambridge University Press, 1988.