Developments in Mathematics

Krishnaswami Alladi
Manjul Bhargava
David Savitt
Pham Huu Tiep *Editors*

Quadratic and Higher Degree Forms

Springer

Developments in Mathematics

VOLUME 31

For further volumes:
http://www.springer.com/series/5834

Krishnaswami Alladi • Manjul Bhargava
David Savitt • Pham Huu Tiep
Editors

Quadratic and Higher Degree Forms

Springer

Editors
Krishnaswami Alladi
Department of Mathematics
University of Florida
Gainesville, FL, USA

Manjul Bhargava
Department of Mathematics
Princeton University
Princeton, NJ, USA

David Savitt
Department of Mathematics
University of Arizona
Tucson, AZ, USA

Pham Huu Tiep
Department of Mathematics
University of Arizona
Tucson, AZ, USA

ISSN 1389-2177
ISBN 978-1-4614-7487-6 ISBN 978-1-4614-7488-3 (eBook)
DOI 10.1007/978-1-4614-7488-3
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013940058

Mathematics Subject Classification (2010): 11E04, 11E16, 11E20, 11E70, 11F03, 05B05, 11H71, 33C70, 33C55, 11F37, 11B65, 05A30, 33E05, 11E12, 11D85, 11E25, 11E76, 11P05, 11N10, 11R52, 11R04

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

There have been dramatic developments in the areas of quadratic and higher degree forms in recent years, and so the time seemed opportune to convene meetings devoted to these topics. During March 2009 there were two major conferences in the area of quadratic forms. One was a research conference at the University of Florida in Gainesville, on “Quadratic forms, sums of squares, and integral lattices” where the latest advances were presented. Immediately after this was the Arizona Winter School on “Quadratic Forms” at the University of Arizona in Tucson, which was an instructional workshop for graduate students with the goal of preparing them for research in this important area. These two conferences were followed by the Conference on Higher Degree Forms at the University of Florida in May 2009.

This volume is an outgrowth of these three conferences, all of which were completely funded by the National Science Foundation. We gratefully acknowledge this support from the NSF. The Tucson conference was the twelfth Arizona Winter School, a longstanding series of NSF-supported workshops on topics in arithmetic geometry. The two Gainesville conferences were in keeping with the tradition there of having annual conferences on various aspects of number theory; they were followed by two Focused Weeks (one on quadratic forms and another on the related topic of integral lattices) at the University of Florida during the Spring of 2010, also fully supported by the NSF. The PIs for the 2009 Florida NSF grant DMS-0753080 were Krishnaswami Alladi and Pham Tiep (then at the University of Florida), with Manjul Bhargava (Princeton) as a consultant. The PIs for the Arizona Winter School NSF grant DMS-0602287 were Matthew Papanikolas, Fernando Rodriguez-Villegas, David Savitt, William Stein, and Dinesh Thakur.

The Arizona Winter School featured instructional lectures by Manjul Bhargava, John Conway, Noam Elkies, Jonathan Hanke, and R. Parimala on various aspects of quadratic forms. The informal (but comprehensive) notes of these lectures are available at the website of the 2009 Arizona Winter School (http://swc.math.arizona.edu). Parimala and Hanke have polished their articles and submitted excellent surveys to this volume.

Even though the Florida conference on quadratic forms was a research conference focusing on the latest developments, there was significant participation by graduate and undergraduate students to help them enter this exciting domain of research. In order to prepare them for the advanced conference lectures, an instructional workshop preceded this conference for which Jonathan Hanke was the main lecturer. Some aspects of his Florida talks are covered in his survey paper in this book.

In his survey, Hanke discusses fundamental connections between the classical theory of quadratic forms over number fields and their rings of integers, and the theory of modular and automorphic forms. In doing so he provides a treatment of theta functions and some aspects of Clifford algebras as well. Hanke's survey is nicely complemented by that of Parimala who provides a lucid introduction to the algebraic theory of quadratic forms, the invariants associated with quadratic forms, and connections with Galois cohomology. She also states some open problems and discusses recent progress. These two surveys are augmented by the survey and research paper of Voight on quaternion algebras and quadratic forms.

The classical theorems of Lagrange that every integer is a sum of four squares and Gauss that every integer is a sum of three triangular numbers motivate the study of "universal forms", namely those that represent all integers, as well as the investigation of ternary forms in general. The papers of Jagy on integral positive ternary quadratic forms, of Berkovich on sums of three squares, and of Chan and Haensch on certain almost universal ternary forms, show that there still are fundamental questions worthy of investigation on very classical topics.

Whereas the study of universal quadratic forms addresses the question of representing all integers, one could consider the question of representing quadratic forms by integral quadratic forms. In 2008 Ellenberg and Venkatesh introduced ergodic theory as a new tool in this study and made dramatic progress going beyond what Eichler and Kneser had achieved using an arithmetic approach. In his survey of such representation problems, Schulze-Pillot sketches three approaches—arithmetic, algebraic and ergodic—and gives a comparative study of them.

The theory of integral lattices has important links with quadratic forms. Bannai and Miezaki discuss a famous conjecture of D. H. Lehmer on the Fourier coefficients of weighted theta series of certain integral lattices and describe recent progress on this classical question. Integral lattices and quadratic forms have links with binary linear codes, and this is investigated by Elkies and Kominers. In doing so, they provide a new structural development of harmonic polynomials on Hamming space analogous to the treatment of harmonic polynomials on Euclidean space, and present several applications.

Finally, the paper of Reznick discusses certain fundamental questions on the length of binary forms of higher degree starting from the seminal work of Sylvester in the mid-nineteenth century. After discussing some current research, he concludes with a list of important open questions.

We hope that this volume, which comprises both introductory survey articles and research papers reporting the latest developments, will be of interest to students and senior mathematicians alike. In conducting the conferences in Florida, we owe a special debt to Frank Garvan as a conference organizer and to Margaret Somers for taking care of all local arrangements. Similarly, we wish to acknowledge Annette

Horn for handling the local arrangements for the 2009 Arizona Winter School. We thank Elizabeth Loew of Springer for her support and interest in including this book in the series Developments in Mathematics.

Gainesville, FL, USA Krishnaswami Alladi
Princeton, NJ, USA Manjul Bhargava
Tucson, AZ, USA David Savitt
Tucson, AZ, USA Pham Huu Tiep

Contents

Toy Models for D. H. Lehmer's Conjecture II

Eiichi Bannai and Tsuyoshi Miezaki*

Abstract In the previous paper under the same title, we showed that the m-th Fourier coefficient of the weighted theta series of the $\mathbb{Z}^2$-lattice and the A_2-lattice does not vanish when the shell of norm m of those lattices is not the empty set. In other words, the spherical 4 (resp. 6)-design does not exist among the nonempty shells in the $\mathbb{Z}^2$-lattice (resp. A_2-lattice). This paper is the sequel to the previous paper. We take 2-dimensional lattices associated to the algebraic integers of imaginary quadratic fields whose class number is either 1 or 2, except for $\mathbb{Q}(\sqrt{-1})$ and $\mathbb{Q}(\sqrt{-3})$, then, show that the m-th Fourier coefficient of the weighted theta series of those lattices does not vanish, when the shell of norm m of those lattices is not the empty set. Equivalently, we show that the corresponding spherical 2-design does not exist among the nonempty shells in those lattices.

Key words and Phrases Weighted theta series • Spherical t-design • Modular forms • Lattices • Hecke operator

Mathematics Subject Classification (2010): Primary 11F03; Secondary 05B30; Tertiary 11R04

*The second author was supported by a JSPS Research Fellowship.

E. Bannai (✉)
Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, China
e-mail: bannai@sjtu.edu.cn

T. Miezaki
Oita National College of Technology, 1666 Oaza-Maki, Oita 870-0152, Japan

Yamagata University, 1-4-12 Kojirakawa, Yamagata 990-8560, Japan
e-mail: miezaki@oita-ct.ac.jp; miezaki@e.yamagata-u.ac.jp

K. Alladi et al. (eds.), *Quadratic and Higher Degree Forms*, Developments in Mathematics 31, DOI 10.1007/978-1-4614-7488-3_1,

1 Introduction

The concept of spherical t-design is due to Delsarte-Goethals-Seidel [7]. For a positive integer t, a finite nonempty subset X of the unit sphere

$$S^{n-1} = \{x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n \mid x_1^2 + x_2^2 + \cdots + x_n^2 = 1\}$$

is called a spherical t-design on S^{n-1} if the following condition is satisfied:

$$\frac{1}{|X|}\sum_{x\in X} f(x) = \frac{1}{|S^{n-1}|}\int_{S^{n-1}} f(x)d\sigma(x),$$

for all polynomials $f(x) = f(x_1, x_2, \cdots, x_n)$ of degree not exceeding t. Here, the righthand side means the surface integral on the sphere, and $|S^{n-1}|$ denotes the surface volume of the sphere S^{n-1}. The meaning of spherical t-design is that the average value of the integral of any polynomial of degree up to t on the sphere is replaced by the average value at a finite set on the sphere. A finite subset X in $S^{n-1}(r)$, the sphere of radius r centered at the origin, is also called a spherical t-design if $\frac{1}{r}X$ is a spherical t-design on the unit sphere S^{n-1}.

We denote by $\mathrm{Harm}_j(\mathbb{R}^n)$ the set of homogeneous harmonic polynomials of degree j on $\mathbb{R}^n$. It is well known that X is a spherical t-design if and only if the condition

$$\sum_{x\in X} P(x) = 0$$

holds for all $P \in \mathrm{Harm}_j(\mathbb{R}^n)$ with $1 \le j \le t$ [7]. If the set X is antipodal, that is $-X = X$, and j is odd, then the above condition is fulfilled automatically. So we reformulate the condition of spherical t-design on the antipodal set as follows:

Proposition 1.1. *A nonempty finite antipodal subset $X \subset S^{n-1}$ is a spherical $2s+1$-design if the condition*

$$\sum_{x\in X} P(x) = 0$$

holds for all $P \in \mathrm{Harm}_{2j}(\mathbb{R}^n)$ with $2 \le 2j \le 2s$.

It is known [7] that there is a natural lower bound (Fisher type inequality) for the size of a spherical t-design in S^{n-1}. Namely, if X is a spherical t-design in S^{n-1}, then

$$|X| \ge \binom{n-1+[t/2]}{[t/2]} + \binom{n+[t/2]-2}{[t/2]-1}$$

if t is even, and

$$|X| \geq 2\binom{n-1+[t/2]}{[t/2]} \tag{1}$$

if t is odd.

A lattice in $\mathbb{R}^n$ is a subset $\Lambda \subset \mathbb{R}^n$ with the property that there exists a basis $\{v_1, \cdots, v_n\}$ of $\mathbb{R}^n$ such that $\Lambda = \mathbb{Z}v_1 \oplus \cdots \oplus \mathbb{Z}v_n$, i.e., Λ consists of all integral linear combinations of the vectors $v_1, \cdots, v_n$. The dual lattice Λ is the lattice

$$\Lambda^\sharp := \{y \in \mathbb{R}^n \mid (y, x) \in \mathbb{Z}, \text{ for all } x \in \Lambda\},$$

where (x, y) is the standard Euclidean inner product. The lattice Λ is called integral if $(x, y) \in \mathbb{Z}$ for all x, $y \in \Lambda$. An integral lattice is called even if $(x, x) \in 2\mathbb{Z}$ for all $x \in \Lambda$, and it is odd otherwise. An integral lattice is called unimodular if $\Lambda^\sharp = \Lambda$. For a lattice Λ and a positive real number $m > 0$, the shell of norm m of Λ is defined by

$$\Lambda_m := \{x \in \Lambda \mid (x, x) = m\} = \Lambda \cap S^{n-1}(\sqrt{m}).$$

Let $\mathbb{H} := \{z \in \mathbb{C} \mid \mathrm{Im}\,(z) > 0\}$ be the upper half-plane.

Definition 1.1. Let Λ be the lattice of $\mathbb{R}^n$. Then for a polynomial P, the function

$$\Theta_{\Lambda,P}(z) := \sum_{x \in \Lambda} P(x) e^{i\pi z(x,x)}$$

is called the theta series of Λ weighted by P.

Remark 1.1 (See Hecke [9], Schoeneberg [19, 20]).

(i) When $P = 1$, we get the classical theta series

$$\Theta_\Lambda(z) = \Theta_{\Lambda,1}(z) = \sum_{m \geq 0} |\Lambda_m| q^m, \text{ where } q = e^{\pi i z}.$$

(ii) The weighted theta series can be written as

$$\begin{aligned}\Theta_{\Lambda,P}(z) &= \sum_{x \in \Lambda} P(x) e^{i\pi z(x,x)} \\ &= \sum_{m \geq 0} a_m^{(P)} q^m, \text{ where } a_m^{(P)} := \sum_{x \in \Lambda_m} P(x).\end{aligned}$$

These weighted theta series have been used efficiently for the study of spherical designs which are the nonempty shells of Euclidean lattices. (See [5, 6, 16, 23, 24]. See also [2].)

Lemma 1.1 (cf. [23, 24], [16, Lemma 5]). *Let Λ be an integral lattice in $\mathbb{R}^n$. Then, for $m > 0$, the non-empty shell Λ_m is a spherical t-design if and only if*

$$a_m^{(P)} = 0$$

for all $P \in \mathrm{Harm}_{2j}(\mathbb{R}^n)$ *with* $1 \le 2j \le t$, *where* $a_m^{(P)}$ *are the Fourier coefficients of the weighted theta series*

$$\Theta_{\Lambda,P}(z) = \sum_{m \ge 0} a_m^{(P)} q^m.$$

We recall the definition of a modular form.

Definition 1.2. Let $\Gamma \subset SL_2(\mathbb{R})$ be a Fuchsian group of the first kind and let χ be a character of Γ. A holomorphic function $f : \mathbb{H} \to \mathbb{C}$ is called a modular form of weight k for Γ with respect to χ, if the following conditions are satisfied:

$$\text{(i)} \qquad f\Big(\frac{az+b}{cz+d}\Big) = \Big(\frac{cz+d}{\chi(\sigma)}\Big)^k f(z) \text{ for all } \sigma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma.$$

$$\text{(ii)} \qquad f(z) \text{ is holomorphic at every cusp of } \Gamma.$$

If $f(z)$ has period N, then $f(z)$ has a Fourier expansion at infinity, [11]:

$$f(z) = \sum_{m=0}^{\infty} a_m q_N^m, \ q_N = e^{2\pi i z/N}.$$

We remark that for $m < 0$, $a_m = 0$, by the condition (ii). A modular form with constant term $a_0 = 0$, is called a cusp form. We denote by $M_k(\Gamma, \chi)$ (resp. $S_k(\Gamma, \chi)$) the space of modular forms (resp. cusp forms) with respect to Γ with the character χ. When f is the normalized eigenform of Hecke operators, p. 163, [11], the Fourier coefficients satisfy the following relations:

Lemma 1.2 (cf. [11], Proposition 32, 37, 40, Exercise 2, p. 164). *Let* $\alpha \in \mathbb{N}$ *and* $f(z) = \sum_{m \ge 1} a(m) q^m \in S_k(\Gamma, \chi)$. *If* $f(z)$ *is the normalized eigenform of Hecke operators, then the Fourier coefficients of* $f(z)$ *satisfy the following relations*:

$$a(mn) = a(m)a(n) \ if \ (m,n) = 1 \tag{2}$$

$$a(p^{\alpha+1}) = a(p)a(p^\alpha) - \chi(p)p^{k-1}a(p^{\alpha-1}) \ if \ p \ is \ a \ prime. \tag{3}$$

We set $f(z) = \sum_{m \ge 1} a(m) q^m \in S_k(\Gamma, \chi)$. When $\dim S_k(\Gamma, \chi) = 1$ and $a(1) = 1$, then $f(z)$ is the normalized eigenform of Hecke operators, [11]. So, the coefficients of $f(z)$ have the relations as mentioned in Lemma 1.2. It is known that

$$|a(p)| < 2p^{(k-1)/2} \tag{4}$$

for all primes p, [11, p. 164], [10]. Note that this is the Ramanujan conjecture and its generalization, called the Ramanujan-Petersson conjecture for cusp forms which are eigenforms of the Hecke operators. These conjectures were proved by Deligne as a consequence of his proof of the Weil conjectures, [11, p. 164], [10]. Moreover, for a prime p with $\chi(p) = 1$ the following equation holds, [12].

$$a(p^\alpha) = p^{(k-1)\alpha/2}\frac{\sin(\alpha+1)\theta_p}{\sin\theta_p}, \tag{5}$$

where $2\cos\theta_p = a(p)p^{-(k-1)/2}$ and $\alpha \in \mathbb{N}$.

It is well known that the theta series of $\Lambda \subset \mathbb{R}^n$ weighted by harmonic polynomial $P \in \mathrm{Harm}_j(\mathbb{R}^n)$ is a modular form of weight $n/2 + j$ for some subgroup $\Gamma \subset SL_2(\mathbb{R})$ [8]. In particular, when $\deg(P) \geq 1$, the theta series of Λ weighted by P is a cusp form.

For example, we consider the even unimodular lattice Λ. Then the theta series of Λ weighted by harmonic polynomial P, $\Theta_{\Lambda,P}(z)$, is a modular form with respect to $SL_2(\mathbb{Z})$.

Example 1.1. Let Λ be the E_8-lattice. This is an even unimodular lattice of $\mathbb{R}^8$, generated by the E_8 root system. The theta series is as follows:

$$\begin{aligned}\Theta_\Lambda(z) = E_4(z) &= 1 + 240\sum_{m=1}^{\infty}\sigma_3(m)q^{2m}\\ &= 1 + 240q^2 + 2{,}160q^4 + 6{,}720q^6 + 17{,}520q^8 + \cdots,\end{aligned}$$

where $\sigma_3(m)$ is a divisor function $\sigma_3(m) = \sum_{0<d|m} d^3$.

For $j = 2, 4$ and 6, the theta series of Λ weighted by $P \in \mathrm{Harm}_j(\mathbb{R}^8)$ is a weight $6, 8$ and 10 cusp form with respect to $SL_2(\mathbb{Z})$. However, it is well known that for $k = 6, 8$ and 10, $\dim S_k(SL_2(\mathbb{Z})) = 0$, that is, $\Theta_{\Lambda,P}(z) = 0$. Then by Lemma 1.1, all the nonempty shells of E_8-lattice are spherical 6-design.

For $j = 8$, the theta series of Λ weighted by P is a weight 12 cusp form with respect to $SL_2(\mathbb{Z})$. Such a cusp form is uniquely determined up to constant, i.e., it is Ramanujan's delta function:

$$\Delta(z) = q^2\prod_{m\geq 1}(1-q^{2m})^{24} = \sum_{m\geq 1}\tau(m)q^{2m}.$$

The following proposition is due to Venkov, de la Harpe and Pache [5, 6, 16, 23].

Proposition 1.2 (cf. [16]). *Let the notation be the same as above. Let Λ be the E_8-lattice. Then the following are equivalent*:

(i) $\tau(m) = 0$.
(ii) $(\Lambda)_{2m}$ *is an* 8*-design.*

It is a famous conjecture of Lehmer that $\tau(m) \neq 0$. So, Proposition 1.2 gives a reformulation of Lehmer's conjecture. Lehmer proved in [12] the following theorem.

Theorem 1.1 (cf. [12]). *Let m_0 be the least value of m for which $\tau(m) = 0$. Then m_0 is a prime if it is finite.*

There are many attempts to study Lehmer's conjecture [12, 21], but it is difficult to prove and it is still open.

Recently, however, we showed the "Toy models for D. H. Lehmer's conjecture" [3]. We take the two cases $\mathbb{Z}^2$-lattice and A_2-lattice. Then, we consider the analogue of Lehmer's conjecture corresponding to the theta series weighted by some harmonic polynomial P. Namely, we show that the m-th coefficient of the weighted theta series of $\mathbb{Z}^2$-lattice does not vanish when the shell of norm m of those lattices is not an empty set. Or equivalently, we show the following result.

Theorem 1.2 (cf. [3]). *The nonempty shells in $\mathbb{Z}^2$-lattice (resp. A_2-lattice) are not spherical 4-designs (resp. 6-designs).*

This paper is sequel to the previous paper [3]. In this paper, we take some lattices related to the imaginary quadratic fields. Let $K = \mathbb{Q}(\sqrt{-d})$ be an imaginary quadratic field, and let $\mathcal{O}_K$ be its ring of algebraic integers. Let Cl_K be the ideal classes. In this paper, we only consider the cases $|\mathrm{Cl}_K| = 1$ and $|\mathrm{Cl}_K| = 2$ except for Sect. 6. So, when we consider the cases $|\mathrm{Cl}_K| = 1$ and $|\mathrm{Cl}_K| = 2$, we denote by $\mathfrak{o}$ (resp. $\mathfrak{a}$) the principal (resp. nonprincipal) ideal class.

We denote by d_K the discriminant of K:

$$d_K = \begin{cases} -4d \text{ if } -d \equiv 2,\ 3 \pmod 4, \\ -d \ \text{ if } -d \equiv 1 \pmod 4. \end{cases}$$

Theorem 1.3 (cf. [25, p. 87]). *Let d be a positive square-free integer, and let $K = \mathbb{Q}(\sqrt{-d})$. Then*

$$\mathcal{O}_K = \begin{cases} \mathbb{Z} + \mathbb{Z}\sqrt{-d} & if\ -d \equiv 2,\ 3 \pmod 4, \\ \mathbb{Z} + \mathbb{Z}\dfrac{-1+\sqrt{-d}}{2} & if\ -d \equiv 1 \pmod 4. \end{cases}$$

Therefore, we consider $\mathcal{O}_K$ to be the lattice in $\mathbb{R}^2$ with the basis

$$\begin{cases} (1,0), (1, \sqrt{-d}) & \text{if } -d \equiv 2,\ 3 \pmod 4, \\ (1,0), \left(-\dfrac{1}{2}, \dfrac{\sqrt{-d}}{2}\right) & \text{if } -d \equiv 1 \pmod 4, \end{cases}$$

denoted by $L_{\mathfrak{o}}$.

Generally, it is well-known that there exists one-to-one correspondence between the set of reduced quadratic forms $f(x, y)$ with a fundamental discriminant $d_K < 0$ and the set of fractional ideal classes of the unique quadratic field $\mathbb{Q}(\sqrt{-d})$ [25, p. 94]. Namely, For a fractional ideal $A = \mathbb{Z}\alpha + \mathbb{Z}\beta$, we obtain the quadratic form $ax^2 + bxy + cy^2$, where $a = \alpha\bar{\alpha}/N(A)$, $b = (\alpha\bar{\beta} + \bar{\alpha}\beta)/N(A)$ and $c = \beta\bar{\beta}/N(A)$. Conversely, for a quadratic form $ax^2 + bxy + cy^2$, we obtain the fractional ideal $\mathbb{Z} + \mathbb{Z}(b + \sqrt{d_K})/2a$. We remark that $N(A)$ is a norm of A and $\bar{\alpha}$ is a complex conjugate of α.

Here, we define the automorphism group of $f(x,y)$ as follows:

$$U_f = \left\{ \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in SL_2(\mathbb{Z}) \,\middle|\, f(\alpha x + \beta y, \gamma x + \delta y) = f(x,y) \right\}.$$

Then, for $n \geq 1$, the number of the nonequivalent solutions of $f(x,y) = n$ under the action of U_f is equal to the number of the integral ideals of norm n [25]. Namely, let $\mathfrak{a}$ be an ideal class and $f_{\mathfrak{a}}(x,y)$ be the reduced quadratic form corresponding to $\mathfrak{a}$. Moreover, let $L_{\mathfrak{a}}$ be the lattice corresponding to $f_{\mathfrak{a}}(x,y)$. Then,

$$\begin{aligned} &\sum_{x \in L_{\mathfrak{a}}} q^{(x,x)} \\ &= 1 + \#U_f \textstyle\sum_{n=1}^{\infty} \#\{A \mid A \text{ is an integral ideal of } \mathfrak{a},\ N(A) = n\}\, q^n, \end{aligned} \tag{6}$$

where $N(A)$ is the norm of an ideal A.

Theorem 1.4 (cf. [25, p. 63]). *Let $f(x,y)$ be the reduced quadratic form with a fundamental discriminant $D < 0$ and U_f be the automorphism group of $f(x,y)$. Then*

$$\#U_f = \begin{cases} 6 & \textit{if } D = -3, \\ 4 & \textit{if } D = -4, \\ 2 & \textit{if } D < -4. \end{cases}$$

These classical results are due to Gauss, Dirichlet, etc.

When $|\mathrm{Cl}_K| = 1$ and 2, we give the generators of $L_{\mathfrak{a}}$ Tables 5 and 6 of Appendix. Here, we remark that when $K = \mathbb{Q}(\sqrt{-1})$ (resp. $K = \mathbb{Q}(\sqrt{-3})$), $L_{\mathfrak{o}}$ is $\mathbb{Z}^2$-lattice (resp. A_2-lattice). We studied the spherical designs of shells of those lattices in the previous paper [3].

In this paper, we take the imaginary quadratic fields $\mathbb{Q}(\sqrt{-d})$, with $d \neq 1$ and $d \neq 3$. Then, we consider the analogue of Lehmer's conjecture corresponding to its theta series weighted by some harmonic polynomial P. Here, we consider the following problem of whether the nonempty shells of $L_{\mathfrak{o}}$ and $L_{\mathfrak{a}}$ are spherical 2-designs (hence 3-designs) or not.

In Sect. 4, we study the case that the class number is 1. We show that the m-th coefficient of the weighted theta series of $L_{\mathfrak{o}}$-lattice does not vanish when the shell of norm m of those lattices is not an empty set. Or equivalently, we show the following result:

Theorem 1.5. *Let $K = \mathbb{Q}(\sqrt{-d})$ be an imaginary quadratic field whose class number is 1 and $d \neq 1, 3$ i.e., d is in the following set: $\{2, 7, 11, 19, 43, 67, 163\}$. Then, the nonempty shells in $L_{\mathfrak{o}}$ are not spherical 2-designs.*

Similarly, in Sect. 5, we study the case that the class number is 2 and show the following result:

Theorem 1.6. *Let $K = \mathbb{Q}(\sqrt{-d})$ be an imaginary quadratic field whose class number is 2 i.e., d is in the following set: $\{5, 6, 10, 13, 15, 22, 35, 37, 51, 58, 91,$*

115, 123, 187, 235, 267, 403, 427$\}$. *Then, the nonempty shells in* $L_{\mathfrak{o}}$ *and* $L_{\mathfrak{a}}$ *are not spherical* 2*-designs.*

In Sect. 6, we consider the case that the class number is 3 and study the property of Hecke characters. In Sect. 7, we give some concluding remarks and state a conjecture for the future study.

2 Preliminaries

In this section, we review the theory of imaginary quadratic fields.

Theorem 2.1 (cf. [4, p. 104, Proposition 5.16]). *We can classify the prime ideals of a quadratic field as follows*:

1. *If* p *is an odd prime and* $(d_K/p) = 1$ *(resp.* $d_K \equiv 1 \pmod 8$*) then* $(p){=}P\overline{P}$ *(resp.* $(2) = P\overline{P}$*), where* P *and* $\overline{P}$ *are prime ideals with* $P \neq \overline{P}$*,* $N(P){=}N(\overline{P}) = p$ *(resp.* $N(P) = 2$*).*
2. *If* p *is an odd prime and* $(d_K/p) = -1$ *(resp.* $d_K \equiv 5 \pmod 8$*) then* $(p){=}P$ *(resp.* $(2) = P$*), where* P *is a prime ideal with* $N(P) = p^2$ *(resp.* $N(P){=}4$*).*
3. *If* $p \mid d_k$ *then* $(p) = P^2$*, where* P *is a prime ideal with* $N(P) = p$*.*

Lemma 2.1. *Let* $|\mathrm{Cl}_K| = 1$ *and* I *be an integral ideal of* K*. For* $n \in \mathbb{N}$*, if* $N(I) = n$ *and* I *is a principal ideal, namely,* $I \in \mathfrak{o}$ *then there exist* a*,* $b \in \mathbb{Z}$ *such that for* $-d \equiv 2,\ 3 \pmod 4$

$$n = a^2 + db^2,$$

for $-d \equiv 1 \pmod 4$

$$n = a^2 + db^2 \quad \textit{or} \quad n = \frac{a^2 + db^2}{4}.$$

If $|\mathrm{Cl}_K| = 2$*,* $N(I) = n$ *and* I *is a nonprincipal ideal, namely,* $I \in \mathfrak{a}$ *and assume that* m *is one of the norms of nonprincipal ideals then there exist* a*,* $b \in \mathbb{Z}$ *such that for* $-d \equiv 2,\ 3 \pmod 4$

$$mn = a^2 + db^2,$$

for $-d \equiv 1 \pmod 4$

$$mn = a^2 + db^2 \quad \textit{or} \quad mn = \frac{a^2 + db^2}{4}.$$

Proof. We assume that $|\mathrm{Cl}_K| = 1$. For $-d \equiv 2,\ 3 \pmod 4$, we can write $I = (a + b\sqrt{-d})$, then $N(I) = a^2 + db^2$. For $-d \equiv 1 \pmod 4$, we can write $I = (a + b\sqrt{-d})$ or $I = ((a + b\sqrt{-d})/2)$, then $N(I) = a^2 + db^2$ or $N(I) = (a^2 + db^2)/4$.

Here, we assume that $|\mathrm{Cl}_K| = 2$. Let J be the nonprincipal ideal of K whose norm is m. If I is a nonprincipal ideal then, JI is a principal ideal of K. Therefore, for $-d \equiv 2,\ 3 \pmod 4$, we can write $JI = (a + b\sqrt{-d})$, then $N(JI) = a^2 + db^2$. Hence, $mn = a^2 + db^2$. for $-d \equiv 1 \pmod 4$, we can write $JI = (a + b\sqrt{-d})$ or $JI = ((a + b\sqrt{-d})/2)$, then $N(JI) = a^2 + db^2$ or $N(JI) = (a^2 + db^2)/4$. Hence, $mn = a^2 + db^2$ or $mn = (a^2 + db^2)/4$. □

Proposition 2.1. *Let $F(m)$ be the number of the integral ideals of norm m of K. Let p be a prime number. Then, if $p \neq 2$*

$$F(p^e) = \begin{cases} e+1 & \textit{if } \ (d_K/p) = 1, \\ (1 + (-1)^e)/2 & \textit{if } \ (d_K/p) = -1, \\ 1 & \textit{if } p \mid d_K, \end{cases}$$

if $p = 2$

$$F(2^e) = \begin{cases} e+1 & \textit{if } d_K \equiv 1 \pmod 8, \\ (1 + (-1)^e)/2 & \textit{if } d_K \equiv 5 \pmod 8, \\ 1 & \textit{if } 2 \mid d_K. \end{cases}$$

Proof. When $(d_K/p) = 1$ i.e., $(p) = P\overline{P}$ and $P \neq \overline{P}$, since P and $\overline{P}$ are the only integral ideals of norm p, we have $F(p) = 2$. Moreover, the integral ideals of norm p^e are as follows: P^e, $P^{e-1}\overline{P}$, …, $(\overline{P})^e$. So, we have $F(p^e) = e + 1$. The other cases can be proved similarly. □

3 Hecke Characters and Theta Series

In this section, we introduce the Hecke character and discuss the relationships between the Hecke character and the weighted theta series of the lattices $L_{\mathfrak{o}}$ and $L_{\mathfrak{a}}$. Then, we show that for $|\mathrm{Cl}_K| = 1$ and $P_1 = (x^2 - y^2)/2$, the weighted theta series $\Theta_{L_{\mathfrak{o}},P_1}$ is a normalized Hecke eigenform. For $|\mathrm{Cl}_K| = 2$ and $P_2 = x^2 - y^2$, a certain sum of the two weighted theta series $c_1\Theta_{L_{\mathfrak{o}},P_2} + c_2\Theta_{L_{\mathfrak{a}},P_2}$ is a normalized Hecke eigenform. Later, we give the explicit values of c_1 and c_2.

For the readers convenience we quote from [15] the notion of the Hecke character (for more information the reader is referred to [15]). A Hecke character ϕ of weight $k \geq 2$ with modulus Λ is defined in the following way. Let Λ be a nontrivial ideal in $\mathcal{O}_K$ and let $I(\Lambda)$ denote the group of fractional ideals prime to Λ. A Hecke character ϕ with modulus Λ is a homomorphism

$$\phi : I(\Lambda) \to \mathbb{C}^\times$$

such that for each $\alpha \in K^\times$ with $\alpha \equiv 1 \pmod \Lambda$ we have

$$\phi(\alpha\mathcal{O}_K) = \alpha^{k-1}. \tag{7}$$

Let ω_ϕ be the Dirichlet character with the property that

$$\omega_\phi(n) := \phi((n))/n^{k-1}$$

for every integer n coprime to Λ.

Theorem 3.1 (cf. [15, p. 9], [14, p. 183]). *Let the notation be the same as above, and define $\Psi_{K,\Lambda}(z)$ by*

$$\Psi_{K,\Lambda}(z) := \sum_{A} \phi(A) q^{N(A)} = \sum_{n=1}^{\infty} a(n) q^n, \tag{8}$$

where the sum is over the integral ideals A that are prime to Λ and $N(A)$ is the norm of the ideal A. Then $\Psi_{K,\Lambda}(z)$ is a cusp form in $S_k(\Gamma_0(d_K \cdot N(\Lambda)), \left(\frac{-d_K}{\bullet}\right) \omega_\phi)$.

We remark that function (8) is a normalized Hecke eigenform [1,22]. Moreover, if the class number of K is h then the character as given in (7) will have h extensions to nonprincipal ideals. Namely, the function (8) has h choices, so we denote by $\Psi^{(1)}_{K,\Lambda}(z), \dots, \Psi^{(h)}_{K,\Lambda}(z)$ these functions (see [17]).

Example 3.1.

(i) $d = 2$.
We calculate $\Psi_{K,\Lambda}(z) = \sum_{m \geq 1} a(m) q^m$, where $\Lambda = (1)$ and the weight of the Hecke character is 3. We remark that $|\mathrm{Cl}_K| = 1$ and ideals are listed in Table 3.

By the definitions (7) and (8), we have $a(1) = 1^2 = 1$, $a(2) = \sqrt{-2}^2 = -2$, $a(3) = (-1+\sqrt{-2})^2 + (-1-\sqrt{-2})^2 = 2$, $a(4) = 2^2, \dots$. Thus, we obtain

$$\Psi^{(1)}_{K,\Lambda}(z) = q - 2q^2 - 2q^3 + 4q^4 + 4q^6 - 8q^8 - 5q^9 + \cdots .$$

(ii) $d = 5$.
We calculate $\Psi_{K,\Lambda}(z) = \sum_{m \geq 1} a(m) q^m$, where $\Lambda = (1)$ and the weight of the Hecke character is 3. We remark that $|\mathrm{Cl}_K| = 2$ and ideals are listed in Table 4. When A of norm m is a nonprincipal ideal, A^2 is a principal ideal, so, $\phi(A^2)$ is computable by the definition (7). For example, $\phi((2, 1+\sqrt{-5}))^2 = \phi((2)) = 4$, so, we can assume that $\phi((2, 1+\sqrt{-5})) = 2$, i.e., $a(2) = 2$. Then, since $(2, 1+\sqrt{-5})(3, 1+\sqrt{-5}) = (1-\sqrt{-5})$ and $(2, 1+\sqrt{-5})(3, 1-\sqrt{-5}) = (-1-\sqrt{-5})$, we have $a(3) = ((1+\sqrt{-5})^2 + (1-\sqrt{-5})^2)/2 = -4$, $a(4) = 2^2$, ... Thus, we obtain

$$\Psi^{(1)}_{K,\Lambda}(z) = q + 2q^2 - 4q^3 + 4q^4 - 5q^5 - 8q^6 + 4q^7 + 8q^8 + 7q^9 + \cdots .$$

On the other hand, we assume that $\phi((2, 1+\sqrt{-5})) = -2$, i.e., $a(2) = -2$. Then, we have

$$\Psi^{(2)}_{K,\Lambda}(z) = q - 2q^2 + 4q^3 + 4q^4 - 5q^5 - 8q^6 - 4q^7 - 8q^8 + 7q^9 + \cdots .$$

Table 1 Coefficients, c_1 and c_2

$-d$	-5	-6	-10	-13	-15	-22	-35	-37	-51
c_1	$1/2$	$1/2$	$1/2$	$1/2$	$1/2$	$1/2$	$1/2$	$1/2$	$1/2$
c_2	$1/2$	$1/2$	$1/2$	$1/2$	2	$1/2$	3	$1/2$	$1/2$
$-d$	-58	-91	-115	-123	-187	-235	-267	-403	-427
c_1	$1/2$	$1/2$	$1/2$	$1/2$	$1/2$	$1/2$	$1/2$	$1/2$	$1/2$
c_2	$1/2$	$5/3$	$1/2$	$1/2$	$7/3$	$1/2$	$1/2$	$11/9$	$1/2$

Here, we discuss the relationships between the Hecke character and the weighted theta series of the lattices $L_{\mathfrak{o}}$ and $L_{\mathfrak{a}}$. First, we quote the following theorem:

Theorem 3.2 (cf. [14, p. 192]). *Let L be a 2-dimensional integral lattice with the Gram matrix A and N be the natural number such that the elements of NA^{-1} are rational integers. Let the character $\chi(d)$ be*

$$\chi(d) = \Big(\frac{(-1)^{(r/2)} \det L}{d}\Big).$$

Then, for $P \in \mathrm{Harm}_2(\mathbb{R}^2)$,

(1) $\Theta_{L,P} \in M_3(\Gamma_0(4N), \chi)$.
(2) *If all the diagonal elements of A are even, then $\Theta_{L,P} \in M_3(\Gamma_0(2N), \chi)$.*
(3) *If all the diagonal elements of A and NA^{-1} are even, then $\Theta_{L,P} \in M_3(\Gamma_0(N), \chi)$.*

Then, we obtain the following lemmas:

Lemma 3.1. *Let K be an imaginary quadratic field whose class number is 1 and $L_{\mathfrak{o}}$ be the lattice corresponding to the principal ideal class $\mathfrak{o}$. Let ϕ be the Hecke character of weight 3 with modulus Λ. Assume that $\Lambda = (1)$ and $P_1 = (x^2 - y^2)/2 \in \mathrm{Harm}_2(\mathbb{R}^2)$. Then, $\Psi_{K,\Lambda}(q) = \Theta_{L_{\mathfrak{o}},P_1}(q)$.*

Lemma 3.2. *Let K be an imaginary quadratic field whose class number is 2 and $L_{\mathfrak{o}}$ (resp. $L_{\mathfrak{a}}$) be the lattice corresponding to the principal ideal class $\mathfrak{o}$ (resp. nonprincipal ideal class $\mathfrak{a}$). Let ϕ be the Hecke character of weight 3 with modulus Λ. Assume that $\Lambda = (1)$ and $P_2 = x^2 - y^2 \in \mathrm{Harm}_2(\mathbb{R}^2)$. Then, $\Psi_{K,\Lambda}(q) = c_1\Theta_{L_{\mathfrak{o}},P_2}(q) + c_2\Theta_{L_{\mathfrak{a}},P_2}(q)$, where c_1 and c_2 are given as in Table 1.*

Proof of Lemmas 3.1 and 3.2. First, we assume that the lattices are integral lattices, if not we multiply the Gram matrix of L by 2.

Because of the Theorems 3.1 and 3.2, $\Psi_{K,\Lambda}(q)$, $\Theta_{L_{\mathfrak{o}},P}(q)$ and $\Theta_{L_{\mathfrak{a}},P}(q)$ with $P = P_1$, P_2 are modular forms of the same group Γ. Therefore, we calculate the basis of the space of modular forms of group Γ and check $\Psi_{K,\Lambda}(q) = \Theta_{L_{\mathfrak{o}},P_1}(q)$ and $\Psi_{K,\Lambda}(q) = c_1\Theta_{L_{\mathfrak{o}},P_2}(q) + c_2\Theta_{L_{\mathfrak{a}},P_2}(q)$ explicitly (using "Sage", Mathematics Software [18]). □

Corollary 3.1. *Let the notation be the same as above. If* $|\mathrm{Cl}_K| = 1$ *then* $\Theta_{L_1,P_1}(q)$ *is a normalized Hecke eigenform. If* $|\mathrm{Cl}_K| = 2$ *then* $c_1\Theta_{L_1,P_2}(q) + c_2\Theta_{L_2,P_2}(q)$ *is a normalized Hecke eigenform.*

Proof. The function (8) is a normalized Hecke eigenform [1, 22]. □

Finally, we give the following proposition, which is an analogue of Theorem 1.1 and the crucial part of the proof of Theorems 1.5 and 1.6.

Proposition 3.1. *Assume that* $\sum_{m\geq 1} a(m)q^m$ *is a normalized Hecke eigenform of* $S_3(\Gamma, \chi)$ *and the coefficients* $a(m)$ *are rational integers. Moreover let* p *be the prime such that* $\chi(p) = 1$*. Let* α_0 *be the least value of* α *for which* $a(p^\alpha) = 0$*. If* $a(p) \neq \pm p$ *then* $\alpha_0 = 1$ *if it is finite.*

Proof. Assume the contrary, that is, $\alpha_0 > 1$, so that $a(p) \neq 0$. By the equation (5),

$$a(p^{\alpha_0}) = 0 = p^{\alpha_0}\frac{\sin(\alpha_0 + 1)\theta_p}{\sin\theta_p}.$$

This shows that θ_p is a real number of the form $\theta_p = \pi k/(1 + \alpha_0)$, where k is an integer. Now the number

$$z = 2\cos\theta_p = a(p)p^{-1}, \tag{9}$$

being twice the cosine of a rational multiple of 2π, is an algebraic integer. On the other hand, z is a root of the equation

$$pz - a(p) = 0. \tag{10}$$

Hence z is a rational integer. By (4) and (9), we have $|z| \leq 1$. Therefore $z = \pm 1$ and the equation (10) becomes $a(p) = \pm p$. By assumption, this is a contradiction. □

4 The Case of $|\mathrm{Cl}_K| = 1$

Let $K := \mathbb{Q}(\sqrt{-d})$ be an imaginary quadratic field. If the class number of K is 1 then d is in the following set $\{1, 2, 3, 7, 11, 19, 43, 67, 163\}$. In particular, we only consider the cases where d is in the set: $\{2, 7, 11, 19, 43, 67, 163\}$ since the cases $d = 1$ and $d = 3$ are considered in [3].

In this section, we assume that $a(m)$ and $b(m)$ are the coefficients of the following functions:

$$\Theta_{L_\mathfrak{o}}(q) = \sum_{m\geq 0} a(m)q^m,\ \Theta_{L_\mathfrak{o},P_1}(q) = \sum_{m\geq 1} b(m)q^m,$$

where $P_1 = (x^2 - y^2)/2 \in \mathrm{Harm}_2(\mathbb{R}^2)$.

Lemma 4.1. *Let p be a prime number. Let d be one of the elements in* $\{2, 7, 11, 19, 43, 67, 163\}$. *We set $a'(m) = a(m)/2$ for all m. Then,*

$$a'(p^e) = \begin{cases} e+1 & if\ (d_K/p) = 1, \\ (1+(-1)^e)/2 & if\ (d_K/p) = -1, \\ 1 & if\ p \mid d_K. \end{cases}$$

Proof. Because of the equation (6), $a'(m)$ is the number of integral ideals of K of norm m. Therefore, it can be proved by Proposition 2.1. □

Lemma 4.2. *Let p be a prime number such that $(d_K/p) = 1$. Then, $b(p) \neq 0$. Moreover, if $p \neq d$ then $b(p) \neq \pm p$.*

Proof. We remark that by Lemma 3.1 and Corollary 3.1, $\Theta_{L_o,P_1}(q) = \Psi_{K,\Lambda}(q)$. So, the numbers $b(m)$ are the coefficients of $\Psi_{K,\Lambda}(q)$.

First, we assume that $d \neq 2$, i.e., $-d \equiv 1 \pmod 4$ and $\mathcal{O}_K = \mathbb{Z} + \mathbb{Z}(1 + \sqrt{-d})/2$. If $N((a + b\sqrt{-d}))$ is equal to p then by Lemma 2.1

$$p = a^2 + db^2.$$

Because of the definition of $\Psi_{K,\Lambda}(q)$,

$$b(p) = (a + b\sqrt{-d})^2 + (a - b\sqrt{-d})^2 = 2(a^2 - db^2).$$

If $b(p) = 0$ then $a^2 = db^2$. This is a contradiction. Assume that $b(p) = \pm p$. Then,

$$2(a^2 - db^2) = \pm(a^2 + db^2),$$

that is, $a^2 = 3db^2$ or $3a^2 = db^2$. This is a contradiction.

If $N(((a + b\sqrt{-d})/2))$ is equal to p then by Lemma 2.1

$$\frac{a^2 + db^2}{4} = p.$$

Because of the definition of $\Psi_{K,\Lambda}(q)$,

$$b(p) = \Big(\frac{a + b\sqrt{-d}}{2}\Big)^2 + \Big(\frac{a - b\sqrt{-d}}{2}\Big)^2 = \frac{a^2 - db^2}{2}.$$

If $b(p) = 0$ then $a^2 = db^2$. This is a contradiction. Assume that $b(p) = \pm p$. Then,

$$\frac{a^2 - db^2}{2} = \pm\frac{a^2 + db^2}{4},$$

that is, $a^2 = 3db^2$ or $3a^2 = db^2$. This is a contradiction.

Next, we assume that $d = 2$ i.e., $-d \equiv 2 \pmod 4$ and $\mathcal{O}_K = \mathbb{Z} + \mathbb{Z}\sqrt{-2}$. If $N((a + b\sqrt{-2}))$ is equal to p then by Lemma 2.1

$$p = a^2 + 2b^2.$$

Because of the definition of $\Psi_{K,\Lambda}(q)$,

$$b(p) = (a + b\sqrt{-2})^2 + (a - b\sqrt{-2})^2 = 2(a^2 - 2b^2).$$

If $b(p) = 0$ then $a^2 = 2b^2$. This is a contradiction. Assume that $b(p) = \pm p$. Then,

$$2(a^2 - 2b^2) = \pm(a^2 + 2b^2),$$

that is, $a^2 = 6b^2$ or $3a^2 = 2b^2$. This is a contradiction. □

Proof of Theorem 1.5. We will show that $b(m) \neq 0$ when $(L_{\mathfrak{o}})_m \neq \emptyset$.

By Theorem 3.1, $\Theta_{L_{\mathfrak{o}},P_1}$ is a normalized Hecke eigenform. So, We assume that m is a power of prime, if not we could apply the equation (2). We will divide our considerations into the following three cases.

(i) Case $m = p^\alpha$ and $p \mid d_K$:
By $a(m) = 2$ and the inequality (1), the shells $(L_{\mathfrak{o}})_m$ are not spherical 2-designs. Hence, $b(m) \neq 0$.

(ii) Case $m = p^\alpha$ and $(d_K/p) = -1$:
By Lemma 4.1,

$$a(p^n) = \begin{cases} 0 & \text{if } n \text{ is odd,} \\ 2 & \text{if } n \text{ is even.} \end{cases}$$

By $a(m) = 2$ and the inequality (1), when n is even, the shells $(L_{\mathfrak{o}})_m$ are not spherical 2-designs. Hence, $b(m) \neq 0$.

(iii) Case $m = p^\alpha$ and $(d_K/p) = 1$:
By Proposition 3.1 and Lemma 4.2, we have $b(m) \neq 0$. This completes the proof of Theorem 1.5. □

5 The Case of $|\mathrm{Cl}_K| = 2$

Let $K := \mathbb{Q}(\sqrt{-d})$ be an imaginary quadratic field. In this section, we assume that the class number of K is 2. So, we consider that d is in the following set: {5, 6, 10, 13, 15, 22, 35, 37, 51, 58, 91, 115, 123, 187, 235, 267, 403, 427}. We denote by $\mathfrak{o}$ (resp. $\mathfrak{a}$) the principal (resp. nonprincipal) ideal class.

In this section, we also assume that $a(m)$ and $b(m)$ are the coefficients of the following functions:

$$\Theta_{L_{\mathfrak{o}}}(q) + \Theta_{L_{\mathfrak{a}}}(q) = \sum_{m \geq 0} a(m)q^m,$$

$$c_1\Theta_{L_{\mathfrak{o}},P_2}(q) + c_2\Theta_{L_{\mathfrak{a}},P_2}(q) = \sum_{m \geq 1} b(m)q^m,$$

where c_1 and c_2 are defined in Lemma 3.2.

Table 2 Values of m and $b(m)$

$-d$	-5	-6	-10	-13	-15	-22	-35	-37	-51
m	2	2	2	2	3	2	5	2	3
$b(m)$	2	2	2	2	-3	2	-5	2	3
$-d$	-58	-91	-115	-123	-187	-235	-267	-403	-427
m	2	7	5	3	11	5	3	13	7
$b(m)$	2	-7	-5	3	-11	5	3	-13	7

Lemma 5.1. *Set* $l_1 := \{N(O) \mid O \in \mathfrak{o}\}$ *and* $l_2 := \{N(A) \mid A \in \mathfrak{a}\}$. *Then,* $l_1 \cap l_2 = \emptyset$. *Therefore, the set* $L_{\mathfrak{o}} \cap L_{\mathfrak{a}}$ *consists of the origin.*

Proof. Let p be a prime number such that $(d_K/p) = 1$. Then there exist prime ideals P and P' such that $(p) = PP'$ and $N(P) = N(P') = p$. Since a class number is 2, we have P and $P' \in \mathfrak{o}$ or P and $P' \in \mathfrak{a}$. If P and $P' \in \mathfrak{o}$ we denote by p_i such a prime. If P and $P' \in \mathfrak{a}$ we denote by p'_i such a prime.

Let p be a prime number such that $(d_K/p) = -1$. Then (p) is a prime ideal and $N((p)) = p^2$. We denote by q_i such a prime.

Let p be a prime number such that $p \mid d_K$. Then there exists a prime ideal P such that $(p) = P^2$ and $N(P) = p$. Since a class number is 2, we have $P \in \mathfrak{o}$ or $P \in \mathfrak{a}$. If $P \in \mathfrak{o}$ we denote by r_i such a prime. If $P \in \mathfrak{a}$ we denote by r'_i such a prime.

We take the element $n \in l_1 \cap l_2$ and perform a prime factorization, $n = p_1 \cdots p'_1 \cdots q_1 \cdots r_1 \cdots r'_1 \cdots$. Then, $p_1 \cdots$, $q_1 \cdots$ and $r_1 \cdots$ correspond to principal ideals. So, if the number of occurrences of each of the primes p' and r' is even then $n \in l_1$ and if the number of occurrences of each of the primes p' and r' is odd then $n \in l_2$. This completes the proof of Lemma 5.1. □

Lemma 5.2. *Let p be a prime number. Let d be one of the elements in* $\{5, 6, 10, 13, 15, 22, 35, 37, 51, 58, 91, 115, 123, 187, 235, 267, 403, 427\}$. *We set* $a'(m) = a(m)/2$ *for all m. Then,*

$$a'(p^e) = \begin{cases} e+1 & \textit{if } \ (d_K/p) = 1, \\ (1+(-1)^e)/2 & \textit{if } \ (d_K/p) = -1, \\ 1 & \textit{if } p \mid d_K. \end{cases}$$

Proof. Because of the equation (6), $a'(m)$ is the number of integral ideals of K of norm m. Therefore, it can be proved by Proposition 2.1. □

Lemma 5.3. *Let p be a prime number such that* $(d_K/p) = 1$. *Then,* $b(p) \neq 0$. *Moreover, if* $p \neq d$ *then* $b(p) \neq \pm p$.

Proof. We remark that by Lemma 3.2 and Corollary 3.1, $c_1\Theta_{L_{\mathfrak{o}},P_2}(q) + c_2\Theta_{L_{\mathfrak{a}},P_2}(q) = \Psi_{K,\Lambda}(q)$. So, the numbers $b(m)$ are the coefficients of $\Psi_{K,\Lambda}(q)$.

We set $N(J) = p$. When J is a principal ideal, it can be proved by the similar method in Lemma 4.2. So, we assume that J is nonprincipal.

We list the smallest prime number m such that $m \mid d_K$ and $m \in \{N(I) \mid I \in \mathfrak{a}\}$, and the values $b(m)$ are in Table 2.

First, we assume that $-d \equiv 2$ or $3 \pmod 4$. If $N(J)$ is equal to p then by Lemma 2.1

$$mp = a^2 + db^2.$$

Because of the definition of $\Psi_{K,\Lambda}(q)$,

$$b(mp) = (a + b\sqrt{-d})^2 + (a - b\sqrt{-d})^2 = 2(a^2 - db^2).$$

Since $b(mp) = b(m)b(p)$ and the value of $b(m)$ in Table 2, we have $b(p) = a^2 - db^2$. If $b(p) = 0$ then $a^2 = db^2$. This is a contradiction. Assume that $b(p) = \pm p$. Then,

$$a^2 - db^2 = \pm \frac{a^2 + db^2}{2},$$

that is, $a^2 = 3db^2$ or $3a^2 = db^2$. This is a contradiction.

Next, we assume that $-d \equiv 1 \pmod 4$. If $N(J)$ is equal to p then by Lemma 2.1 there exist a, $b \in \mathbb{Z}$ such that

$$mp = a^2 + db^2 \quad \textit{or} \quad mp = \frac{a^2 + db^2}{4}.$$

Because of the definition of $\Psi_{K,\Lambda}(q)$,

$$b(mp) = (a + b\sqrt{-d})^2 + (a - b\sqrt{-d})^2 = 2(a^2 - db^2).$$

or

$$b(mp) = \Big(\frac{a + b\sqrt{-d}}{2}\Big)^2 + \Big(\frac{a - b\sqrt{-d}}{2}\Big)^2 = \frac{a^2 - db^2}{2}.$$

Since $b(mp) = b(m)b(p)$ and the value of $b(m)$ in Table 2, we have $b(p) = 2/b(m) \times (a^2 - db^2)$ or $b(p) = 1/b(m) \times (a^2 - db^2)/2$. If $b(p) = 0$ then $a^2 = db^2$. This is a contradiction. Assume that $b(p) = \pm p$. Then,

$$\frac{2(a^2 - db^2)}{b(m)} = \pm \frac{a^2 + db^2}{m},$$

or

$$\frac{a^2 - db^2}{2b(m)} = \pm \frac{a^2 + db^2}{4m},$$

that is, $a^2 = 3db^2$ or $3a^2 = db^2$ since $m = \pm b(m)$ for $-d \equiv 1 \pmod 4$. This is a contradiction. □

Proof of Theorem 1.6. Because of Lemma 5.1, it is enough to show that $b(m) \neq 0$ when $(L_{\mathfrak{o}})_m \neq \emptyset$ or $(L_{\mathfrak{a}})_m \neq \emptyset$.

By Theorem 3.1, $c_1\Theta_{L_{\mathfrak{o}},P_2}+c_2\Theta_{L_{\mathfrak{a}},P_2}$ is a normalized Hecke eigenform. So, We assume that m is a power of prime, if not we could apply the equation (2). We will divide into the three cases.

(i) Case $m=p^{\alpha}$ and $p \mid d_K$:
By $a(m)=2$ and (1), the shells $(L)_m$ are not spherical 2-designs. Hence, $b(m)\neq 0$.

(ii) Case $m=p^{\alpha}$ and $(d_K/p)=-1$:
By Lemma (4.1),

$$a(p^n)=\begin{cases}0 & \text{if } n \text{ is odd,}\\ 2 & \text{if } n \text{ is even.}\end{cases}$$

By $a(m)=2$ and (1), when n is even, the shells $(L)_m$ are not spherical 2-designs. Hence, $b(m)\neq 0$.

(iii) Case $m=p^{\alpha}$ and $(d_K/p)=1$:
By Proposition 3.1 and Lemma 5.3, $b(m)\neq 0$. This completes the proof of Theorem 1.6. □

6 The Case of $|\mathrm{Cl}_K|=3$

In the previous sections, we studied the cases of class number $h=|\mathrm{Cl}_K|$ is either 1 or 2. However, it seems that the situation is somewhat different for the cases of class numbers $h\geq 3$. In this section, we discuss briefly how it is different, by considering the case of $d=23$ $(h=3)$.

We first remark that one reason of success for the cases $h=1$ and $h=2$ is that the coefficients $a(m)$ of the Hecke eigenform $\Psi_{K,\Lambda}$ are all integers. Therefore, by formula (10), we have that $z=a(p)/p$ is a rational number (and since it is an algebraic integer), and so it must be a rational integer. It seems that this situation is no longer true in general for the cases of $h\geq 3$. We will give more detailed information, concentrating the special (and typical) case of $d=23$.

We denote by $\mathfrak{o}$, $\mathfrak{a}_1$ and $\mathfrak{a}_2$ the ideal classes. The corresponding quadratic forms are $x^2+xy+6y^2$, $2x^2-xy+3y^2$ and $2x^2+xy+3y^2$, namely, $L_{\mathfrak{o}}=\langle(1,0),(1/2,\sqrt{23}/2)\rangle$, $L_{\mathfrak{a}_1}=\langle(2,0),(1/2,\sqrt{23}/2)\rangle$ and $L_{\mathfrak{a}_2}=\langle(2,0),(-1/2,\sqrt{23}/2)\rangle$, respectively. We give the weighted theta series of those ideal lattices. We set $P_1=x^2-y^2$ and $P_2=xy$ in this section.

$\Theta_{L_{\mathfrak{o}}}=1+2q+2q^4+4q^6+4q^8+2q^9+4q^{12}+2q^{16}+4q^{18}+2q^{23}+4q^{24}+2q^{25}+4q^{26}+4q^{27}+4q^{32}+6q^{36}+4q^{39}+8q^{48}+2q^{49}+4q^{52}+4q^{54}+4q^{58}+4q^{59}+4q^{62}+6q^{64}+8q^{72}+4q^{78}+2q^{81}+4q^{82}+4q^{87}+2q^{92}+4q^{93}+4q^{94}+8q^{96}+2q^{100}+O[q]^{101}$.

$\frac{1}{2}\times\Theta_{L_{\mathfrak{o}},P_1}=q+4q^4-11q^6-7q^8+9q^9+q^{12}+16q^{16}+13q^{18}-23q^{23}-44q^{24}+25q^{25}+29q^{26}-38q^{27}-28q^{32}+85q^{36}-14q^{39}+77q^{48}+49q^{49}-103q^{52}-99q^{54}-91q^{58}+26q^{59}+101q^{62}-15q^{64}-11q^{72}+133q^{78}+81q^{81}-43q^{82}+82q^{87}-92q^{92}-182q^{93}-19q^{94}-7q^{96}+100q^{100}+O[q]^{101}$.

$\Theta_{L_{\mathfrak{o}},P_2}=0$.

$\Theta_{L_{\mathfrak{a}_1}} = 1+2q^2+2q^3+2q^4+2q^6+2q^8+2q^9+4q^{12}+2q^{13}+4q^{16}+4q^{18}+6q^{24}+2q^{26}+2q^{27}+2q^{29}+2q^{31}+4q^{32}+6q^{36}+2q^{39}+2q^{41}+2q^{46}+2q^{47}+6q^{48}+2q^{50}+4q^{52}+6q^{54}+2q^{58}+2q^{62}+4q^{64}+2q^{69}+2q^{71}+8q^{72}+2q^{73}+2q^{75}+6q^{78}+4q^{81}+2q^{82}+2q^{87}+2q^{92}+2q^{93}+2q^{94}+8q^{96}+2q^{98}+2q^{100}+O[q]^{101}$.
$2 \times \Theta_{L_{\mathfrak{a}_1},P_1} = 8q^2-11q^3-7q^4+q^6+32q^8+13q^9-88q^{12}+29q^{13}-56q^{16}+121q^{18}+81q^{24}-103q^{26}-99q^{27}-91q^{29}+101q^{31}+49q^{32}+41q^{36}+133q^{39}-43q^{41}-184q^{46}-19q^{47}-183q^{48}+200q^{50}+232q^{52}-295q^{54}+209q^{58}+41q^{62}-224q^{64}+253q^{69}+77q^{71}+393q^{72}-283q^{73}-275q^{75}-375q^{78}+418q^{81}-247q^{82}-227q^{87}+161q^{92}-203q^{93}+353q^{94}+616q^{96}+392q^{98}-175q^{100}+O[q]^{101}$.
$\frac{4}{\sqrt{23}} \times \Theta_{L_{\mathfrak{a}_1},P_2} = q^3-3q^4+5q^6-7q^9+9q^{13}-11q^{18}+13q^{24}-3q^{26}+9q^{27}-15q^{29}-15q^{31}+21q^{32}-27q^{36}+17q^{39}+33q^{41}-39q^{47}-19q^{48}+45q^{54}+21q^{58}-51q^{62}-23q^{69}+57q^{71}+5q^{72}-15q^{73}+25q^{75}-35q^{78}-38q^{81}+45q^{82}-55q^{87}+69q^{92}+65q^{93}-27q^{94}-75q^{100}+O[q]^{101}$.
$\Theta_{L_{\mathfrak{a}_2}} = 1+2q^2+2q^3+2q^4+2q^6+2q^8+2q^9+4q^{12}+2q^{13}+4q^{16}+4q^{18}+6q^{24}+2q^{26}+2q^{27}+2q^{29}+2q^{31}+4q^{32}+6q^{36}+2q^{39}+2q^{41}+2q^{46}+2q^{47}+6q^{48}+2q^{50}+4q^{52}+6q^{54}+2q^{58}+2q^{62}+4q^{64}+2q^{69}+2q^{71}+8q^{72}+2q^{73}+2q^{75}+6q^{78}+4q^{81}+2q^{82}+2q^{87}+2q^{92}+2q^{93}+2q^{94}+8q^{96}+2q^{98}+2q^{100}+O[q]^{101}$.
$2 \times \Theta_{L_{\mathfrak{a}_2},P_1} = 8q^2-11q^3-7q^4+q^6+32q^8+13q^9-88q^{12}+29q^{13}-56q^{16}+121q^{18}+81q^{24}-103q^{26}-99q^{27}-91q^{29}+101q^{31}+49q^{32}+41q^{36}+133q^{39}-43q^{41}-184q^{46}-19q^{47}-183q^{48}+200q^{50}+232q^{52}-295q^{54}+209q^{58}+41q^{62}-224q^{64}+253q^{69}+77q^{71}+393q^{72}-283q^{73}-275q^{75}-375q^{78}+418q^{81}-247q^{82}-227q^{87}+161q^{92}-203q^{93}+353q^{94}+616q^{96}+392q^{98}-175q^{100}+O[q]^{101}$.
$\frac{4}{\sqrt{23}} \times \Theta_{L_{\mathfrak{a}_2},P_2} = -q^3+3q^4-5q^6+7q^9-9q^{13}+11q^{18}-13q^{24}+3q^{26}-9q^{27}+15q^{29}+15q^{31}-21q^{32}+27q^{36}-17q^{39}-33q^{41}+39q^{47}+19q^{48}-45q^{54}-21q^{58}+51q^{62}+23q^{69}-57q^{71}-5q^{72}+15q^{73}-25q^{75}+35q^{78}+38q^{81}-45q^{82}+55q^{87}-69q^{92}-65q^{93}+27q^{94}+75q^{100}+O[q]^{101}$.

We calculate the Hecke character of weight 3 and modulus (1), i.e, we calculate $\Psi_{K,\Lambda} = \sum_{m \geq 1} a(m)q^m$, where $\Lambda = (1)$ and $k = 3$. When A of norm m is a nonprincipal ideal, A^3 is a principal ideal. Then we set $\phi(A)^3 = \phi(A^3)$. For example, $(2, -1/2 + \sqrt{-23}/2)^3 = (-3/2 - \sqrt{-23}/2)$. Because of

$$\phi\Big(\Big(\frac{-3-\sqrt{-23}}{2}\Big)\Big) = \Big(\frac{-3-\sqrt{-23}}{2}\Big)^2 = \frac{-7+3\sqrt{-23}}{2},$$

$\phi((2, -1/2 + \sqrt{-23}/2))$ is one of the roots of

$$x^3 - \Big(\frac{-7+3\sqrt{-23}}{2}\Big) = 0. \tag{11}$$

We denote by α_1, α_2 and α_3 the roots of equation (11), namely, $\alpha_1 \sim -1.86272 + 0.728188i$, $\alpha_2 \sim 0.300733 - 1.97726i$ and $\alpha_3 \sim 1.56199 + 1.24907i$, respectively. Then, $\phi((2, -1/2 + \sqrt{-23}/2))$ is one of α_1, α_2 or α_3. (Actually there are three different Hecke characters in this case.) First let us set $\phi((2, -1/2 + \sqrt{-23}/2)) = \alpha_1$. By the equation $(2, -1/2 + \sqrt{-23}/2) \times (2, 1/2 + \sqrt{-23}/2) = (2)$,

$$\phi\Big(\Big(2, \frac{-1+\sqrt{-23}}{2}\Big)\Big) \times \phi\Big(\Big(2, \frac{1+\sqrt{-23}}{2}\Big)\Big) = \phi((2)).$$

We get

$$\alpha_1 \times \phi\Big(\Big(2, \frac{1+\sqrt{-23}}{2}\Big)\Big) = 4,$$

hence, $\phi((2, 1/2 + \sqrt{-23}/2)) = 4/\alpha_1$. So,

$$a(2) = \phi\Big(\Big(2, \frac{-1+\sqrt{-23}}{2}\Big)\Big) + \phi\Big(\Big(2, \frac{1+\sqrt{-23}}{2}\Big)\Big) = \alpha_1 + 4/\alpha_1.$$

By the equation $(2, -1/2+\sqrt{-23}/2) \times (3, 1/2-\sqrt{-23}/2) = (1/2-\sqrt{-23}/2)$,

$$\phi\Big(\Big(2, \frac{-1+\sqrt{-23}}{2}\Big)\Big) \times \phi\Big(\Big(3, \frac{1-\sqrt{-23}}{2}\Big)\Big) = \phi\Big(\Big(\frac{1-\sqrt{-23}}{2}\Big)\Big).$$

We get

$$\alpha_1 \times \phi\Big(\Big(3, \frac{1-\sqrt{-23}}{2}\Big)\Big) = \Big(\frac{1-\sqrt{-23}}{2}\Big)^2 = \frac{-11-\sqrt{-23}}{2},$$

hence, $\phi((3, 1/2-\sqrt{-23}/2)) = (-11-\sqrt{-23})/2 \times 1/\alpha_1$. Similarly, $\phi((3, -1/2-\sqrt{-23}/2)) = (-11+\sqrt{-23})/2 \times \alpha_1/(\alpha_1^2+4)$. So,

$$\begin{aligned} a(3) &= \phi\Big(\Big(3, \frac{1-\sqrt{-23}}{2}\Big)\Big) + \phi\Big(\Big(3, \frac{-1-\sqrt{-23}}{2}\Big)\Big) \\ &= \frac{-11-\sqrt{-23}}{2} \times \frac{1}{\alpha_1} + \frac{-11+\sqrt{-23}}{2} \times \frac{\alpha_1}{\alpha_1^2+4}. \end{aligned}$$

We recall $\alpha_1 \sim -1.86272 + 0.728188i$. Then, we obtain

$$\Psi^{(1)}_{K,\Lambda} = q - 3.72545q^2 + 4.24943q^3 + \cdots .$$

Actually, it is possible to continue this calculation, but we need the information on the basis of all the ideals, which is rather complicated. So, we determine the Hecke eigenforms $\Psi^{(i)}_{K,\Lambda}$ by a different method. By computer calculation (using "Sage" [18]), we know that the space of the modular forms of weight 3 where $\Psi_{K,\Lambda}$ belongs is of dimension 3. We can calculate the basis of this modular form explicitly, and the three basis elements are of the form:

$$\begin{aligned} &q + 4q^4 - 11q^6 - 7q^8 + 9q^9 + \cdots , \\ &q^2 - 5q^4 + 7q^6 + 4q^8 - 8q^9 + \cdots , \\ &q^3 - 3q^4 + 5q^6 - 7q^9 + \cdots . \end{aligned}$$

On the other hand, because of Theorems 3.1 and 3.2, $\Theta_{L_{\mathfrak{o}},P_1}$, $\Theta_{L_{\mathfrak{a}_1},P_1}$ and $\Theta_{L_{\mathfrak{a}_2},P_2}$ are in the same space of Hecke eigenforms $\Psi_{K,\Lambda}^{(i)}$. Therefore, comparing the first three coefficients of the following equation:

$$\Psi_{K,\Lambda}^{(1)}(q) = \frac{1}{2}\Theta_{L_{\mathfrak{o}},P}(q) + a2\Theta_{L_{\mathfrak{a}_1},P}(q) + b\frac{4}{\sqrt{23}}\Theta_{L_{\mathfrak{a}_2},P}(q),$$

we can find numbers a and b as follows:

$$(a,b) = \begin{cases} (A_1, B_2), \\ (A_2, B_1), \\ (A_3, B_3), \end{cases}$$

where A_1, A_2 and A_3 are the elements defined by

$$\begin{aligned} &\{x \mid 512x^3 - 96x + 7 = 0\} \\ &\qquad = \{A_1 = -0.465681, A_2 = 0.0751832, A_2 = 0.390498\}, \end{aligned}$$

respectively, and B_1, B_2 and B_3 are the elements defined by

$$\begin{aligned} &\{x \mid 512x^3 - 2208x + 1587 = 0\} \\ &\qquad = \{B_1 = -2.37065, B_2 = 0.873067, B_3 = 1.49759\}, \end{aligned}$$

respectively.

In this way, we can calculate the Hecke eigenforms $\Psi_{K,\Lambda}^{(i)}$. Namely, $\Psi_{K,\Lambda}^{(1)} = q - 3.72545q^2 + 4.24943q^3 + 9.87897q^4 - 15.831q^6 - 21.9018q^8 + 9.05761q^9 + 41.9799q^{12} - 21.3624q^{13} + 42.0781q^{16} - 33.7437q^{18} - 23q^{23} - 93.07q^{24} + 25q^{25} + 79.5844q^{26} + 0.244826q^{27} + 55.473q^{29} - 33.9378q^{31} - 69.1528q^{32} + 89.4799q^{36} - 90.7777q^{39} - 8.78692q^{41} + 85.6853q^{46} + 42.8975q^{47} + 178.808q^{48} + 49q^{49} - 93.1362q^{50} + O[q]^{51}$.

$\Psi_{K,\Lambda}^{(2)} = q + 0.601466q^2 + 1.54364q^3 - 3.63824q^4 + 0.928445q^6 - 4.59414q^8 - 6.61718q^9 - 5.61612q^{12} + 23.5162q^{13} + 11.7897q^{16} - 3.98001q^{18} - 23q^{23} - 7.09168q^{24} + 25q^{25} + 14.1442q^{26} - 24.1073q^{27} - 42.4015q^{29} - 27.9663q^{31} + 25.4677q^{32} + 24.0749q^{36} + 36.3005q^{39} + 74.9986q^{41} - 13.8337q^{46} - 93.8839q^{47} + 18.1991q^{48} + 49q^{49} + 15.0366q^{50} + O[q]^{51}$.

$\Psi_{K,\Lambda}^{(3)} = q + 3.12398q^2 - 5.79306q^3 + 5.75927q^4 - 18.0974q^6 + 5.49593q^8 + 24.5596q^9 - 33.3638q^{12} - 2.15383q^{13} - 5.86788q^{16} + 76.7237q^{18} - 23q^{23} - 31.8383q^{24} + 25q^{25} - 6.72853q^{26} - 90.1376q^{27} - 13.0715q^{29} + 61.9041q^{31} - 40.3149q^{32} + 141.445q^{36} + 12.4773q^{39} - 66.2117q^{41} - 71.8516q^{46} + 50.9864q^{47} + 33.993q^{48} + 49q^{49} + 78.0996q^{50} + O[q]^{51}$.

The coefficients $a(m)$ for this case are far from integers. In fact they are not elements in a cyclotomic number field in general. So, it seems difficult to use the

Hecke eigenforms obtained this way to apply for the case of the class number 3 or more in general. Some new additional ideas will be needed to treat the case of $d = 23$ or more generally the cases of class numbers $h \geq 3$. We have included the presentation of the results (although they are not conclusive) for $d = 23$, hoping that it might help the reader for the future study on this topic.

Remark 6.1. We remark that the coefficients of $\Psi_{K,\Lambda}^{(i)}$ in above calculator results are not exact values but approximate values.

7 Concluding Remarks

1. In this paper, we use the mathematics software "Sage" [18]. In particular, the results in Tables 3 and 4 are compute by "Sage" using the command "K.ideals_of_bdd_norm()". We remark that this command does not always give a $\mathbb{Z}$-basis of an ideal. We must use the command "(ideal).basis()".
2. In Appendix C, we list theta series of lattices obtained from $\mathbb{Q}(\sqrt{-5})$. The other cases are listed on one of the author's website [13].

Table 3 Integral ideals of small norm of $d = 2$ and $d = 5$

$N(A)$	A: ideal
1	(1)
2	$(\sqrt{-2})$
3	$(-1+\sqrt{-2})$
	$(-1-\sqrt{-2})$
4	(2)

$N(A)$	A: ideal
1	(1)
2	$(2, 1+\sqrt{-5})$
3	$(3, 1+\sqrt{-5})$
	$(3, 1-\sqrt{-5})$
4	(2)
5	$(\sqrt{-5})$
6	$(1-\sqrt{-5})$
	$(-1-\sqrt{-5})$

Table 4 Integral ideals of small norm of $d = 23$

$N(A)$	A: ideal
1	(1)
2	$(2, -1/2+\sqrt{-23}/2)$
	$(2, 1/2+\sqrt{-23}/2)$
3	$(3, 1/2-\sqrt{-23}/2)$
	$(3, -1/2-\sqrt{-23}/2)$
4	$(4, 3/2+\sqrt{-23}/2)$
	(2)
	$(4, 5/2+\sqrt{-23}/2)$
5	–
6	$(1/2-\sqrt{-23}/2)$
	$(6, 5/2+\sqrt{-23}/2)$
	$(6, 7/2+\sqrt{-23}/2)$
	$(1/2+\sqrt{-23}/2)$
7	–
8	$(-3/2-\sqrt{-23}/2)$
	$(4, -1+\sqrt{-23})$
	$(4, 1+\sqrt{-23})$
	$(-3/2+\sqrt{-23}/2)$
9	$(9, 11/2+\sqrt{-23}/2)$
	(3)
	$(9, 7/2+\sqrt{-23}/2)$
10	–

3. In the previous paper [3], we studied the spherical designs in the nonempty shells of the $\mathbb{Z}^2$-lattice and A_2-lattice. The results state that any shells in the $\mathbb{Z}^2$-lattice (resp. A_2-lattice) are spherical 2-designs (resp. 4-designs). However, the nonempty shells in the $\mathbb{Z}^2$-lattice (resp. A_2-lattice) are not spherical 4-designs (resp. 6-designs). It is interesting to note that no spherical 6-designs among the nonempty shells of any Euclidean lattice of 2-dimensions is known. It is an interesting open problem to prove or disprove whether these exist any 6-designs which is a shell of a Euclidean lattice of 2-dimensions.

   Responding to the authors' request, Junichi Shigezumi performed computer calculations to determine whether there are spherical t-designs for bigger t, in the 2- and 3-dimensional cases. His calculation shows that among the nonempty shells of integral lattices in 2-dimensions (with relatively small discriminant and small norms), there are only 4-designs. That is, no 6-designs were found. (So far, all examples of such 4-designs are the union of vertices of regular 6-gons, although they are the nonempty shells of many different lattices). In the 3-dimensional case, all examples obtained are only 2-designs. No 4-designs which are shells of a lattice were found. It is an interesting open problem whether this is true in general for the dimensions 2 and 3. Moreover, it is interesting to note that no spherical 12-designs among the nonempty shells of any Euclidean lattice (of any dimensions) is known. It is also an interesting open problem to prove or disprove whether these exist any 12-designs which is a shell of a Euclidean lattice.

   Finally, we state the following conjecture for the 2-dimensional lattices.

Conjecture 7.1. *Let L be the Euclidean lattice of 2-dimensions, whose quadratic form is $ax^2 + bxy + cy^2$.*

(i) *Assume that $b^2 - 4ac = (\textit{Integer})^2 \times (-3)$. Then, all the nonempty shells of L are not spherical 6-designs and some of the nonempty shells of L are spherical 4-designs. Moreover, if all the nonempty shells of L are spherical 4-designs then*
$b^2 - 4ac = -3$, that is, A_2-lattice.

(ii) *Assume that $b^2 - 4ac = (\textit{Integer})^2 \times (-4)$. Then, all the nonempty shells of L are not spherical 4-designs and some of the nonempty shells of L are spherical 2-designs. Moreover, if all the nonempty shells of L are spherical 2-designs then*
$b^2 - 4ac = -4$, that is, $\mathbb{Z}^2$-lattice.

(iii) *Otherwise, all the nonempty shells of L are not spherical 2-designs.*

A The Case of $|\mathrm{Cl}_K| = 1$

Table 5 $|\mathrm{Cl}_K| = 1$

$-d$	$-d \pmod 4$	d_K	$L_{\mathfrak{o}}$
-1	3	-2^2	$[1, \sqrt{-1}]$
-2	2	-2^3	$[1, \sqrt{-2}]$
-3	1	-3	$[1, (1 + \sqrt{-3})/2]$
-7	1	-7	$[1, (1 + \sqrt{-7})/2]$
-11	1	-11	$[1, (1 + \sqrt{-11})/2]$
-19	1	-19	$[1, (1 + \sqrt{-19})/2]$
-43	1	-43	$[1, (1 + \sqrt{-43})/2]$
-67	1	-67	$[1, (1 + \sqrt{-67})/2]$
-163	1	-163	$[1, (1 + \sqrt{-163})/2]$

B The Case of $|\mathrm{Cl}_K| = 2$

Table 6 $|\mathrm{Cl}_K| = 2$

$-d$	$-d \pmod 4$	d_K	$L_{\mathfrak{o}}$	$L_{\mathfrak{a}}$
-5	3	$-2^2 \times 5$	$[1, \sqrt{-5}]$	$[2, 1 + \sqrt{-5}]$
-6	2	$-2^3 \times 3$	$[1, \sqrt{-6}]$	$[2, \sqrt{-6}]$
-10	2	$-2^3 \times 5$	$[1, \sqrt{-10}]$	$[2, \sqrt{-10}]$
-13	3	$-2^2 \times 13$	$[1, \sqrt{-13}]$	$[2, 1 + \sqrt{-13}]$
-15	1	-3×5	$[1, (1 + \sqrt{-15})/2]$	$[2, (1 + \sqrt{-15})/2]$
-22	2	$-2^3 \times 11$	$[1, \sqrt{-22}]$	$[2, \sqrt{-22}]$
-35	1	-5×7	$[1, (1 + \sqrt{-35})/2]$	$[3, (1 + \sqrt{-35})/2]$
-37	3	$-2^2 \times 37$	$[1, \sqrt{-37}]$	$[2, 1 + \sqrt{-37}]$
-51	1	-3×17	$[1, (1 + \sqrt{-51})/2]$	$[3, (3 + \sqrt{-51})/2]$
-58	2	$-2^3 \times 29$	$[1, \sqrt{-58}]$	$[2, \sqrt{-58}]$
-91	1	-7×13	$[1, (1 + \sqrt{-91})/2]$	$[5, (3 + \sqrt{-91})/2]$
-115	1	-5×23	$[1, (1 + \sqrt{-115})/2]$	$[5, (5 + \sqrt{-115})/2]$
-123	1	-3×41	$[1, (1 + \sqrt{-123})/2]$	$[3, (3 + \sqrt{-123})/2]$
-187	1	-11×17	$[1, (1 + \sqrt{-187})/2]$	$[7, (3 + \sqrt{-187})/2]$
-235	1	-5×47	$[1, (1 + \sqrt{-235})/2]$	$[5, (5 + \sqrt{-235})/2]$
-267	1	-3×89	$[1, (1 + \sqrt{-267})/2]$	$[3, (3 + \sqrt{-267})/2]$
-403	1	-13×31	$[1, (1 + \sqrt{-403})/2]$	$[11, (9 + \sqrt{-403})/2]$
-427	1	-7×61	$[1, (1 + \sqrt{-427})/2]$	$[7, (7 + \sqrt{-427})/2]$

C Theta Series of $L_{\mathfrak{o}}$ and $L_{\mathfrak{a}}$ of $\mathbb{Q}(\sqrt{-5})$

$$\Theta_{L_{\mathfrak{o}}} = 1 + 2q + 2q^{4} + 2q^{5} + 4q^{6} + 6q^{9} + 4q^{14} + 2q^{16} + 2q^{20} + 8q^{21} + 4q^{24} + 2q^{25} + 4q^{29} + 4q^{30} + 6q^{36} + 4q^{41} + 6q^{45} + 4q^{46} + 6q^{49} + 8q^{54} + 4q^{56} + 4q^{61} + 2q^{64} + 8q^{69} + 4q^{70} + 2q^{80} + 10q^{81} + 8q^{84} + 4q^{86} + 4q^{89} + 4q^{94} + 4q^{96} + 2q^{100} + 4q^{101} + 8q^{105} + 4q^{109} + 4q^{116} + 4q^{120} + 2q^{121} + 2q^{125} + 12q^{126} + 8q^{129} + 4q^{134} + 8q^{141} + 6q^{144} + 4q^{145} + 4q^{149} + 4q^{150} + 8q^{161} + 4q^{164} + 4q^{166} + 2q^{169} + 8q^{174} + 6q^{180} + 4q^{181} + 4q^{184} + 16q^{189} + 6q^{196} + 8q^{201} + 4q^{205} + 4q^{206} + 4q^{214} + 8q^{216} + 4q^{224} + 6q^{225} + 4q^{229} + 4q^{230} + 4q^{241} + 4q^{244} + 6q^{245} + 8q^{246} + 8q^{249} + 4q^{254} + 2q^{256} + 12q^{261} + 4q^{269} + 8q^{270} + 8q^{276} + 4q^{280} + 4q^{281} + 2q^{289} + 12q^{294} + 8q^{301} + 4q^{305} + 8q^{309} + 2q^{320} + 8q^{321} + 10q^{324} + 4q^{326} + 8q^{329} + 4q^{334} + 8q^{336} + 4q^{344} + 8q^{345} + 4q^{349} + 4q^{350} + 4q^{356} + 2q^{361} + 8q^{366} + 12q^{369} + 4q^{376} + 8q^{381} + 4q^{384} + 4q^{389} + 2q^{400} + 4q^{401} + 4q^{404} + 10q^{405} + 8q^{406} + 4q^{409} + 12q^{414} + 8q^{420} + 4q^{421} + 4q^{430} + 4q^{436} + 18q^{441} + 4q^{445} + 4q^{446} + 4q^{449} + 4q^{454} + 4q^{461} + 4q^{464} + 8q^{469} + 4q^{470} + 4q^{480} + 2q^{484} + 12q^{486} + 8q^{489} + 2q^{500} + O[q]^{501}$$

$$\Theta_{L_{\mathfrak{a}}} = 1 + 2q^{2} + 4q^{3} + 4q^{7} + 2q^{8} + 2q^{10} + 4q^{12} + 4q^{15} + 6q^{18} + 4q^{23} + 8q^{27} + 4q^{28} + 2q^{32} + 4q^{35} + 2q^{40} + 8q^{42} + 4q^{43} + 4q^{47} + 4q^{48} + 2q^{50} + 4q^{58} + 4q^{60} + 12q^{63} + 4q^{67} + 6q^{72} + 4q^{75} + 4q^{82} + 4q^{83} + 8q^{87} + 6q^{90} + 4q^{92} + 6q^{98} + 4q^{103} + 4q^{107} + 8q^{108} + 4q^{112} + 4q^{115} + 4q^{122} + 8q^{123} + 4q^{127} + 2q^{128} + 8q^{135} + 8q^{138} + 4q^{140} + 12q^{147} + 2q^{160} + 10q^{162} + 4q^{163} + 4q^{167} + 8q^{168} + 4q^{172} + 4q^{175} + 4q^{178} + 8q^{183} + 4q^{188} + 4q^{192} + 2q^{200} + 4q^{202} + 8q^{203} + 12q^{207} + 8q^{210} + 4q^{215} + 4q^{218} + 4q^{223} + 4q^{227} + 4q^{232} + 4q^{235} + 4q^{240} + 2q^{242} + 12q^{243} + 2q^{250} + 12q^{252} + 8q^{258} + 4q^{263} + 8q^{267} + 4q^{268} + 8q^{282} + 4q^{283} + 8q^{287} + 6q^{288} + 4q^{290} + 4q^{298} + 4q^{300} + 8q^{303} + 4q^{307} + 12q^{315} + 8q^{322} + 8q^{327} + 4q^{328} + 4q^{332} + 4q^{335} + 2q^{338} + 8q^{343} + 4q^{347} + 8q^{348} + 6q^{360} + 4q^{362} + 4q^{363} + 4q^{367} + 4q^{368} + 4q^{375} + 16q^{378} + 4q^{383} + 12q^{387} + 6q^{392} + 8q^{402} + 4q^{410} + 4q^{412} + 4q^{415} + 12q^{423} + 8q^{427} + 4q^{428} + 8q^{432} + 8q^{435} + 4q^{443} + 8q^{447} + 4q^{448} + 6q^{450} + 4q^{458} + 4q^{460} + 4q^{463} + 4q^{467} + 4q^{482} + 16q^{483} + 4q^{487} + 4q^{488} + 6q^{490} + 8q^{492} + 8q^{498} + O[q]^{501}$$

$$\Theta_{L_{\mathfrak{o}},P} = q + 4q^{4} - 5q^{5} - 8q^{6} + 7q^{9} + 8q^{14} + 16q^{16} - 20q^{20} - 16q^{21} - 32q^{24} + 25q^{25} - 22q^{29} + 40q^{30} + 28q^{36} + 62q^{41} - 35q^{45} - 88q^{46} - 33q^{49} + 16q^{54} + 32q^{56} - 58q^{61} + 64q^{64} + 176q^{69} - 40q^{70} - 80q^{80} - 95q^{81} - 64q^{84} + 152q^{86} - 142q^{89} + 8q^{94} - 128q^{96} + 100q^{100} + 122q^{101} + 80q^{105} + 38q^{109} - 88q^{116} + 160q^{120} + 121q^{121} - 125q^{125} + 56q^{126} - 304q^{129} - 232q^{134} - 16q^{141} + 112q^{144} + 110q^{145} + 278q^{149} - 200q^{150} - 176q^{161} + 248q^{164} + 152q^{166} + 169q^{169} + 176q^{174} - 140q^{180} - 358q^{181} - 352q^{184} + 32q^{189} - 132q^{196} + 464q^{201} - 310q^{205} - 88q^{206} + 248q^{214} + 64q^{216} + 128q^{224} + 175q^{225} - 262q^{229} + 440q^{230} + 302q^{241} - 232q^{244} + 165q^{245} - 496q^{246} - 304q^{249} - 472q^{254} + 256q^{256} - 154q^{261} + 38q^{269} - 80q^{270} + 704q^{276} - 160q^{280} - 418q^{281} + 289q^{289} + 264q^{294} + 304q^{301} + 290q^{305} + 176q^{309} - 320q^{320} - 496q^{321} - 380q^{324} - 328q^{326} + 16q^{329} + 488q^{334} - 256q^{336} + 608q^{344} - 880q^{345} - 22q^{349} + 200q^{350} - 568q^{356} + 361q^{361} + 464q^{366} + 434q^{369} + 32q^{376} + 944q^{381} - 512q^{384} - 202q^{389} + 400q^{400} - 478q^{401} + 488q^{404} + 475q^{405} - 176q^{406} - 802q^{409} -$$

$$616q^{414}+320q^{420}-778q^{421}-760q^{430}+152q^{436}-231q^{441}+710q^{445}+872q^{446}+398q^{449}-712q^{454}+842q^{461}-352q^{464}-464q^{469}-40q^{470}+640q^{480}+484q^{484}+616q^{486}+656q^{489}-500q^{500}+O[q]^{501}$$

$$\Theta_{L_{\mathfrak{a},P}}=2q^2-4q^3+4q^7+8q^8-10q^{10}-16q^{12}+20q^{15}+14q^{18}-44q^{23}+8q^{27}+16q^{28}+32q^{32}-20q^{35}-40q^{40}-32q^{42}+76q^{43}+4q^{47}-64q^{48}+50q^{50}-44q^{58}+80q^{60}+28q^{63}-116q^{67}+56q^{72}-100q^{75}+124q^{82}+76q^{83}+88q^{87}-70q^{90}-176q^{92}-66q^{98}-44q^{103}+124q^{107}+32q^{108}+64q^{112}+220q^{115}-116q^{122}-248q^{123}-236q^{127}+128q^{128}-40q^{135}+352q^{138}-80q^{140}+132q^{147}-160q^{160}-190q^{162}-164q^{163}+244q^{167}-128q^{168}+304q^{172}+100q^{175}-284q^{178}+232q^{183}+16q^{188}-256q^{192}+200q^{200}+244q^{202}-88q^{203}-308q^{207}+160q^{210}-380q^{215}+76q^{218}+436q^{223}-356q^{227}-176q^{232}-20q^{235}+320q^{240}+242q^{242}+308q^{243}-250q^{250}+112q^{252}-608q^{258}-284q^{263}+568q^{267}-464q^{268}-32q^{282}+316q^{283}+248q^{287}+224q^{288}+220q^{290}+556q^{298}-400q^{300}-488q^{303}-596q^{307}-140q^{315}-352q^{322}-152q^{327}+496q^{328}+304q^{332}+580q^{335}+338q^{338}-328q^{343}-116q^{347}+352q^{348}-280q^{360}-716q^{362}-484q^{363}+724q^{367}-704q^{368}+500q^{375}+64q^{378}-44q^{383}+532q^{387}-264q^{392}+928q^{402}-620q^{410}-176q^{412}-380q^{415}+28q^{423}-232q^{427}+496q^{428}+128q^{432}-440q^{435}+796q^{443}-1112q^{447}+256q^{448}+350q^{450}-524q^{458}+880q^{460}-764q^{463}+124q^{467}+604q^{482}+704q^{483}+484q^{487}-464q^{488}+330q^{490}-992q^{492}-608q^{498}+O[q]^{501}$$

$$\Psi^{(1)}_{K,\Lambda}(z)=q+2q^2-4q^3+4q^4-5q^5-8q^6+4q^7+8q^8+7q^9-10q^{10}-16q^{12}+8q^{14}+20q^{15}+16q^{16}+14q^{18}-20q^{20}-16q^{21}-44q^{23}-32q^{24}+25q^{25}+8q^{27}+16q^{28}-22q^{29}+40q^{30}+32q^{32}-20q^{35}+28q^{36}-40q^{40}+62q^{41}-32q^{42}+76q^{43}-35q^{45}-88q^{46}+4q^{47}-64q^{48}-33q^{49}+50q^{50}+16q^{54}+32q^{56}-44q^{58}+80q^{60}-58q^{61}+28q^{63}+64q^{64}-116q^{67}+176q^{69}-40q^{70}+56q^{72}-100q^{75}-80q^{80}-95q^{81}+124q^{82}+76q^{83}-64q^{84}+152q^{86}+88q^{87}-142q^{89}-70q^{90}-176q^{92}+8q^{94}-128q^{96}-66q^{98}+100q^{100}+122q^{101}-44q^{103}+80q^{105}+124q^{107}+32q^{108}+38q^{109}+64q^{112}+220q^{115}-88q^{116}+160q^{120}+121q^{121}-116q^{122}-248q^{123}-125q^{125}+56q^{126}-236q^{127}+128q^{128}-304q^{129}-232q^{134}-40q^{135}+352q^{138}-80q^{140}-16q^{141}+112q^{144}+110q^{145}+132q^{147}+278q^{149}-200q^{150}-160q^{160}-176q^{161}-190q^{162}-164q^{163}+248q^{164}+152q^{166}+244q^{167}-128q^{168}+169q^{169}+304q^{172}+176q^{174}+100q^{175}-284q^{178}-140q^{180}-358q^{181}+232q^{183}-352q^{184}+16q^{188}+32q^{189}-256q^{192}-132q^{196}+200q^{200}+464q^{201}+244q^{202}-88q^{203}-310q^{205}-88q^{206}-308q^{207}+160q^{210}+248q^{214}-380q^{215}+64q^{216}+76q^{218}+436q^{223}+128q^{224}+175q^{225}-356q^{227}-262q^{229}+440q^{230}-176q^{232}-20q^{235}+320q^{240}+302q^{241}+242q^{242}+308q^{243}-232q^{244}+165q^{245}-496q^{246}-304q^{249}-250q^{250}+112q^{252}-472q^{254}+256q^{256}-608q^{258}-154q^{261}-284q^{263}+568q^{267}-464q^{268}+38q^{269}-80q^{270}+704q^{276}-160q^{280}-418q^{281}-32q^{282}+316q^{283}+248q^{287}+224q^{288}+289q^{289}+220q^{290}+264q^{294}+556q^{298}-400q^{300}+304q^{301}-488q^{303}+290q^{305}-596q^{307}+176q^{309}-140q^{315}-320q^{320}-496q^{321}-352q^{322}-380q^{324}-328q^{326}-152q^{327}+496q^{328}+16q^{329}+304q^{332}+488q^{334}+580q^{335}-256q^{336}+338q^{338}-328q^{343}+608q^{344}-880q^{345}-116q^{347}+352q^{348}-22q^{349}+200q^{350}-568q^{356}-280q^{360}+361q^{361}-716q^{362}-484q^{363}+464q^{366}+724q^{367}-704q^{368}+434q^{369}+500q^{375}+32q^{376}+64q^{378}+944q^{381}-$$

$$44q^{383}-512q^{384}+532q^{387}-202q^{389}-264q^{392}+400q^{400}-478q^{401}+928q^{402}+488q^{404}+475q^{405}-176q^{406}-802q^{409}-620q^{410}-176q^{412}-616q^{414}-380q^{415}+320q^{420}-778q^{421}+28q^{423}-232q^{427}+496q^{428}-760q^{430}+128q^{432}-440q^{435}+152q^{436}-231q^{441}+796q^{443}+710q^{445}+872q^{446}-1112q^{447}+256q^{448}+398q^{449}+350q^{450}-712q^{454}-524q^{458}+880q^{460}+842q^{461}-764q^{463}-352q^{464}+124q^{467}-464q^{469}-40q^{470}+640q^{480}+604q^{482}+704q^{483}+484q^{484}+616q^{486}+484q^{487}-464q^{488}+656q^{489}+330q^{490}-992q^{492}-608q^{498}-500q^{500}+O[q]^{501}$$

Acknowledgements The authors thank Masao Koike for informing us that our previous results in [3] can be interpreted in terms of the cusp forms attached to L-functions with a Hecke character of CM fields, i.e., the imaginary quadratic fields $\mathbb{Q}(\sqrt{-1})$ and $\mathbb{Q}(\sqrt{-3})$, and in particular for bringing our attention to Theorem 1.31 in [15]. The authors also thank Junichi Shigezumi for his helpful discussions and computations on this research.

References

1. S. Ahlgren, Multiplicative Relations in Powers of Euler's Product, *Journal of Number Theory*, **89** (2001), 222–233.
2. E. Bannai, M. Koike, M. Shinohara, M. Tagami, Spherical designs attached to extremal lattices and the modulo p property of Fourier coefficients of extremal modular forms, *Mosc. Math. J.*, **6-2** (2006), 225–264.
3. E. Bannai, T. Miezaki, Toy models for D. H. Lehmer's conjecture, *J. Math. Soc. Japan*, **62-3** (2010), 687–705.
4. D. A. Cox, *Primes of the form x^2+ny^2. Fermat, class field theory and complex multiplication.*, A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1989.
5. P. de la Harpe and C. Pache, Cubature formulas, geometrical designs, reproducing kernels, and Markov operators, *Infinite groups: geometric, combinatorial and dynamical aspects*, *Progr. Math.*, Birkhäuser, Basel, **248** (2005), 219–267.
6. P. de la Harpe, C. Pache, B. Venkov, Construction of spherical cubature formulas using lattices, *Algebra i Analiz*, **18-1** (2006), 162–186, ; translation in *St. Petersburg Math. J.* **18-1** (2007), 119–139.
7. P. Delsarte, J.-M. Goethals, and J. J. Seidel, Spherical codes and designs, *Geom. Dedicata* **6** (1977), 363–388.
8. W. Ebeling, *Lattices and codes.*, Advanced Lectures in Mathematics. Friedr. Vieweg & Sohn, Braunschweig, 2002.
9. E. Hecke, *Mathematische Werke*, Vandenhoeck & Ruprecht, Göttingen, 1983.
10. N. Katz, An overview of Deligne's proof of the Riemann hypothesis for varieties over finite fields, *Amer. Math. Soc. Proc. Symp. Pure Math.*, **28** (1976), 275–305.
11. N. Koblitz, *Introduction to Elliptic Curves and Modular Forms*, Springer-Verlag, Berlin/New York, 1984.
12. D. H. Lehmer, The vanishing of Ramanujan's $\tau(n)$, *Duke Math. J.* **14** (1947), 429–433.
13. *T. Miezaki's website*, http://sites.google.com/site/tmiezaki/home/
14. T. Miyake, *Modular forms*, Translated from the Japanese by Yoshitaka Maeda. Springer-Verlag, Berlin, 1989.
15. K. Ono, *The web of modularity:arithmetic of the coefficients of modular forms and q-series*, CBMS Regional Conference Series in Mathematics, vol. 102, Published for the Conference Board of the Mathematical Sciences, Washington, DC, 2004.
16. C. Pache, Shells of selfdual lattices viewed as spherical designs, *International Journal of Algebra and Computation* **5** (2005), 1085–1127.

17. R. Prime, Hecke character, http://www.math.uconn.edu/~prime/whatthehecke.pdf
18. *Sage*, http://www.sagemath.org/
19. B. Schoeneberg, Das Verhalten von mehrfachen Thetareihen bei Modulsubstitutionen, (German) *Math. Ann.* **116-1** (1939), 511–523.
20. B. Schoeneberg, *Elliptic modular functions: an introduction*, Translated from the German by J. R. Smart and E. A. Schwandt. Die Grundlehren der mathematischen Wissenschaften, Band 203. Springer-Verlag, New York-Heidelberg, 1974.
21. J.-P. Serre, *A course in arithmetic*, Translated from the French. Graduate Texts in Mathematics, **7**, Springer-Verlag, New York-Heidelberg, 1973.
22. J.-P. Serre, Sur la lacunarité des puissances de η, *Glasgow Math. J.* **27** (1985), 203–221.
23. B. B. Venkov, Even unimodular extremal lattices, (Russian) *Algebraic geometry and its applications. Trudy Mat. Inst. Steklov.*, **165** (1984), 43–48; translation in *Proc. Steklov Inst. Math.* **165** (1985) 47–52.
24. B. B. Venkov, Réseaux et designs sphériques, (French) [Lattices and spherical designs], *Réseaux euclidiens, designs sphériques et formes modulaires*, Monogr. Enseign. Math. **37** (2001), 10–86, Enseignement Math., Geneva.
25. D. B. Zagier, *Zetafunktionen und quadratische Körper : eine Einführung in die höhere Zahlentheorie*, Springer-Verlag, Berlin-New York, 1981.

On Representation of an Integer by $X^2 + Y^2 + Z^2$ and the Modular Equations of Degree 3 and 5

Alexander Berkovich

There are always flowers for those who want to see them

Abstract I discuss a variety of results involving $s(n)$, the number of representations of n as a sum of three squares. One of my objectives is to reveal numerous interesting connections between the properties of this function and certain modular equations of degree 3 and 5. In particular, I show that

$$s(25n) = (6 - (-n|5))\, s(n) - 5s\left(\frac{n}{25}\right)$$

follows easily from the well known Ramanujan modular equation of degree 5. Moreover, I establish new relations between $s(n)$ and $h(n)$, $g(n)$, the number of representations of n by the ternary quadratic forms

$$2x^2 + 2y^2 + 2z^2 - yz + zx + xy, \quad x^2 + y^2 + 3z^2 + xy,$$

respectively.

Finally, I propose a remarkable new identity for $s(p^2n) - ps(n)$ with p being an odd prime. This identity makes nontrivial use of the ternary quadratic forms with discriminants p^2, $16p^2$.

Key words and Phrases Ternary quadratic forms • Sum of three squares • Modular equations • θ-function identities

Mathematics Subject Classification (2010): Primary 11E20, 11F37, 11B65; Secondary 05A30, 33 E05

A. Berkovich (✉)
Department of Mathematics, University of Florida, Gainesville, FL 32611-8105, USA
e-mail: alexb@ufl.edu

K. Alladi et al. (eds.), *Quadratic and Higher Degree Forms*, Developments in Mathematics 31, DOI 10.1007/978-1-4614-7488-3_2,

1 Introduction

Let $(a,b,c,d,e,f)(n)$ denote the number of representations of n by the ternary form $ax^2+by^2+cz^2+dyz+ezx+fxy$. I will assume that $(a,b,c,d,e,f)(n)=0$, whenever $n\notin Z$. Let $s(n)$ denote the number of representations of n by ternary form $x^2+y^2+z^2$. In [14], Hirschhorn and Sellers proved in a completely elementary manner that

$$s(p^2n)=(p+1-(-n|p))\,s(n)-ps\left(\frac{n}{p^2}\right), \tag{1.1}$$

when $p=3$. Here $(a|p)$ denotes the Legendre symbol. It should be pointed out that the authors of [14] proved (1.1) for all odd prime numbers p by an appeal to the theory of modular forms.

In Sect. 2, I will show that (1.1) with $p=5$ follows easily from the well-known identity for $\phi(q)^2-\phi(q^5)^2$ with

$$\phi(q)=\sum_{n=-\infty}^{\infty} q^{n^2}. \tag{1.2}$$

Here and throughout, q is a complex number with $|q|<1$. I will also provide an elementary proof of the following

Theorem 1.1. *If* $n\equiv 1,2$ *mod* 4, *then*

$$s(25n)-5s(n)=4(2,2,2,-1,1,1)(n), \tag{1.3}$$

and

Theorem 1.2. *If* $n\equiv 1,2$ *mod* 4, *then*

$$s(9n)-3s(n)=2(1,1,3,0,0,1)(n). \tag{1.4}$$

In Sect. 5, I will show how to remove the parity restrictions in the above theorems by proving Theorems 5.2 and 5.3. Section 6 contains my new Proposition 6.1, which generalizes Theorems 1.1, 1.2, 5.2 and 5.3. A reader with no vested interest in q-series may want to proceed directly to Sect. 6. However, a motivated reader may decide to walk slowly through the initial sections to experience suffering which will later turn into joy.

Let me point out that two ternary forms $2x^2+2y^2+2z^2-yz+zx+xy$ and $x^2+y^2+3z^2+xy$ both have class number one. This implies that these forms are both regular [11, 16, 17]. For a recent discussion of the relation between the Ramanujan modular equations and certain ternary quadratic forms the reader is invited to examine [2]. And it goes without saying that one should not forget the timeless classic [1].

I begin by recalling some standard notations, definitions, and useful formulas.

$$(a;q)_\infty := \prod_{j\geq 0}(1-aq^j), \tag{1.5}$$

and

$$E(q) := \prod_{j\geq 1}(1-q^j). \tag{1.6}$$

Note that

$$E(-q) = \frac{E(q^2)^3}{E(q^4)E(q)}, \tag{1.7}$$

Ramanujan's general theta-function $f(a,b)$ is defined by

$$f(a,b) = \sum_{n=-\infty}^{\infty} a^{\frac{(n-1)n}{2}} b^{\frac{(n+1)n}{2}}, \quad |ab|<1. \tag{1.8}$$

In Ramanujan's notation, the celebrated Jacobi triple product identity takes the shape [5], p. 35

$$f(a,b) = (-a;ab)_\infty(-b;ab)_\infty(ab;ab)_\infty, \quad |ab|<1. \tag{1.9}$$

Note that $\phi(q)$ can be interpreted as

$$\phi(q) = f(q,q) = \frac{E(q^2)^5}{E(q^4)^2E(q)^2}, \tag{1.10}$$

where the product on the right follows easily from (1.8). We shall also require

$$\phi(-q) = \frac{E(q)^2}{E(q^2)}. \tag{1.11}$$

Next we define

$$\psi(q) = f(q,q^3) = \sum_{n=-\infty}^{\infty} q^{2n^2+n}. \tag{1.12}$$

It is not hard to check that

$$\psi(q) = \frac{1}{2}f(1,q) = \sum_{n\geq 0} q^{\frac{(n+1)n}{2}} = \frac{E(q^2)^2}{E(q)}, \tag{1.13}$$

$$\sum_{n=-\infty}^{\infty} q^{(4n+1)^2} = \sum_{n=-\infty}^{\infty} q^{(4n+3)^2} = q\psi(q^8), \tag{1.14}$$

and that

$$f(q,q^9)f(q^3,q^7)=\frac{E(q^{20})E(q^5)E(q^2)^2}{E(q^4)E(q)}, \tag{1.15}$$

$$f(q,q^4)f(q^2,q^3)=\frac{E(q^5)^3E(q^2)}{E(q^{10})E(q)}. \tag{1.16}$$

The function $f(a,b)$ may be dissected in many different ways. We will use the following trivial dissections [5], pp. 40, 49

$$\phi(q)=\phi(q^4)+2q\psi(q^8), \tag{1.17}$$

$$\phi(q)=\phi(q^9)+2qf(q^3,q^{15}), \tag{1.18}$$

$$\phi(q)=\phi(q^{25})+2qf(q^{15},q^{35})+2q^4f(q^5,q^{45}). \tag{1.19}$$

We will also require a special case of Schröter's formula [5], p. 45

$$f(a,b)f(c,d)=f(ac,bd)f(ad,bc)+af\left(\frac{b}{c},ac^2d\right)f\left(\frac{b}{d},acd^2\right), \tag{1.20}$$

provided $ab=cd$. Setting $a=b=c=d=q$ in (1.20) we obtain

$$\phi(q)^2=\phi(q^2)^2+4q\psi(q^4)^2. \tag{1.21}$$

Iterating, we find that

$$\phi(q)^2=\phi(q^4)^2+4q\psi(q^4)^2+4q^2\psi(q^8)^2. \tag{1.22}$$

Next, we set $a=q, b=q^9, c=q^3, d=q^7$ in (1.20) and square the result. This way we have

$$\begin{aligned}f(q,q^9)^2f(q^3,q^7)^2&=f(q^4,q^{16})^2f(q^8,q^{12})^2\\&+2qf(q^4,q^{16})f(q^8,q^{12})f(q^6,q^{14})f(q^2,q^{18})+q^2f(q^6,q^{14})^2f(q^2,q^{18})^2.\end{aligned} \tag{1.23}$$

Finally, we multiply both sides in (1.23) by

$$\frac{E(q^4)\phi(q^5)}{E(q^{20})E(q^{10})^2},$$

and use (1.10), (1.13), (1.15) and (1.16) to arrive at

$$\begin{aligned}\phi(q)f(q^2,q^8)f(q^4,q^6)&=\psi(q^4)\phi(q^5)\phi(q^{10})\\&+2q\psi(q^2)\psi(q^{10})\phi(q^5)+q^2\psi(q^{20})\phi(q^2)\phi(q^5).\end{aligned} \tag{1.24}$$

This result will come in handy in my proof of (1.3) with $n \equiv 2 \bmod 4$. To deal with the case $n \equiv 1 \bmod 4$ in (1.3) I will require another identity

$$\phi(q)\phi(q^5) + \sum_{m,n} q^{2m^2+2nm+3n^2} = 2\Pi_1(q), \tag{1.25}$$

where

$$\Pi_1(q) = \frac{E(q^{10})E(q^5)E(q^4)E(q^2)}{E(q^{20})E(q)}. \tag{1.26}$$

This formula was discovered and proven in [4]. The proof of (1.25), given in [4], used only a special case of the Ramanujan ${}_1\psi_1$ summation formula [6], p. 64. Multiplying both sides in (1.25) by $\psi(q^{10})$ and utilizing (1.13) and (1.15) we can rewrite (1.25) as

$$\psi(q^{10})\phi(q)\phi(q^5)+\psi(q^{10})\sum_{m,n} q^{2m^2+2nm+3n^2} = 2\psi(q^2)f(q,q^9)f(q^3,q^7). \tag{1.27}$$

2 The Ternary Implications of the Fundamental Modular Equation of Degree 5

In this section we will make an extensive use of a well-known modular equation of degree 5

$$\phi(q)^2 - \phi(q^5)^2 = 4qf(q,q^9)f(q^3,q^7) \tag{2.1}$$

to prove (1.1) with $p = 5$. We note that (2.1) has an attractive companion

$$5\phi(q^5)^2 - \phi(q)^2 = 4\Pi_2(q), \tag{2.2}$$

where

$$\Pi_2(q) = \frac{E(q^{10})^2E(q^4)E(q)}{E(q^{20})E(q^5)}. \tag{2.3}$$

Both (2.1) and (2.2) are discussed in [5]. We remark that the right hand side of (2.1) was interpreted in terms of so-called self-conjugate 5-cores in [12]. To proceed further I will need a sifting operator $S_{t,s}$. It is defined by its action on power series as follows

$$S_{t,s}\sum_{n\geq 0} c(n)q^n = \sum_{k\geq 0} c(tk+s)q^k. \tag{2.4}$$

Here t, s are integers such that $0 \leq s < t$. Making use of (1.19), we find that

$$S_{5,0}\phi(q)^2 = \phi(q^5)^2 + 8qf(q,q^9)f(q^3,q^7). \tag{2.5}$$

And so

$$S_{5,0}(\phi(q)^2-\phi(q^5)^2)=-(\phi(q)^2-\phi(q^5)^2)+8qf(q,q^9)f(q^3,q^7). \tag{2.6}$$

Employing (2.1) twice, we see that

$$S_{5,0}(qf(q,q^9)f(q^3,q^7))=qf(q,q^9)f(q^3,q^7). \tag{2.7}$$

Analogously, we can check that

$$S_{5,0}\phi(q)^3=\phi(q^5)^3+24q\phi(q^5)f(q,q^9)f(q^3,q^7), \tag{2.8}$$

and that

$$S_{5,1}\phi(q)^3=6f(q^3,q^7)(\phi(q^5)^2+4qf(q,q^9)f(q^3,q^7))=6f(q^3,q^7)\phi(q)^2, \tag{2.9}$$

$$S_{5,4}\phi(q)^3=6f(q,q^9)(\phi(q^5)^2+4qf(q,q^9)f(q^3,q^7))=6f(q,q^9)\phi(q)^2. \tag{2.10}$$

We note, in passing, that thanks to (1.9), the right hand side in (2.9) can be rewritten as an infinite product

$$\sum_{n=0}^{\infty}s(5n+1)q^n=6\prod_{j=1}^{\infty}(1-q^{2j})^2(1-q^{10j})$$
$$(1+q^{-1+2j})^4(1+q^{-3+10j})(1+q^{-7+10j}).$$

Cooper and Hirschhorn studied the generating functions of subsequences of $s(n)$ that could be represented by a single, simple infinite product. For example, (2.9), (2.10) and (4.17) are the formulas (3.1), (3.2) and (1.1) in [10].

With the aid of (1.19) we can combine (2.9) and (2.10) into a single elegant statement

$$S_{5,r}(\phi(q)^3-3\phi(q)\phi(q^5)^2)=0, \tag{2.11}$$

where $r=1,4$. Next, we apply $S_{5,0}$ to both sides of (2.8) to obtain, with a little help from (2.7)

$$S_{25,0}\phi(q)^3=\phi(q)^3+24q\phi(q)f(q,q^9)f(q^3,q^7). \tag{2.12}$$

Subtracting $5\phi(q)^3$ and making use of (2.1) again, we deduce that

$$\begin{aligned}S_{25,0}\phi(q)^3-5\phi(q)^3&=-4\phi(q)^3+6\phi(q)(\phi(q)^2-\phi(q^5)^2)\\&=2(\phi(q)^3-3\phi(q)\phi(q^5)^2).\end{aligned} \tag{2.13}$$

Finally, we apply $S_{5,r}$ with $r=1,4$ to both sides of (2.13) to find that

$$S_{125,25r}\phi(q)^3-5S_{5,r}\phi(q)^3=0. \tag{2.14}$$

But it is plain that

$$\phi(q)^3 = \sum_{n=0}^{\infty} s(n)q^n. \tag{2.15}$$

And so the equation (2.14) can be interpreted as

$$s(25n) - 5s(n) = 0, \tag{2.16}$$

when $n \equiv 1, 4 \bmod 5$. Thus, the proof of (1.1) with $p = 5$ and $n \equiv 1, 4 \bmod 5$ is complete.

We now turn our attention to the $n \equiv 2, 3 \bmod 5$ case. Subtracting $2\phi(q)^3$ from the extremes of (2.13), we end up with the formula

$$S_{25,0}\phi(q)^3 - 7\phi(q)^3 = -6\phi(q)\phi(q^5)^2. \tag{2.17}$$

It is now clear that for $r = 2, 3$

$$S_{5,r}(S_{25,0}\phi(q)^3 - 7\phi(q)^3) = -6\phi(q)^2 S_{5,r}\phi(q) = 0, \tag{2.18}$$

where in the last step we took advantage of the dissection formula (1.19). Obviously, (2.18) is equivalent to

$$s(25n) - 7s(n) = 0, \tag{2.19}$$

when $n \equiv 2, 3 \bmod 5$. And so we completed the proof of (1.1) with $p = 5$ and $n \equiv 2, 3 \bmod 5$. All that remains to do is to take care of the $n \equiv 0 \bmod 5$ case. Adding $\phi(q)^3$ to both sides of (2.17) and applying $S_{5,0}$ to the result, we get

$$S_{5,0}(S_{25,0}\phi(q)^3 - 6\phi(q)^3) = S_{5,0}(\phi(q)^3 - 6\phi(q)\phi(q^5)^2). \tag{2.20}$$

Next, we utilize (1.19), (2.1) and (2.8) to process the right hand side of (2.20) as follows

$$\begin{aligned} S_{5,0}(\phi(q)^3 - 6\phi(q)\phi(q^5)^2) &= \phi(q^5)^3 + 6\phi(q^5)(\phi(q)^2 - \phi(q^5)^2) - 6\phi(q^5)\phi(q)^2 \\ &= -5\phi(q^5)^3. \end{aligned}$$

Hence, we have shown that

$$S_{125,0}\phi(q)^3 - 6S_{5,0}\phi(q)^3 = -5\phi(q^5)^3. \tag{2.21}$$

Consequently,

$$s(25n) - 6s(n) = -5s\left(\frac{n}{25}\right), \tag{2.22}$$

when $5|n$. This concludes our proof of (1.1) with $p = 5$.

3 Proof of Theorem 1.1

I begin by observing that Theorem 1.1 is equivalent to the following statement

$$S_{100,25r}\phi(q)^3 - 5S_{4,r}\phi(q)^3 = 4S_{4,r}T(q), \tag{3.1}$$

where

$$T(q) := \sum_{x,y,z} q^{2x^2+2y^2+2z^2-yz+zx+xy} \tag{3.2}$$

and $r = 1, 2$. It is not hard to verify that

$$S_{4,1}T(q) = 6S_{4,1}X(1,q), \tag{3.3}$$

and that

$$S_{4,2}T(q) = 3S_{4,2}(X(0,q) + X(2,q)). \tag{3.4}$$

Here

$$X(r,q) := \sum_{\substack{x,\\ y\equiv -z\equiv r \bmod 4}} q^{2x^2+2y^2+2z^2-yz+zx+xy}. \tag{3.5}$$

It takes very little effort to check that

$$2x^2+2y^2+2z^2-zy+zx+xy = 2\left(x+\frac{y+z}{4}\right)^2+\frac{5}{8}(y+z)^2+\frac{5}{4}(y-z)^2. \tag{3.6}$$

Hence

$$\begin{aligned} X(r,q) &= \sum_{\substack{x,\\ y\equiv -z\equiv r \bmod 4}} q^{2\left(x+\frac{y+z}{4}\right)^2+10(\frac{y+z}{4})^2+20(\frac{y-z}{4})^2} \\ &= \sum_{\substack{u,\\ w\equiv v+\frac{r}{2} \bmod 2}} q^{2u^2+10v^2+20w^2}, \end{aligned} \tag{3.7}$$

for $r = 0, 2$. It is now evident that

$$X(0,q) + X(2,q) = \sum_{u,v,w} q^{2u^2+10v^2+20w^2} = \phi(q^2)\phi(q^{10})\phi(q^{20}). \tag{3.8}$$

Using this last result in (3.4), we find that

$$S_{4,2}T(q) = 3\phi(q^5)S_{4,2}(\phi(q^2)\phi(q^{10})). \tag{3.9}$$

Recalling (1.17), we obtain at once that

$$4S_{4,2}T(q) = 24\phi(q^5)(\psi(q^4)\phi(q^{10}) + 6q^2\phi(q^2)\psi(q^{20})). \tag{3.10}$$

We now consider $X(r,q)$ with $r=1,3$.

$$X(r,q)=\sum_{\substack{u,\\ v\equiv w \bmod 2}} q^{2u^2+10v^2+5(2w+r)^2}.$$

Recalling (1.14), we get

$$X(1,q)=X(3,q)=\sum_{u,v,\tilde{w}} q^{2n^2+10v^2+5(4\tilde{w}+1)^2}=q^5\phi(q^2)\phi(q^{10})\psi(q^{40}). \quad (3.11)$$

Using (1.17), (3.3) and (3.11), we deduce that

$$S_{4,1}T(q)=6q\psi(q^{10})S_{4,0}(\phi(q^2)\phi(q^{10}))=6q\psi(q^{10})(\phi(q^2)\phi(q^{10})+4q^3\psi(q^4)\psi(q^{20})).$$

Also, it is not hard to check that

$$\begin{aligned}\sum_{m,n} q^{2m^2+2nm+3n^2}&=\sum_{m,n} q^{2(m+n)^2+10n^2}+q^3\sum_{m,n} q^{2(m+n+1)(m+n)+10(n+1)n}\\ &=\phi(q^2)\phi(q^{10})+4q^3\psi(q^4)\psi(q^{20}). \quad (3.12)\end{aligned}$$

This implies that

$$4S_{4,1}T(q)=24q\psi(q^{10})\sum_{m,n} q^{2m^2+2nm+3n^2}. \quad (3.13)$$

Next, we employ (2.13) to get

$$S_{100,25r}\phi(q)^3-5S_{4,r}\phi(q)^3=2S_{4,r}(\phi(q)^3-3\phi(q)\phi(q^5)^2). \quad (3.14)$$

With the aid of (1.17), (1.22), (2.1) and (2.2) we verify that

$$S_{4,1}(\phi(q)^3-3\phi(q)\phi(q^5)^2)=24q\psi(q^2)f(q,q^9)f(q^3,q^7)-12q\phi(q)\phi(q^5)\psi(q^{10}), \quad (3.15)$$

$$S_{4,2}(\phi(q)^3-3\phi(q)\phi(q^5)^2)=-24q\psi(q^2)\psi(q^5)^2+12\phi(q)f(q^2,q^8)f(q^4,q^6). \quad (3.16)$$

Utilizing these results in (3.14) we obtain

$$S_{100,25}\phi(q)^3-5S_{4,1}\phi(q)^3=48q\psi(q^2)f(q,q^9)f(q^3,q^7)-24q\phi(q)\phi(q^5)\psi(q^{10}), \quad (3.17)$$

$$S_{100,50}\phi(q)^3-5S_{4,2}\phi(q)^3=-48q\psi(q^2)\psi(q^5)^2+24\phi(q)f(q^2,q^8)f(q^4,q^6). \quad (3.18)$$

Recalling (3.13), we see that (3.1) with $r=1$ is equivalent to

$$2\psi(q^2)f(q,q^9)f(q^3,q^7)-\phi(q)\phi(q^5)\psi(q^{10})=\psi(q^{10})\sum_{m,n} q^{2m^2+2nm+3n^2},$$

which is, essentially, (1.27). Analogously, employing (3.10), we find that (3.1) with $r = 2$ is equivalent to

$$-2q\psi(q^2)\psi(q^5)^2+\phi(q)f(q^2,q^8)f(q^4,q^6) = \phi(q^5)\psi(q^4)\phi(q^{10})+6q^2\phi(q^2)\phi(q^5)\psi(q^{20}),$$

which is, essentially, (1.24). The proof of Theorem 1.1 is now complete.

In Sect. 5 we will generalize Theorem 1.1. To this end we need to define

$$Y(r,q) := \sum_{\substack{x,\\ y+z\equiv r \bmod 4}} q^{2x^2+2y^2+2z^2-yz+zx+xy}, \tag{3.19}$$

where $r = 0,1,2,3$. Observe that the condition $y+z \equiv r \bmod 4$ allows us to introduce new summation variables u, v, w, defined as $x = w-v$, $y = 2u+v+r$, $z = 2u-v$. Using (3.6), it is easy to see that

$$2x^2+2y^2+2z^2-zy+zx+xy = 2r^2+w(2w+r)+5v(v+r)+5u(2u+r).$$

Hence

$$Y(0,q) = \phi(q^2)\phi(q^5)\phi(q^{10}), \tag{3.20}$$

$$Y(2,q) = 4q^3\phi(q^5)\psi(q^4)\psi(q^{20}), \tag{3.21}$$

$$Y(1,q) = Y(3,q) = 2q^2\psi(q)\psi(q^5)\psi(q^{10}). \tag{3.22}$$

Employing (3.12), (3.20)–(3.22), we derive

$$T(q) = \sum_{r=0}^{3} Y(r,q) = \phi(q^5)\sum_{m,n} q^{2m^2+2nm+3n^2} + 4q^2\psi(q)\psi(q^5)\psi(q^{10}). \tag{3.23}$$

It is easy to see that

$$\sum_{\substack{x,\\ y+z\equiv 1 \bmod 2}} q^{2x^2+2y^2+2z^2-yz+zx+xy} = Y(1,q)+Y(3,q) = 4q^2\psi(q)\psi(q^5)\psi(q^{10}), \tag{3.24}$$

and that

$$\sum_{\substack{x,\\ y+z\equiv 1 \bmod 2}} q^{2x^2+2y^2+2z^2-yz+zx+xy} = 2Z(q), \tag{3.25}$$

where

$$Z(q) := \sum_{\substack{x,\\ y\equiv 0 \bmod 2,\\ z\equiv 1 \bmod 2}} q^{2x^2+2y^2+2z^2-yz+zx+xy}. \tag{3.26}$$

It is worthwhile to point out that $Z(q)$ has six equivalent representations. For example, one has

$$Z(q) := \sum_{\substack{x\equiv 0 \bmod 2,\\ y\equiv 1 \bmod 2,\\ z}} q^{2x^2+2y^2+2z^2-yz+zx+xy}.$$

From (3.24), (3.25) we deduce that

$$Z(q) = 2q^2\psi(q)\psi(q^5)\psi(q^{10}). \tag{3.27}$$

We conclude this Section that by proving that

$$\sum_{\substack{x+y\equiv 1 \bmod 2,\\ y\equiv z \bmod 2}} q^{2x^2+2y^2+2z^2-yz+zx+xy} = Z(q). \tag{3.28}$$

Indeed, the left hand side of (3.28) can be rewritten as

$$\sum_{\substack{x\equiv 0 \bmod 2,\\ y\equiv 1 \bmod 2,\\ z\equiv 1 \bmod 2}} q^{2x^2+2y^2+2z^2-yz+zx+xy} + \sum_{\substack{x\equiv 1 \bmod 2,\\ y\equiv 0 \bmod 2,\\ z\equiv 0 \bmod 2}} q^{2x^2+2y^2+2z^2-yz+zx+xy}.$$

Now observe that

$$\sum_{\substack{x\equiv 1 \bmod 2,\\ y\equiv 0 \bmod 2,\\ z\equiv 0 \bmod 2}} q^{2x^2+2y^2+2z^2-yz+zx+xy} = \sum_{\substack{x\equiv 0 \bmod 2,\\ y\equiv 1 \bmod 2,\\ z\equiv 0 \bmod 2}} q^{2x^2+2y^2+2z^2-yz+zx+xy}.$$

And so the left hand side of (3.28) becomes

$$\sum_{\substack{x\equiv 0 \bmod 2,\\ y\equiv 1 \bmod 2,\\ z}} q^{2x^2+2y^2+2z^2-yz+zx+xy} = \sum_{\substack{x,\\ y\equiv 0 \bmod 2,\\ z\equiv 1 \bmod 2}} q^{2x^2+2y^2+2z^2-yz+zx+xy} = Z(q),$$

as desired.

4 Cubic Modular Identities Revisited

As in the last section, I begin by observing that Theorem 1.2 is equivalent to the following statement

$$S_{36,9r}\phi(q)^3 - 3S_{4,r}\phi(q)^3 = 4S_{4,r}\phi(q^3)a(q), \tag{4.1}$$

where

$$a(q) := \sum_{x,y} q^{x^2+xy+y^2},$$

and $r = 1, 2$. The function $a(q)$ was extensively studied in the literature [7–9, 13]. It appeared in Borwein's cubic analogue of Jacobi's celebrated theta function identity [8]. I will record below some useful formulas

$$4a(q^2)\phi(q^3) = \phi(q)^3 + 3\tfrac{\phi(q^3)^4}{\phi(q)}, \tag{4.2}$$

$$a(q) = a(q^3) + 6q\tfrac{E(q^9)^3}{E(q^3)}, \tag{4.3}$$

$$a(q) = \phi(q)\phi(q^3) + 4q\psi(q^2)\psi(q^6), \tag{4.4}$$

$$a(q) = 2\phi(q)\phi(q^3) - \phi(-q)\phi(-q^3), \tag{4.5}$$

$$2a(q^2) - a(q) = \tfrac{\phi(-q)^3}{\phi(-q^3)} \tag{4.6}$$

$$a(q) = a(q^4) + 6q\psi(q^2)\psi(q^6). \tag{4.7}$$

Formula (4.2) appears as equation (6.4) in [7]. Identities (4.3)–(4.6) are discussed in [9]. In order to prove (4.7), the authors of [13] have shown that

$$2q\psi(q^2)\psi(q^6) = \sum_{u \not\equiv v \bmod 2} q^{u^2+3v^2}. \tag{4.8}$$

We have at once that

$$\begin{aligned} 2q\psi(q^2)\psi(q^6) &= \sum_{\substack{u\equiv 1 \bmod 2,\\ v\equiv 0 \bmod 2}} q^{u^2+3v^2} + \sum_{\substack{u\equiv 0 \bmod 2,\\ v\equiv 1 \bmod 2}} q^{u^2+3v^2} \\ &= 2q\psi(q^8)\phi(q^{12}) + 2q^3\phi(q^4)\psi(q^{24}). \end{aligned} \tag{4.9}$$

Combining (4.7) and (4.9), we have a pretty neat dissection of $a(q)$ mod 4

$$a(q) = a(q^4) + 6q\psi(q^8)\phi(q^{12}) + 6q^3\phi(q^4)\psi(q^{24}). \tag{4.10}$$

In [19], L.C. Shen discussed two well-known modular identities of degree 3

$$\phi(q)^2 - \phi(q^3)^2 = 4q\frac{\psi(q)\psi(q^3)\psi(q^6)}{\psi(q^2)}, \tag{4.11}$$

and

$$\phi(q)^2 + \phi(q^3)^2 = 2\frac{\psi(q)f(q,q^2)f(q^2,q^4)}{\psi(q^2)}. \tag{4.12}$$

Multiplying (4.11) and (4.12), and using

$$f(q,q^2)=\frac{E(q^3)^2E(q^2)}{E(q^6)E(q)}, \tag{4.13}$$

$$f(q,q^5)=\frac{E(q^{12})E(q^3)E(q^2)^2}{E(q^6)E(q^4)E(q)} \tag{4.14}$$

together with (1.13) we have

$$\phi(q)^4-\phi(q^3)^4=8q\phi(q^3)f(q,q^5)^3. \tag{4.15}$$

Next, we rewrite (4.15) as

$$\frac{\phi(q)^4}{\phi(q^3)}=\phi(q^3)^3+8qf(q,q^5)^3. \tag{4.16}$$

Recalling (1.18), we can recognize the expression on the right as

$$\phi(q^3)^3+8qf(q,q^5)^3=S_{3,0}(\phi(q^9)+2qf(q^3,q^{15}))^3=S_{3,0}\phi(q)^3.$$

And so

$$S_{3,0}\phi(q)^3=\frac{\phi(q)^4}{\phi(q^3)}. \tag{4.17}$$

Next, we want to show that

$$S_{9,0}\phi(q)^3=\frac{4\phi(q)^4-3\phi(q^3)^4}{\phi(q)}. \tag{4.18}$$

To this end, we apply $S_{3,0}$ to both sides of (4.17). Utilizing (1.18), we find that

$$S_{9,0}\phi(q)^3=\frac{\phi(q^3)^4+4(8q\phi(q^3)f(q,q^5)^3)}{\phi(q)}. \tag{4.19}$$

The statement in (4.18) follows immediately from (4.15) and (4.19). Moreover, we have

$$S_{9,0}\phi(q)^3-5\phi(q)^3=-\phi(q)^3-3\frac{\phi(q^3)^4}{\phi(q)}=-4a(q^2)\phi(q^3), \tag{4.20}$$

where we used (4.2) in the last step. Adding $2\phi(q)^3$ to the extremes in (4.20) we derive

$$S_{9,0}\phi(q)^3-3\phi(q)^3=2\phi(q)^3-4a(q^2)\phi(q^3). \tag{4.21}$$

This result will come in handy in my proof of Theorem 5.2 in the next section.

5 Proof of Theorems 1.2, 5.2 and 5.3

I begin this section by providing an easy proof of two formulas in (4.1). All I need is the following

Lemma 5.1. *If* $r = 1, 2$, *then*

$$S_{4,r}(\phi(q)^3 - 2a(q^2)\phi(q^3)) = S_{4,r}(a(q)\phi(q^3)). \tag{5.1}$$

Proof: This lemma is a straightforward corollary of (1.17), (4.7) and (4.10). Next, we apply $S_{4,r}$ with $r = 1, 2$ to (4.21) and use (5.1) to obtain

$$S_{36,9r}\phi(q)^3 - 3S_{4,r}\phi(q)^3 = 2S_{4,r}(\phi(q)^3 - 2\phi(q^3)a(q^2)) = 2S_{4,r}(a(q)\phi(q^3)), \tag{5.2}$$

which is (4.1), as desired. The proof of Theorem 1.2 is now complete. We can do much better, if we realize that (5.1) is an immediate consequence of the following elegant result

$$\phi(q)^3 = \phi(q^3)(a(q) + 2a(q^2) - 2a(q^4)). \tag{5.3}$$

To prove it, we divide both sides by $\phi(q^3)$ and obtain

$$\frac{\phi(q)^3}{\phi(q^3)} = 2a(q^2) - a(q) + 2(a(q) - a(q^4)). \tag{5.4}$$

Using (4.6) and (4.7) in (5.4), we see that (5.3) is equivalent to

$$\frac{\phi(q)^3}{\phi(q^3)} - \frac{\phi(-q)^3}{\phi(-q^3)} = 12q\psi(q^2)\psi(q^6). \tag{5.5}$$

To verify (5.5), I replace q by $-q$ in (4.6) and subtract (4.6) to find with the aid of (4.5) the following

$$\frac{\phi(q)^3}{\phi(q^3)} - \frac{\phi(-q)^3}{\phi(-q^3)} = a(q) - a(-q) = 3(\phi(q)\phi(q^3) - \phi(-q)\phi(-q^3)). \tag{5.6}$$

Subtracting (4.4) from (4.5) we obtain

$$\phi(q)\phi(q^3) - \phi(-q)\phi(-q^3) = 4q\psi(q^2)\psi(q^6). \tag{5.7}$$

Hence,

$$\frac{\phi(q)^3}{\phi(q^3)} - \frac{\phi(-q)^3}{\phi(-q^3)} = 12q\psi(q^2)\psi(q^6), \tag{5.8}$$

as desired. This completes the proof of (5.3). We are now in a position to improve on (5.2). Indeed, it follows from (4.21) and (5.3) that

$$S_{9,0}\phi(q)^3 - 3\phi(q)^3 = 2\phi(q^3)a(q) - 4\phi(q^3)a(q^4). \tag{5.9}$$

Consequently, we can extend Theorem 1.2 as

Theorem 5.2.

$$s(9n) - 3s(n) = 2(1,1,3,0,0,1)(n) - 4(4,3,4,0,4,0)(n). \tag{5.10}$$

It is worthwhile to point out that Theorem 1.1 can be extended in a similar manner as

Theorem 5.3.

$$s(25n) - 5s(n) = 4(2,2,2,-1,1,1)(n) - 8(7,8,8,-4,8,8)(n). \tag{5.11}$$

It is easy to check that $(7,8,8,-4,8,8)(n) = 0$ when $n \equiv 1,2 \bmod 4$. And so (5.11) reduces to (1.3) when $n \equiv 1,2 \bmod 4$. Recalling (2.13), we see that all that is required to prove Theorem 5.3 is

$$\phi(q)^3 - 3\phi(q)\phi(q^5)^2 = 2T(q) - 4\tilde{T}(q), \tag{5.12}$$

where $T(q)$ was defined in (3.2), and

$$\tilde{T}(q) := \sum_{x,y,z} q^{7x^2+8y^2+8z^2-4yz+8zx+8xy}. \tag{5.13}$$

Making easy changes of summation variables $y \to x+y$ and $z \to x+z$ in (3.2) we find that

$$T(q) = \sum_{x,y,z} q^{7x^2+2y^2+2z^2-yz+4zx+4xy}. \tag{5.14}$$

In a similar fashion one can prove that

$$\tilde{T}(q) = \sum_{x\equiv y\equiv z \bmod 2} q^{2x^2+2y^2+2z^2-yz+zx+xy}. \tag{5.15}$$

Combining (3.2), (3.25), (3.27), (3.28) and (5.15), we can easily derive that

$$T(q) - \tilde{T}(q) = 2Z(q) + Z(q) = 6q^2\psi(q)\psi(q^5)\psi(q^{10}). \tag{5.16}$$

Hence we can rewite the right hand side of (5.12) as

$$2T(q) - 4\tilde{T}(q) = 24q^2\psi(q)\psi(q^5)\psi(q^{10}) - 2T(q).$$

Recalling (3.23), we see that (5.12) is equivalent to

$$\phi(q)^3 - 3\phi(q)\phi(q^5)^2 = 16q^2\psi(q)\psi(q^5)\psi(q^{10}) - 2\phi(q^5)\sum_{m,n} q^{2m^2+2nm+3n^2}. \tag{5.17}$$

To prove the above identity we subtract $2\phi(q)\phi(q^5)^2$ from both sides and use (1.25), (2.2) to find that

$$\phi(q)\Pi_2(q) = \phi(q^5)\Pi_1(q) - 4q^2\psi(q)\psi(q^5)\psi(q^{10}). \tag{5.18}$$

Next, we multiply both sides of (5.18) by

$$\frac{E(q^{20})E(q^5)E(q)}{E(q^{10})^2E(q^4)E(q^2)},$$

and use (1.11) to end up with

$$\phi(-q^2)^2 - \phi(-q^{10})^2 = -4q^2\frac{E(q^{20})^3E(q^2)}{E(q^{10})E(q^4)}.$$

Finally, replacing q^2 by q in the above, we deduce that (5.12) is equivalent to

$$\phi(-q)^2 - \phi(-q^5)^2 = -4q\frac{E(q^{10})^3E(q)}{E(q^5)E(q^2)}.$$

Employing (1.7) and (1.15), we see that the last identity is nothing else but (2.1) with q replaced by $-q$. Hence (5.12) is true. This completes my proof of the Theorem 5.3.

6 Bold Proposition

I now proceed to describe the generalization of Theorem 1.2 for any odd prime p. Observe that the ternary quadratic form $x^2+y^2+3z^2+xy$ in this theorem has the discriminants 3^2. We remind the reader that a discriminant of a ternary form $ax^2+by^2+cz^2+dyz+ezx+fxy$ is defined as

$$\frac{1}{2}\det\begin{bmatrix}2a & f & e\\ f & 2b & d\\ e & d & 2c\end{bmatrix}.$$

Using [18] it is easy to check that all ternary forms with the discriminant p^2 belong to the same genus, say $TG_{1,p}$. Let $|\mathrm{Aut}(f)|$ denote the number of integral automorphs of a ternary quadratic form f, and let $R_f(n)$ denote the number of representations of n by f. Let p be an odd prime and $n \not\equiv 3 \bmod 4$. I propose that

$$s(p^2n) - ps(n) = 48\sum_{f\in TG_{1,p}}\frac{R_f(n)}{|\mathrm{Aut}(f)|} - 96\sum_{f\in TG_{1,p}}\frac{R_f\left(\frac{n}{4}\right)}{|\mathrm{Aut}(f)|}. \tag{6.1}$$

Clearly, one wants to know if the parity restriction on n in (6.1) can be removed. In other words, the question is whether a straightforward generalizion of Theorem 5.2 exists. Fortunately, the answer is “yes”. However, the answer involves the second genus of ternary forms $TG_{2,p}$ with discriminant $16p^2$. Note that, in general, there are 12 genera of the ternary forms with the discriminant $16p^2$ [18]. However, when $p \equiv 3 \bmod 4$ one can create $TG_{2,p}$ from some binary quadratic form of discriminant $-p$. It is a well known fact that all binary forms with the discriminant $-p$ belong to the same genus, say BG_p. Let $ax^2 + bxz + cz^2$ be some binary form $\in BG_p$. We can convert it into ternary form

$$f(x, y, z) := 4ax^2 + py^2 + 4cz^2 + 4|b|xz.$$

Next, we extend f to a genus that contains f. This genus is, in fact, $TG_{2,p}$ when $p \equiv 3 \bmod 4$. It can be shown that the map

$$BG_p \to TG_{2,p}$$

does not depend on which specific binary form from BG_p we have choosen as our starting point. I would like to comment that somewhat similar construction was employed in [2] to define the so-called S-genus. Let me illustrate this map for $p = 23$. In this case,

$$BG_{23} = \{x^2 + xz + 6z^2, 2x^2 + xz + 3z^2, 2x^2 - xz + 3z^2\}.$$

Choosing a binary form $x^2 + xz + 6z^2$ as a starting point one gets

$$\begin{aligned}\{x^2 + xz + 6z^2\} &\to \{4x^2 + 23y^2 + 24z^2 + 4xz\} \to \\ &\{4x^2 + 23y^2 + 24z^2 + 4xz, 8x^2 + 23y^2 + 12z^2 + 4xz, 3x^2 \\ &+ 31y^2 + 31z^2 - 30yz + 2zx + 2xy\}.\end{aligned}$$

We note that

$$\begin{aligned}TG_{2,23} : = \{&4x^2 + 23y^2 + 24z^2 + 4xz, 8x^2 + 23y^2 + 12z^2 + 4xz, 3x^2 \\ &+ 31y^2 + 31z^2 - 30yz + 2zx + 2xy\}\end{aligned}$$

is just one out of 12 possible genera of the ternary form with the discriminant $8{,}464$. It is instructive to compare $TG_{2,23}$ and

$$TG_{1,23} := \{x^2+6y^2+23z^2+xy, 2x^2+3y^2+23z^2+xy, 3x^2+8y^2+8z^2-7yz+2zx+2xy\}.$$

Clearly,

$$|TG_{1,23}| = |TG_{2,23}|.$$

Moreover,

$$\begin{aligned}&|\mathrm{Aut}(3x^2 + 8y^2 + 8z^2 - 7yz + 2zx + 2xy)| \\ &\quad = |\mathrm{Aut}(3x^2 + 31y^2 + 31z^2 - 30yz + 2zx + 2xy)| = 12,\end{aligned}$$

$$|\text{Aut}(x^2+6y^2+23z^2+xy)| = |\text{Aut}(4x^2+23y^2+24z^2+4xz)| = 8,$$
$$|\text{Aut}(2x^2+3y^2+23z^2+xy)| = |\text{Aut}(8x^2+23y^2+12z^2+4xz)| = 4.$$

It is a bit less obvious that

$$(3,31,31,-30,2,2)(4n) = (3,8,8,-7,2,2)(n),$$

$$(4,23,24,0,4,0)(4n) = (1,6,23,0,0,1)(n),$$
$$(8,23,12,0,4,0)(4n) = (2,3,23,0,0,1)(n),$$

and that

$$(3,31,31,-30,2,2)(m) = (4,23,24,0,4,0)(m) = (8,12,23,0,0,4)(m) = 0,$$

whenever $m \equiv 1,2 \bmod 4$. I propose that the above properties are, in fact, the signature properties of $TG_{2,p}$. In other words, for any odd prime p there exists an automorphism preserving bijection

$$H : TG_{2,p} \to TG_{1,p},$$

such that , for any $f \in TG_{2,p}$,

$$|\text{Aut}(f)| = |\text{Aut}H(f)|,$$

$$R_f(4n) = R_{H(f)}(n), \tag{6.2}$$

and

$$R_f(m) = 0, \quad \text{when} \quad m \equiv 1,2 \bmod 4. \tag{6.3}$$

Jagy [15] suggested that $TG_{1,p} \cup TG_{2,p}$ does not represent any integer that is quadratic residue mod p when $p \equiv 1 \bmod 4$, and when $p \equiv 3 \bmod 4$ this union does not represent any integer that is a quadratic nonresidue mod p. That is for any $f \in TG_{1,p} \cup TG_{2,p}$

$$R_f(n) = 0,$$

when $(-n|p) = 1$. In addition, he pointed out that $TG_{2,p}$ represents a proper subset of those numbers represented by $TG_{1,p}$. Lastly, he observed that both $TG_{1,p}$ and $TG_{2,p}$ are anisotropic at p. I discuss one more example. This time I choose $p = 17$. Here one has

$$TG_{1,17} := \{3x^2+5y^2+6z^2+yz+2zx+3xy, 3x^2+6y^2+6z^2-5yz+2zx+2xy\},$$

and

$$TG_{2,17} := \{7x^2+11y^2+20z^2-8yz+4zx+6xy, 3x^2+23y^2+23z^2-22yz+2zx+2xy\}.$$

Note that

$$|\mathrm{Aut}(3x^2+5y^2+6z^2+yz+2zx+3xy)| = |\mathrm{Aut}(7x^2+11y^2+20z^2-8yz+4zx+6xy)| = 4,$$

$$|\mathrm{Aut}(3x^2+6y^2+6z^2-5yz+2zx+2xy)| = |\mathrm{Aut}(3x^2+23y^2+23z^2-22yz+2zx+2xy)| = 12,$$

$$(3,23,23,-22,2,2)(4n) = (3,6,6,-5,2,2)(n),$$
$$(7,11,20,-8,4,6)(4n) = (3,5,6,1,2,3)(n),$$
$$(7,11,20,-8,4,6)(m) = (3,23,23,-22,2,2)(m) = 0,$$

whenever $m \equiv 1,2 \bmod 4$. It is worthwhile to point out that there are exactly 12 genera with the discriminant 4,624. Only three of those have the correct cardinality

$$|TG_{2,17}| = 2,$$

$$|\{3x^2+6y^2+68z^2+2xy, \quad 10x^2+11y^2+14z^2+2yz+4zx+10xy\}| = 2,$$
$$|\{5x^2+7y^2+34z^2+2xy, \quad 6x^2+12y^2+17z^2+4xy\}| = 2.$$

Note, however, that

$$|\mathrm{Aut}(3x^2+6y^2+68z^2+2xy)| = |\mathrm{Aut}(10x^2+11y^2+14z^2+2yz+4zx+10xy)| = 4,$$

and

$$|\mathrm{Aut}(5x^2+7y^2+34z^2+2xy)| = |\mathrm{Aut}(6x^2+12y^2+17z^2+4xy)| = 4.$$

And so, $TG_{2,17}$ is a unique genus with the desired properties.

I would like to conclude this discussion of $TG_{2,p}$ by providing a more explicit description valid in three special cases. If $p \equiv 3 \bmod 4$, then $TG_{2,p}$ is the genus that contains

$$4x^2 + py^2 + (p+1)z^2 + 4zx.$$

I remark that the above form was obtained from the principal binary form $x^2 + xz + \frac{p+1}{4}z^2$. If $p \equiv 2 \bmod 3$, then $TG_{2,p}$ is the genus that contains

$$x^2 + \frac{4p+1}{3}y^2 + \frac{4p+1}{3}z^2 + \frac{2-4p}{3}yz + 2zx + 2xy.$$

If $p \equiv 5 \bmod 8$, then $TG_{2,p}$ is the genus that contains

$$8x^2 + \frac{p+1}{2}y^2 + (p+2)z^2 + 2yz + 8zx + 4xy.$$

Observe that the smallest prime to escape the above net of three special cases is $p = 73$. I am now ready to unveil the promised extension of (6.1).

Proposition 6.1. *Let p be an odd prime, then*

$$s(p^2n) - ps(n) = 48 \sum_{f \in TG_{1,p}} \frac{R_f(n)}{|Aut(f)|} - 96 \sum_{f \in TG_{2,p}} \frac{R_f(n)}{|Aut(f)|}. \tag{6.4}$$

The proof of this neat result with $p \geq 7$ is beyond the scope of this paper and will be given in [3]. Note, that (6.1) follows easily from (6.2) to (6.4).
Below I illustrate Proposition 6.1 with some initial examples

$$s(7^2n) - 7s(n) = 6(1,2,7,0,0,1)(n) - 12(4,7,8,0,4,0)(n), \tag{6.5}$$

$$\begin{aligned} s(11^2n) - 11s(n) &= 4(3,4,4,-3,2,2)(n) + 6(1,3,11,0,0,1)(n) \\ &\quad - 8(3,15,15-14,2,2)(n) - 12(4,11,12,0,4,0)(n), \end{aligned} \tag{6.6}$$

$$s(13^2n) - 13s(n) = 12(2,5,5,-3,1,1)(n) - 24(8,7,15,2,8,4)(n), \tag{6.7}$$

$$\begin{aligned} s(17^2n) - 17s(n) &= 12(3,5,6,1,2,3)(n) + 4(3,6,6,-5,2,2)(n) \\ &\quad - 24(7,11,20,-8,4,6)(n) - 8(3,23,23,-22,2,2)(n), \end{aligned} \tag{6.8}$$

$$\begin{aligned} s(19^2n) - 19s(n) &= 6(1,5,19,0,0,1)(n) + 12(4,5,6,5,1,2)(n) \\ &\quad - 12(4,19,20,0,4,0)(n) - 24(7,11,23,-10,6,2)(n), \end{aligned} \tag{6.9}$$

$$\begin{aligned} s(23^2n) - 23s(n) &= 4(3,8,8,-7,2,2)(n) + 6(1,6,23,0,0,1)(n) \\ &\quad + 12(2,3,23,0,0,1)(n) - 8(3,31,31,-30,2,2)(n) \\ &\quad - 12(4,23,24,0,4,0)(n) - 24(8,23,12,0,4,0)(n), \end{aligned} \tag{6.10}$$

Finally, I note that (6.5) implies the following impressive identity

$$\begin{aligned} &8q\psi(-q)E(q^2)^2 S_{7,5}(-q;q^2)_\infty \\ &\quad = \phi(q)^3 + \phi(q^7) \sum_{m,n} (q^{m^2+mn+2n^2} - 2q^{4m^2+4mn+8n^2}). \end{aligned}$$

Acknowledgements I would like to thank Bruce Berndt, Shaun Cooper, Will Jagy, Rainer Schulze-Pillot for their kind interest and helpful discussions.

References

1. P.T. Bateman, *On the representations of a number as the sum of three squares*, Trans. Amer. Math. Soc. 71 (1951), no. 1, 70–101.
2. A. Berkovich, W.C. Jagy, *Ternary Quadratic Forms, Modular Equations and Certain Positivity Conjectures*, in: The Legacy of Alladi Ramakrishnan in the mathematical sciences, (K. Alladi, J. R. Klauder, and C. R. Rao, Eds.), 211–241, Springer, NY, 2010.
3. A. Berkovich, W.C. Jagy, *On representation of an integer as the sum of three squares and the ternary quadratic forms with the discriminants* p^2, $16p^2$, J. of Number Theory 132 (2012), no. 1, 258–274.
4. A. Berkovich, H. Yesilyurt, *Ramanujan's Identities and Representation of Integers by Certain Binary and Quaternary Quadratic Forms*, Ramanujan J. 20 (2009), no. 3, 375–408.
5. B.C. Berndt, *Ramanujan's Notebooks, Part III*, Springer, New York, 1991.
6. B.C. Berndt, *Number Theory in the Spirit of Ramanujan*, Student Mathematical Library, 34 AMS, Providence, RI, 2006.
7. B.C. Berndt, S. Bhargava, F.G. Garvan, *Ramanujan's theories of elliptic functions to alternative bases*, Trans. Amer. Math. Soc. 347 (1995), no. 11, 4163–4244.
8. J.M. Borwein, P.B. Borwein, *A cubic counterpart of Jacobi's identity and AGM*, Trans. Amer. Math. Soc. 323 (1991), no. 2, 691–701.
9. J.M. Borwein, P.B. Borwein, F.G. Garvan, *Some cubic modular identities of Ramanujan*, Trans. Amer. Math. Soc. 343 (1994), no. 1, 35–47.
10. S. Cooper, M.D. Hirschhorn, *Results of Hurwitz type for three squares*, Discrete Math. 274 (2004), 9–24.
11. L.E. Dickson, *Modern Elementary Theory of Numbers*, The University of Chicago Press, 1939.
12. F.G. Garvan, D. Kim, D. Stanton, *Cranks and t-cores*, Invent. Math. 101 (1990), no. 1, 1–17.
13. M.D. Hirschhorn, F.G. Garvan, J.M. Borwein, *Cubic analogues of the Jacobian theta function* $\theta(z, q)$, Canad. J. Math. 45 (1993), no. 4, 673–694.
14. M.D. Hirschhorn, J.A. Sellers, *On representation of a number as a sum of three squares*, Discrete Math. 199 (1999), 85–101.
15. W.C. Jagy, *Private communication.*
16. W.C. Jagy, I. Kaplansky, A. Schiemann, *There are* 913 *regular ternary forms*, Mathematika 44 (1997), 332–341.
17. B.W. Jones, *The Arithmetic Theory of Quadratic Forms*, Mathematical Association of America, 1950.
18. J.L. Lehman, *Levels of positive definite ternary quadratic forms*, Math. of Comput. 58 (1992) no. 197, 399–417.
19. L.C. Shen, *On the modular equations of degree* 3, Proc. Amer. Math. Soc. 122 (1994), no. 4, 1101–1114.

Almost Universal Ternary Sums of Squares and Triangular Numbers

Wai Kiu Chan and Anna Haensch

Abstract For any integer x, let T_x denote the triangular number $\frac{x(x+1)}{2}$. In this paper we give a complete characterization of all the triples of positive integers (α, β, γ) for which the ternary sums $\alpha x^2 + \beta T_y + \gamma T_z$ represent all but finitely many positive integers. This resolves a conjecture of Kane and Sun (Trans Am Math Soc 362:6425–6455, 2010, Conjecture 1.19(i)) and complete the characterization of all almost universal ternary mixed sums of squares and triangular numbers.

Key words and Phrases Primitive spinor exceptions • Ternary quadratic forms • Triangular numbers

Mathematics Subject Classification (2010): Primary 11E12, 11E20

1 Introduction

In the Focused Week on Integral Lattices hosted by the Department of Mathematics in University of Florida in February 2010, the first author presented the result in [1] which is a complete characterization of all triples (α, β, γ) of positive integers for which the polynomials $\alpha T_x + \beta T_y + \gamma T_z$ are almost universal, that is, representing all but finitely many positive integers. Here T_x denotes the triangular number $x(x+1)/2$. This resolves a conjecture made by Kane and Sun in [7, Conjecture 1.19(ii)]. In [7] they also study other types of almost universal mixed sums of squares and triangular numbers. In particular, they determine all almost

W.K. Chan (✉) • A. Haensch
Department of Mathematics and Computer Science, Wesleyan University,
Middletown, CT 06459, USA
e-mail: wkchan@wesleyan.edu; ahaensch@wesleyan.edu

K. Alladi et al. (eds.), *Quadratic and Higher Degree Forms*, Developments
in Mathematics 31, DOI 10.1007/978-1-4614-7488-3_3,

universal ternary sums $\alpha x^2 + \beta y^2 + \gamma T_z$ [7, Theorem 1.6]. They also formulate a conjecture [7, Conjecture 1.19(i)] about almost universal ternary sums of the form $\alpha x^2 + \beta T_y + \gamma T_z$, and an affirmative answer to this conjecture would complete their classification of those almost universal sums. The goal of this paper is to give such a complete characterization via the geometric approach used in [1]. As consequences, we resolve Kane and Sun's conjecture and complete the task of characterizing all almost universal ternary mixed sums of squares and triangular numbers.

The basic difference and similarity between our geometric approach and the theta series approach in [7] is briefly explained in [1]. We would like to add a few more comments for the ternary sums $\alpha x^2 + \beta T_y + \gamma T_z$ we are considering here. Let $r(n)$ be the number of representations of an integer n by $\alpha x^2 + \beta T_y + \gamma T_z$. By the inclusion and exclusion principle $r(n)$ is the $(\beta + \gamma + 8n)$-th coefficient of the linear combination of theta series

$$\theta_{f(x,y,z)} - \theta_{f(x,2y,z)} - \theta_{f(x,y,2z)} + \theta_{f(x,2y,2z)},$$

where $\theta_{f(x,y,z)}$ is the theta series of the diagonal quadratic form $f(x, y, z) = 8\alpha x^2 + \beta y^2 + \gamma z^2$. Using this linear combination of theta series, the authors in [7] determine triples (α, β, γ) for which almost all the $r(n)$ are nonzero. On the other hand, our geometric approach is built upon only one ternary $\mathbb{Z}$-lattice (not necessarily diagonal) on the quadratic $\mathbb{Q}$-space associated to the quadratic form $2\alpha x^2 + \beta y^2 + \gamma z^2$, whose representations of integers of the form $\beta + \gamma + 8n$ will correspond to the representation of n by $\alpha x^2 + \beta T_y + \gamma T_z$. This changes the original problem to the question of determining which ternary quadratic forms represent all sufficiently large integers in an infinite family of positive integers that has some specific arithmetic interest (in our case, they are the integers congruent to $\beta + \gamma$ mod 8). Two powerful tools from the theory of representations of ternary quadratic forms have been proven to be useful in dealing with this kind of questions. The first one is the theorem of Duke and Schulze-Pillot [2] which asserts that a sufficiently large integer is represented by a positive definite ternary quadratic form if that integer is primitively represented by the spinor genus of the quadratic form. The second one is the established theory of primitive spinor exceptions, especially the formulae for computing the primitive relative spinor norm groups given by Earnest et al. in [5] which allows us to determine effectively the set of primitive spinor exceptions of a genus. The readers can find all relevant material in [5] and [9].

Our main results will be divided into four theorems which altogether will characterize all triples (α, β, γ) for which $\alpha x^2 + \beta T_y + \gamma T_z$ are almost universal. Kane and Sun's conjecture only concerns the cases in which $\max\{\text{ord}_2(\beta), \text{ord}_2(\gamma)\}$ is either 3 or 4. However, we opt to present the proofs of all cases since it does not take too much extra effort to do so. This also provides a better understanding of the geometric setting we described above.

2 Preliminaries

Henceforth, the language of quadratic spaces and lattices as in [8] will be adopted. We will follow mostly the notations used in [1]. Any unexplained notation and terminology can be found there and in [8]. All the $\mathbb{Z}$-lattices discussed below are positive definite. If K is a $\mathbb{Z}$-lattice and A is a symmetric matrix, we shall write "$K \cong A$" if A is the Gram matrix for K with respect to some basis of K. The discriminant of K, denoted dK, is the determinant of A. An $n \times n$ diagonal matrix with $a_1, \ldots, a_n$ as the diagonal entries is written as $\langle a_1, \ldots, a_n \rangle$. However, we use the notation $[a_1, \ldots, a_n]$ to denote a quadratic space (over any field) with an orthogonal basis whose associated Gram matrix is the diagonal matrix $\langle a_1, \ldots, a_n \rangle$.

The subsequent discussion involves the computation of the spinor norm groups of local integral rotations and the relative spinor norm groups of primitive representations of integers by ternary quadratic forms. The formulae for all these computations can be found in [3–6]. A correction of some of these formulae can be found in [1, Footnote 1]. The symbol θ always denotes the spinor norm map. If t is an integer represented primitively by gen(K) and p is a prime, then $\theta^*(K_p, t)$ is the primitive relative spinor norm group of the $\mathbb{Z}_p$-lattice K_p. If E is a quadratic extension of $\mathbb{Q}$, $N_p(E)$ denotes the group of local norms from $E_{\mathfrak{p}}$ to $\mathbb{Q}_p$, where $\mathfrak{p}$ is an extension of p to E.

Let a, b, c be relatively prime positive odd integers, m, r, and s be nonnegative integers such that $r \leq s$. Let L be the $\mathbb{Z}$-lattice $\langle 2^{m+1}a, 2^r b, 2^s c \rangle$ in the orthogonal basis $\{e_1, e_2, e_3\}$. It can be shown easily that an integer n is represented by the ternary sum $2^m ax^2 + 2^r bT_y + 2^s cT_z$ if and only if $2^r b + 2^s c + 8n$ is represented by the coset $\omega + 2L$ where $\omega = e_2 + e_3$.

Lemma 2.1. (1) *If $2^m ax^2 + 2^r bT_y + 2^s cT_z$ is almost universal, then L_p represents all p-adic integers for every odd prime p.*

(2) *If L_p represents all p-adic integers for every odd prime p, then*

(i) *$L_p \cong \langle 1, -1, -dL \rangle$ and $\theta(O^+(L_p)) \supseteq \mathbb{Z}_p^\times$ for all odd primes p;*

(ii) *$\mathbb{Q}_2 L$ is anisotropic.*

Proof. See [1, Lemma 2.1]. □

As is explained in [1], the fact that L_p represents all p-adic integers for every odd prime p is equivalent to the following more elementary statement:

> (i) a, b, c are pairwise relatively prime, and (ii) if an odd prime p divides one of $2^m a$, $2^r b$ or $2^s c$, then the negative of the product of the other two is a square modulo p.

Let M be the $\mathbb{Z}$-lattice $\mathbb{Z}\omega + 2L$. Relative to the basis $\{2e_1, 2e_2, \omega\}$, M has the following Gram matrix representation:

$$\begin{pmatrix} 2^{m+3}a & 0 & 0 \\ 0 & 2^{r+2}b & 2^{r+1}b \\ 0 & 2^{r+1}b & 2^r b + 2^s c \end{pmatrix}.$$

The discriminant of M is $2^{m+r+s+5}abc$. Also of interest is the binary sublattice $P = \mathbb{Z}2e_2 + \mathbb{Z}\omega$ whose Gram matrix is

$$\begin{pmatrix} 2^{r+2}b & 2^{r+1}b \\ 2^{r+1}b & 2^r b + 2^s c \end{pmatrix}$$

and $dP = 2^{r+s+2}bc$.

Lemma 2.2. *Suppose that $2^m ax^2 + 2^r bT_y + 2^s cT_z$ is almost universal. Then*

(1) *$r = 0$ when $m > 0$;*
(2) *$2^r b + 2^s c$ is not divisible by 8 and $r < 2$.*

Proof. (1) This is clear; otherwise $2^m ax^2 + 2^r bT_y + 2^s cT_z$ only represents even integers.

(2) Suppose that $2^m ax^2 + 2^r bT_y + 2^s cT_z$ is almost universal. This means that all but finitely many positive integers of the form $2^r b + 2^s c + 8n$ are represented by the coset $\omega + 2L$ and hence by the lattice M. If $2^r b + 2^s c \equiv 0 \bmod 8$, then M_2 represents all 2-adic integers in $8\mathbb{Z}_2$ because $\mathbb{Z}$ is dense in $\mathbb{Z}_2$. But then M_2 would be isotropic and this contradicts Lemma 2.1(2)(ii).

If $r = 2$, then $m = 0$, $s > 2$ and $2^r b + 2^s c \equiv 4 \mod 8$. In this case,

$$M_2 \cong \langle 4b + 2^s c, 8a, 2^s(4b + 2^s c)bc \rangle$$

which does not represent every element in $4\mathbb{Z}_2^\times$. Therefore, $ax^2 + 4bT_y + 2^s cT_z$ is not almost universal. □

Lemma 2.3. *Suppose that $2^{m+1}ax^2 + 2^r by^2 + 2^s cz^2$ represents all p-adic integers for every odd prime p, and that conditions (1) and (2) in Lemma 2.2 are satisfied. Then*

(1) *Every positive integer of the form $2^r b + 2^s c + 8n$ is represented primitively by gen(M);*
(2) *If t is a primitive spinor exception of gen(M), then $\mathbb{Q}(\sqrt{-tdM})$ is either $\mathbb{Q}(\sqrt{-1})$ or $\mathbb{Q}(\sqrt{-2})$.*

Proof. (1) By virtue of Lemma 2.1(2)(i), it suffices to check that M_2 primitively represents all 2-adic integers of the form $2^r b + 2^s c + 8n$. This is clear when $2^r b + 2^s c$ is odd. Suppose that it is even. We first consider the case $r = s = 0$. Since $b + c \not\equiv 0 \bmod 8$, $P_2^{\frac{1}{2}}$ is a unimodular $\mathbb{Z}_2$-lattice, which is proper when $b + c \equiv 2 \bmod 4$ and improper when $b + c \equiv 4 \bmod 8$. In either case, it is direct to check that P_2 primitively represents all 2-adic integers of the form $b + c + 8n$.

Suppose that $r = 1$. If $s > 1$, then

$$M_2^{\frac{1}{2}} \cong \langle 4a \rangle \perp \langle b + 2^{s-1}c \rangle \perp \langle 2^{s+1}bc(b + 2^{s-1}c) \rangle.$$

The binary sublattice $\langle 4a, b + 2^{s-1}c\rangle$ represents all units that are congruent to $b + 2^{s-1}c$ mod 4. Hence M_2 primitively represents all integers of the form $2b + 2^s c + 8n$.

However, when $r = 1$ and $s = 1$, ord$_2(b + c) = 1$ by our assumption. So, $bc \equiv 1$ mod 4 and $u := (b + c)/2$ is a unit in $\mathbb{Z}_2$. In this case,

$$M_2^{\frac{1}{4}} \cong \langle 2a, u, ubc\rangle$$

which represents all units in $\mathbb{Z}_2$. Therefore, M_2 primitively represents all 2-adic integers of the form $2b + 2c + 8n$.

(2) This is the same as the proof of [1, Lemma 2.2(2)]. □

Lemma 2.4. *Suppose that L_p represents all p-adic integers for every odd prime p. If we are not in the exceptional case where $r = s = 0$ and $b + c \equiv 4 \mod 8$, and if $2^r b + 2^s c + 8n$ is not a primitive spinor exception of gen(M) for all $n \geq 0$, then $2^m ax^2 + 2^r bT_y + 2^s cT_z$ is almost universal.*

Proof. The proof is parallel to that of [1, Lemma 2.3]. We leave the detail to the readers. Note that since we are not in the exceptional case ($r = s = 0$ and $b+c \equiv 4$ mod 8), any representation of $2^r b + 2^s c + 8n$ by M must lie in $\omega + 2L$. □

3 Main Results

We continue to assume that a, b, c are relatively prime positive odd integers. By Lemma 2.2, we only need to address the following four cases:

(i) $m = 0$, $r = 0$, and $s \geq 1$.
(ii) $m = 0$, $r = 1$, $s \geq 1$, and ord$_2(b + c) = 1$ if $s = 1$.
(iii) $m > 0$, $r = 0$, and $s > 0$.
(iv) $m \geq 0$, $r = s = 0$, and ord$_2(b + c) \leq 2$.

They will be covered by Theorems 3.1–3.4 accordingly. We will provide full detail in the proof of the first theorem. For the proofs of the other three theorems, since the strategy and techniques involved are the same or very similar to the first one, we will only present the argument that may be less transparent to the readers. In below, the squarefree part of an integer α is denoted by sf(α).

Theorem 3.1. *Suppose that $s \geq 1$. Then $ax^2 + bT_y + 2^s cT_z$ is almost universal if and only if $2ax^2 + by^2 + 2^s cz^2$ represents all p-adic integers over $\mathbb{Z}_p$ for every odd p, and one of the following holds:*

(1) *s is odd, or $s = 2$;*
(2) *sf(abc) is divisible by a prime $p \equiv 5, 7 \mod 8$;*
(3) *$bc \not\equiv 1 \mod 8$;*
(4) *$\frac{sf(abc)-(b+2^s c)}{8}$ is represented by $ax^2 + bT_y + 2^s cT_z$.*

Proof. We assume throughout that $2ax^2 + by^2 + 2^s cz^2$ represents all p-adic integers over $\mathbb{Z}_p$ for every odd p. Then by Lemma 2.3(1), gen(M) represents all $b + 2^s c + 8n$ primitively. Note that

$$M_2 \cong \langle b + 2^s c, 8a, (b + 2^s c)2^{s+2}bc\rangle.$$

In below, unless stated otherwise, t is always assumed to be an odd primitive spinor exception of gen(M) and E is the quadratic field $\mathbb{Q}(\sqrt{-tdM})$. By Lemma 2.3(2), E is either $\mathbb{Q}(\sqrt{-1})$ or $\mathbb{Q}(\sqrt{-2})$. The strategy of the proof is to show that under (1), (2), or (3), gen(M) does not have any primitive spinor exception of the form $b + 2^s c + 8n$ and hence Lemma 2.4 applies. At the end we show that if (1), (2) and (3) all fail, then $ax^2 + bT_y + 2^s cT_z$ is almost universal if and only if (4) holds.

Suppose that s is odd. Since $\text{ord}_2(-tdM_2) = s + 5$ which is even, E must be $\mathbb{Q}(\sqrt{-1})$. But by Earnest et al. [5, Theorem 2(b)(iv)] $\theta^*(M_2, t) \neq N_2(E)$. Therefore, gen$(M)$ does not have any odd primitive spinor exceptions.

If $s = 2$, M_2 is of *Type E* as defined in [4, page 531] and so $\theta(O^+(M_2)) = \mathbb{Q}_2^\times$ [3, Theorem 2.2]. Together with Lemma 2.1(2)(i) these show that gen(M) has only one spinor genus, and hence in this case gen(M) does not have any primitive spinor exception at all.

Now let us assume that s is even and $s \geq 4$. Then E must be $\mathbb{Q}(\sqrt{-2})$. At the primes p where $\left(\frac{-2}{p}\right) = -1$, $\theta(O^+(M_p)) \subseteq N_p(E)$ if and only if $p \nmid \text{sf}(abc)$, as a consequence of [5, Theorem 1] and Lemma 2.1(2)(i). This means that if sf(abc) is divisible by a prime $p \equiv 5, 7 \mod 8$, then gen(M) does not have any odd primitive spinor exceptions.

Suppose that (1) and (2) do not hold, but (3) holds. Let t be a primitive spinor exception of gen(M) of the form $t = b + 2^s c + 8n$. Since $E = \mathbb{Q}(\sqrt{-tdM}) = \mathbb{Q}(\sqrt{-2})$, it follows that $tabc \equiv 1 \mod 8$, and hence $bc \equiv 3 \mod 8$ and $ta \equiv 3 \mod 8$. Over $\mathbb{Z}_2$, $M_2^t \cong \langle 1, 8at, 2^{s+2}bc\rangle$. Let U be the $\mathbb{Z}_2$-lattice $\langle 1, 24\rangle$. Then $\theta(O^+(M_2)) \supseteq \theta(O^+(U)) = \{1, 6, -1, -6\}\mathbb{Q}_2^{\times 2}$ by Earnest and Hsia [3, 1.9], and so $\theta(O^+(M_2))$ is not contained in $N_2(E) = \{1, 2, 3, 6\}\mathbb{Q}_2^{\times 2}$. As a result, none of the positive integers $b + 2^s c + 8n$ is a primitive spinor exception of gen(M).

Let us suppose further that (1), (2), and (3) all fail. Then $bc \equiv 1 \mod 8$ and hence $b \equiv c \equiv 1$ or $3 \mod 8$. So the $\mathbb{Q}_2$ quadratic space underlying M_2 is isometric to either $[2a, 1, 1]$ or $[2a, 3, 3]$ with $a \equiv 1$ or $3 \mod 8$. They are isotropic when $a \equiv 3 \mod 8$ and $a \equiv 1 \mod 8$, respectively, and this is impossible by Lemma 2.1(2)(ii). Therefore, we must have $a \equiv c \mod 8$, implying $ac \equiv 1 \mod 8$, and hence $ab \equiv 1 \mod 8$ as well.

Now we claim that sf(abc) is a primitive spinor exception of gen(M). Since sf$(abc) \equiv b \equiv b + 2^s c \mod 8$, sf$(abc)$ is represented primitively by gen(M). Without causing any confusion, let E denote the field $\mathbb{Q}(\sqrt{-\text{sf}(abc)dM})$, which is just $\mathbb{Q}(\sqrt{-2})$. When p is an odd prime, it follows from [5, Theorem 1] that $\theta(O^+(M_p)) \subseteq N_p(E)$ and $\theta^*(M_p, \text{sf}(abc)) = N_p(E)$. For the prime 2, $M_2 \cong \langle b, 8a, 2^{s+2}c\rangle$ which is not of *Type E*. Hence $\theta(O^+(M_2))$ can be computed as before, and the calculation shows that $\theta(O^+(M_2))$ is exactly equal to $N_2(E)$.

By Earnest et al. [5, Theorem 2(c)(iii)], we see that $\theta^*(M_2, \mathrm{sf}(abc)) = N_2(E)$. This proves our claim that $\mathrm{sf}(abc)$ is a primitive spinor exception of $\mathrm{gen}(M)$.

Suppose that $\mathrm{sf}(abc)$ is represented by M. If $b+2^sc+8n$ is not a primitive spinor exception of $\mathrm{gen}(M)$, then $b+2^sc+8n$ is represented primitively by $\mathrm{spn}(M)$ and hence it is represented by M when n is sufficiently large. Otherwise, $b+2^sc+8n$ must be a square multiple of $\mathrm{sf}(abc)$ and hence $b+2^sc+8n$ is represented by M. So we can conclude that $b+2^sc+8n$ is represented by M for almost all n. But any representation of $b+2^sc+8n$ must be in $\omega+2L$; hence $ax^2+bT_y+2^sT_z$ is almost universal.

Conversely, suppose that $\mathrm{sf}(abc)$ is not presented by M. Then there exist, as shown in [9], infinitely many primes p such that $p^2\mathrm{sf}(abc)$ is not represented by M. For each such p, we have $\mathrm{sf}(abc)p^2 \equiv b \equiv b+2^sc \mod 8$. Therefore, $n := \frac{\mathrm{sf}(abc)p^2-(b+2^sc)}{8}$ is a positive integer for which $b+2^sc+8n$ is not represented by $\omega+2L$. Therefore, $ax^2+bT_y+2^scT_z$ cannot be almost universal. □

Theorem 3.2. *Suppose $s \geq 1$ and $ord_2(b+c) = 1$ if $s = 1$. Then $ax^2+2bT_y+2^scT_z$ is almost universal if and only if $2ax^2+2by^2+2^scz^2$ represents all p-adic integers over $\mathbb{Z}_p$ for every odd p, and one of the following holds:*

(1) *s is even or $s = 1$;*
(2) *$sf(abc)$ is divisible by a prime $p \equiv 3 \mod 4$;*
(3) *$\frac{sf(abc)-(b+2^{s-1}c)}{4}$ is represented by $ax^2+2bT_y+2^scT_z$.*

Proof. As in the proof of Theorem 3.1, we assume throughout that $2ax^2+2by^2+2^scz^2$ represents all p-adic integers over $\mathbb{Z}_p$ for every odd p. If $s = 1$, then

$$M_2^{\frac{1}{4}} \cong \langle 2a, \epsilon, \epsilon bc\rangle,$$

where $\epsilon := \frac{b+c}{2}$ is a unit, since $b+c \equiv 2 \mod 4$ under our assumption. It follows from [5, Theorem 2(c)] that $\theta(O^+(M_2^{\frac{1}{4}})) \not\subseteq N_2(\mathbb{Q}(\sqrt{-tdM_2^{\frac{1}{4}}}))$ for any odd t, and hence $2b+2c+8n$ is not a primitive spinor exception of $\mathrm{gen}(M)$ for any n. When s is even, it is more convenient to work with $G := M^{\frac{1}{2}}$. Over $\mathbb{Z}_2$,

$$G_2 \cong \langle 4a, b+2^{s-1}c, (b+2^{s-1}c)2^{s+1}bc\rangle.$$

By Earnest et al. [5, Theorem 2(c)(i)], either $\theta(O^+(G_2)) \not\subseteq N_2(\mathbb{Q}(\sqrt{-tdG_2}))$ or $\theta^*(G_2, t) \neq N_2(\mathbb{Q}(\sqrt{-tdG_2}))$ for any odd t. Therefore, $\mathrm{gen}(G)$ cannot have any odd primitive spinor exceptions and hence $2b+2^sc+8n$ is not a primitive spinor exception of $\mathrm{gen}(M)$.

Suppose that (1) is false. Since s is odd, we know that $\mathbb{Q}(\sqrt{-tdG}) = \mathbb{Q}(\sqrt{-1})$ for any odd primitive spinor exception t of $\mathrm{gen}(G)$. By Earnest et al. [5, Theorem 1], we see that if $p \equiv 3 \mod 4$, then p does not divide $\mathrm{sf}(abc)$. Therefore, if (2) holds, then $2b+2^sc+8n$ is not a primitive spinor exception of $\mathrm{gen}(M)$.

Now we assume that both (1) and (2) do not hold. We claim that $\mathrm{sf}(abc)$ is a primitive spinor exception of $\mathrm{gen}(G)$. It is clear that G_p represents $\mathrm{sf}(abc)$ primitively for every odd prime p. At the prime 2, $G_2 \cong \langle t, 4a, t2^{s+1}bc\rangle$, where $t = b + 2^{s-1}c \equiv b \equiv 1 \mod 4$. The binary component $\langle t, 4a\rangle$ represents all 2-adic units that are congruent to t mod 4. Since $\mathrm{sf}(abc) \equiv 1 \equiv t \mod 4$, G_2 represents $\mathrm{sf}(abc)$ primitively.

We then need to show that $\theta(O^+(G_p)) \subseteq N_p(E) = \theta^*(G_p, \mathrm{sf}(abc))$ for all primes p, where $E = \mathbb{Q}(\sqrt{-\mathrm{sf}(abc)dG})$ which is $\mathbb{Q}(\sqrt{-1})$. The argument is similar to the proof of Theorem 3.1. For odd primes, we apply [5, Theorem 1]. For the prime 2, since $G_2^t \cong \langle 1, 4at, 2^{s+1}bc\rangle$ is not of *Type E*, $\theta(O^+(G_2)) = Q(P(U))Q(P(W))\mathbb{Q}_2^{\times 2}$ by Earnest and Hsia [3, Theorem 2.7], with $U \cong \langle 1, 4at\rangle$ and $W \cong 4at\langle 1, 2^{s-1}bc\rangle$. It follows from [3, 1.9] that $\theta(O^+(G_2))$ is always inside $N_2(E) = \{1, 2, 5, 10\}\mathbb{Q}_2^{\times 2}$. We then apply [5, Theorem 2(b)] to show that $\theta^*(G_2, \mathrm{sf}(abc)) = N_2(E)$.

The last step is to show that if (1) and (2) do not hold, then (3) holds if and only if $ax^2 + 2bT_y + 2^s cT_z$ is almost universal. This can be done as in the last step of the proof of Theorem 3.1. □

Theorem 3.3. *Suppose that $m > 0$ and $s > 0$. Then $2^m ax^2 + bT_y + 2^s cT_z$ is almost universal if and only if $2^{m+1}ax^2 + by^2 + 2^s cz^2$ represents all p-adic integers over $\mathbb{Z}_p$ for every odd p, and one of the following holds:*

(1) *m is even and $s = 1$ or 2; or, $m = 1$ and s is odd;*
(2) *sf(abc) is divisible by a prime p for which $\left(\frac{-\delta}{p}\right) = -1$, where $\delta = 1$ or 2 when $s + m$ is odd or even accordingly;*
(3) *$(b + 2^s c)$sf$(abc) \not\equiv 1 \mod 8$;*
(4) *$\frac{sf(abc)-(b+2^s c)}{8}$ is represented by $2^m ax^2 + bT_y + 2^s cT_z$.*

Proof. Again, we assume throughout that $2^{m+1}ax^2 + by^2 + 2^s cz^2$ represents all p-adic integers over $\mathbb{Z}_p$ for every odd p. We first show that under condition (1), $\mathrm{gen}(M)$ does not have any odd primitive spinor exception. When $s = m = 1$,

$$M_2 \cong \langle b + 2c, (b + 2c)2^3bc, 2^4a\rangle$$

which is of *Type E*. This means $\theta(O^+(M_2)) = \mathbb{Q}_2^\times$ and hence $\mathrm{gen}(M)$ has only one spinor genus. Consequently, $\mathrm{gen}(M)$ cannot have any primitive spinor exceptions.

When $s = 1$ and $m > 0$ is even, then $m + 3 > s + 2 = 3$ and

$$M_2 \cong \langle b + 2c, (b + 2c)2^3bc, 2^{m+3}a\rangle.$$

If t is an odd primitive spinor exception of $\mathrm{gen}(M)$, then $\mathbb{Q}(\sqrt{-tdM})$ must be $\mathbb{Q}(\sqrt{-1})$. However, by Earnest et al. [5, Theorem 2(b)], $\theta^*(M_2, t) \neq N_2(\mathbb{Q}(\sqrt{-tdM}))$; hence $\mathrm{gen}(M)$ does not have any odd primitive spinor exception. The two other cases in (1), namely when $m = 1$ and $s \geq 3$ is odd, and when $s = 2$ and $m > 0$ is even, are done similarly but by Earnest et al. [5, Theorem 2(c)] instead, since $\mathrm{ord}_2(tdM) = s + 6$ and $m + 7$, respectively, are odd.

We then move on to show that if (1) fails but (2) holds, then $\text{gen}(M)$ also does not have any odd primitive spinor exception. After that, we assume that both (1) and (2) fail but (3) holds, and show that none of the integers $b + 2^s c + 8n$ is a primitive spinor exception of $\text{gen}(M)$. These two steps can be done as in the proofs of the previous theorems.

Now, assume that (1), (2), and (3) all fail. We claim that $\text{sf}(abc)$ is a primitive spinor exception. First we need to show that $\text{sf}(abc)$ is primitively represented by $\text{gen}(M)$. There is no problem presented at the odd primes because of Lemma 2.1(2)(i). Over $\mathbb{Z}_2$, $M_2 \cong \langle b + 2^s c, 2^{m+3}a, (b + 2^s c)2^{s+2}bc\rangle$ where $m + 3 > 3$ and $s + 2 \geq 3$. Since $(b + 2^s c)\text{sf}(abc) \equiv 1 \mod 8$, M_2 represents $\text{sf}(abc)$ primitively.

Let E be the field $\mathbb{Q}(\sqrt{-\text{sf}(abc)dM})$ which is $\mathbb{Q}(\sqrt{-1})$ when $s + m$ is odd, and $\mathbb{Q}(\sqrt{-2})$ when $s + m$ is even. Using Lemma 2.1(2)(i) and [5, Theorem 1], one can show that $\theta(O^+(M_p)) \subseteq N_p(E) = \theta^*(M_p, \text{sf}(abc))$ for every odd prime p. For the prime 2, we first treat the case $s + 2 \neq m + 3$. This case is further divided into two subcases according to $s + m$ is odd or even. When $s + m$ is odd and $m + 3 > s + 2$, for simplicity we let $t = b + 2^s c$ in the following discussion, we have

$$M_2 \cong \langle t, t2^{s+2}bc, 2^{m+3}a\rangle \cong \langle t, 2^{s+2}a, 2^{m+3}a\rangle.$$

Observe that M_2 is not of *Type E*; otherwise $s = 1$ or $m - s = 0, 2$ and both are impossible under our assumption. Let

$$U \cong \langle 1, 2^{s+2}at\rangle \text{ and } W \cong 2^{s+2}at\langle 1, 2^{m-s+1}\rangle.$$

Then $\theta(O^+(M_2)) = Q(P(U))Q(P(W))\mathbb{Q}_2^{\times 2}$. Since $at \equiv 1 \mod 4$, it follows from [3, Remark 1.9] that both $Q(P(U))\mathbb{Q}_2^{\times 2}$ and $Q(P(W))\mathbb{Q}_2^{\times 2}$ are inside $\{1, 2, 5, 10\}\mathbb{Q}_2^{\times 2}$ which is the norm group $N_2(E)$ (note that $E = \mathbb{Q}(\sqrt{-1})$ when $s + m$ is odd). Further, we are in the situation of [5, Theorem 2(b)], where the "r" there is our $s + 2$ which is strictly bigger than 3. Thus $\theta^*(M_2, \text{sf}(abc)) = N_2(E)$ and hence $\text{sf}(abc)$ is a primitive spinor exception of $\text{gen}(M)$. The case where $s + m$ is odd and $s + 2 > m + 3$ is done similarly, with the Jordan decomposition of M_2 switched to $M_2 \cong \langle t, 2^{m+3}a, 2^{s+2}a\rangle$. The case where $s + m$ is even follows the same argument but with E changed to $\mathbb{Q}(\sqrt{-2})$.

Finally, there is the case when $s + 2 = m + 3$ and

$$M_2 \cong \langle t\rangle \perp 2^{s+2}\langle a, tbc\rangle.$$

Note that $s + 2 > 3$ because $m > 0$, and that $E = \mathbb{Q}(\sqrt{-1})$ since $s + m$ is odd. The spinor norm group $\theta(O^+(M_2))$ is computed using [3, Theorem 3.14(iv)] and [4, Theorem 1.2] (see [1, Footnote 1] for corrections), and the computation shows that $\theta(O^+(M_2)) = \{\gamma \in \mathbb{Q}_2^\times : (\gamma, -1) = 1\}$ which is equal to $N_2(E)$. Further, by Earnest et al. [5, Theorem 2(b)] we know that $\theta^*(M_2, \text{sf}(abc)) = N_2(E)$, since $\text{sf}(abc)$ will not be in any ideal generated by 2.

So we establish that $\mathrm{sf}(abc)$ is a primitive spinor exception of $\mathrm{gen}(M)$. The last step is to show that (4) holds if and only if $2^m ax^2 + bT_y + 2^s cT_z$ is almost universal. This is done by the same argument used in the proof of Theorem 3.1. □

Theorem 3.4. *Suppose that $m \geq 0$ and $ord_2(b+c) \leq 2$. Then $2^m ax^2 + bT_y + cT_z$ is almost universal if and only if $2^{m+1}ax^2 + by^2 + cz^2$ represents all p-adic integers over $\mathbb{Z}_p$ for every odd p, and one of the following holds:*

(1) *$ord_2(b+c) = 2$;*
(2) *m is odd or $m = 0$;*
(3) *$sf(abc)$ is divisible by a prime $p \equiv 3 \mod 4$;*
(4) *$\frac{(b+c)}{2} sf(abc) \not\equiv 1 \mod 4$;*
(5) *$\frac{2sf(abc)-(b+c)}{8}$ is represented by $ax^2 + 2bT_y + 2^s cT_z$.*

Proof. As before, we assume throughout that $2^{m+1}ax^2 + by^2 + cz^2$ represents all p-adic integers over $\mathbb{Z}_p$ for every odd p.

Suppose that $\mathrm{ord}_2(b+c) = 2$. In this case, a representation of an integer of the form $b+c+8n$ by the lattice M is not necessarily from the coset $\omega + 2L$. To amend this, we consider the sublattice R of M such that $R_p = M_p$ for every odd prime p, and

$$R_2 = \mathbb{Z}_2[2e_1, bw - (b+c)e_2, w] \cong \langle 2^{m+3}a, (b+c), (b+c)bc\rangle.$$

Since $b+c \equiv 4$ mod 8, $bc \equiv 3$ mod 8 and hence the binary $\mathbb{Z}_2$-lattice $\langle (b+c), (b+c)bc\rangle$ represents all elements in $4\mathbb{Z}_2^\times$. This shows that all positive integers of the form $b+c+8n$ are represented primitively by $\mathrm{gen}(R)$. But it is straightforward to show that the norm ideal of the $\mathbb{Z}$-lattice $R \cap 2L$ is contained in $8\mathbb{Z}$, and hence any representation of an integer $b+c+8n$ by R must lie inside the subset $R \cap (\omega + 2L)$. Therefore, we can replace M by R in our discussion. It is more convenient to work with $H := R^{\frac{1}{4}}$, and

$$H_2 \cong \langle u, ubc, 2^{m+1}a\rangle$$

where $u := (b+c)/4$, which is a 2-adic unit.

When m is even, $\mathrm{ord}_2(-tdH)$ is odd for any odd integer t. By Earnest et al. [5, Theorem 2(c)], we see that $\mathrm{gen}(H)$ does not have any odd primitive spinor exception, and hence $b+c+8n$ cannot be a primitive spinor exception of $\mathrm{gen}(R)$ for all positive integers n. When m is odd, $\mathbb{Q}(\sqrt{-tdH})$ is the field $\mathbb{Q}(\sqrt{-1})$ for any odd primitive spinor exception t of $\mathrm{gen}(H)$. In this case,

$$H_2^u \cong \langle 1, 3, 2^{m+1}au\rangle \cong A(1,4) \perp \langle 2^{m+1}au\rangle$$

with both Jordan components having even order (see [3, Definition 3.1]). It then follows from [4, 1.2(b)(2)] that $\theta(O^+(H_2)) = \mathbb{Z}_2^\times \mathbb{Q}_2^{\times 2}$ which is not the norm group of $\mathbb{Q}_2(\sqrt{-1})$. This shows that H cannot have any odd primitive spinor exception, and hence none of those integers $b+c+8n$ is a primitive spinor exception of $\mathrm{gen}(R)$. By Duke and Schulze-Pillot [2, Corollary], $2^m ax^2 + bT_y + cT_z$ is almost universal when $\mathrm{ord}_2(b+c) = 2$.

Suppose that $\mathrm{ord}_2(b+c)=1$. Let G be the lattice $M^{\frac{1}{2}}$, and

$$G_2 \cong \langle 2^{m+2}a, \epsilon, \epsilon bc\rangle,$$

where $\epsilon := (b+c)/2$ is a 2-adic unit. Since $b+c \equiv 2 \mod 4$, we have $bc \equiv 1 \mod 4$. If $m=0$, the spinor norm of $O^+(G_2)$ is computed using [4, 1.2(b)(3)] (see the correction in [1, Footnote 1]), and the computation shows that $\theta(O^+(G_2)) = \mathbb{Q}_2^\times$. This means that $\mathrm{gen}(G)$ has only one spinor genus. In particular, $\mathrm{gen}(G)$ does not have any primitive spinor exception, and hence none of the integers $b+c+8n$ is a primitive spinor exception of $\mathrm{gen}(M)$.

Suppose that m is odd. Then $\mathrm{ord}_2(-tdG)$ is odd for any odd integer t. It follows from [5, Theorem 2(c)] that $\theta(O^+(G_2)) \not\subseteq N_2(\mathbb{Q}(\sqrt{-tdG}))$. Therefore, $\mathrm{gen}(G)$ does not have any odd primitive spinor exception, and hence once again for any $n \geq 0$, $b+c+8n$ is never a primitive spinor exception of $\mathrm{gen}(M)$.

If (1) and (2) fail but (3) holds, then the usual argument shows that $b+c+8n$ is not a primitive spinor exception for $\mathrm{gen}(M)$. Suppose now that (1), (2), and (3) all fail. If t is a primitive spinor exception of $\mathrm{gen}(G)$ of the form $\epsilon + 4n$, then $\mathbb{Q}(\sqrt{-tdG}) = \mathbb{Q}(\sqrt{-1})$ and hence $t\mathrm{sf}(abc) \equiv \epsilon\mathrm{sf}(abc) \equiv 1 \mod 4$. Thus, if (4) holds, then $\mathrm{gen}(M)$ does not have any primitive spinor exception of the form $b+c+8n$.

When (1) to (4) all fail, we need to show that $\mathrm{sf}(abc)$ is a primitive spinor exception of $\mathrm{gen}(G)$. As before, there is no problem at the odd primes. Since $\epsilon = (b+c)/2 \equiv 1 \mod 4$, we have $b+c \equiv 2 \mod 8$ and hence $bc \equiv 1 \mod 8$ (not only $\equiv 1 \mod 4$). Therefore $G_2 \cong \langle \epsilon, \epsilon, 2^{m+2}a\rangle \cong \langle 1, 1, 2^{m+2}a\rangle$ which represents $\mathrm{sf}(abc)$ because $\mathrm{sf}(abc) \equiv \epsilon \mod 4$, and then $\theta(O^+(G)) = N_2(E)$ by Earnest and Hsia [4, 1.2(b)(3)], where $E = \mathbb{Q}(\sqrt{-\mathrm{sf}(abc)dG}) = \mathbb{Q}(\sqrt{-1})$. To show that $\theta^*(G_2, \mathrm{sf}(abc)) = N_2(E)$, we need to analyze the conditions in [5, Theorem 2(b)]. The "r" in that theorem is 0, and the lattices K and K' are $\langle 2^{-2}\epsilon, \epsilon, 2^{m+2}a\rangle$ and $\langle \epsilon, \epsilon, 2^{m+2}a\rangle$ respectively. Clearly, $\theta(O^+(K')) = \theta(O^+(G_2)) \subseteq N_2(E)$. For $\theta(O^+(K))$, note that $K^{4\epsilon} \cong \langle 1, 4, 2^{m+4}a\epsilon\rangle$ which is not of type E. Hence $\theta(O^+(K))$ can be computed as we did several times before, and the computations shows that $\theta(O^+(K)) = \{1,5\}\mathbb{Q}_2^{\times 2} \subseteq N_2(E)$. Therefore, by Earnest et al. [5, Theorem 2(b)(iii)], $\theta^*(G_2, \mathrm{sf}(abc))$ is equal to $N_2(E)$.

The last step is to show that when (1) to (4) fail, $2^m ax^2 + bT_y + cT_z$ is almost universal if and only if (5) holds. This is done as in the proofs of the previous theorems. □

The following corollary is conjectured by Kane and Sun in [7, Conjecture 1.19(i)]. We state it in an equivalent form that is more in line with our presentation. In below, the odd part of an integer α is denoted by α'.

Corollary 3.5. *When $s \in \{3,4\}$, if $2^m ax^2 + 2^r bT_y + 2^s cT_z$ is almost universal, then one of the following holds:*

(a) $4 \mid (2^r b + 2^s c)$ *or sf*$(abc) \not\equiv (2^r b + 2^s c)' \mod 2^{3-\nu}$ *where* $\nu := \mathit{ord}_2(2^r b + 2^s c) < 2$.

(b) *If* $sf(2^{m+r+s}abc) \equiv (2^r b + 2^s c) \mod 2$ *then* $sf(abc)$ *is divisible by a prime* $p \equiv 5, 7 \mod 8$. *Otherwise,* $sf(abc)$ *is divisible by a prime* $p \equiv 3 \mod 4$.

(c) $2^{m+3}ax^2 + 2^r by^2 + 2^s cz^2 = 2^\nu sf(abc)$ *has integral solutions with* y *and* z *odd.*

(d) $\begin{cases} s = 3 \textit{ implies } 4 \nmid 2^m a \textit{ and } 2 \nmid 2^r b, \\ s = 4 \textit{ implies } 2 \nmid 2^m a \textit{ and } 2^m a \not\equiv 2^r b \mod 8. \end{cases}$

Proof. Suppose that $2^m ax^2 + 2^r bT_y + 2^s cT_z$ is almost universal. When $r = 1$, then m is necessarily equal to 0. If, in addition, $s = 4$, then (d) is always true. Let us consider the case $r = 1$ and $s = 3$. Theorem 3.2 applies to this case and either (2) or (3) there holds. But it is clear that (2) implies (b) and (3) implies (c).

Suppose that $r = 0$ and $s = 3$. When $m = 0$, $4 \nmid a$ and $2 \nmid b$ which implies that (d) always holds. The case $m > 0$ is covered by Theorem 3.3. One of (2), (3), and (4) there must hold. It is then clear that (b), (a), or (c) is true accordingly.

The case $r = 0$ and $s = 4$ can be verified similarly. If $m > 0$, we apply Theorem 3.3 and it is easy to see that one of (a), (b), and (c) must be true in this case. When $m = 0$ we apply Theorem 3.1 instead. One of (2), (3), and (4) in Theorem 3.1 must be true. It is clear that (2) and (b) are the same; so are (4) and (c). Suppose that (3) of Theorem 3.1 holds, that is, $bc \not\equiv 1 \mod 8$. We may assume that (a) fails; otherwise we are done. Then $\mathrm{sf}(abc) \equiv b + 2^4 c \equiv b \mod 8$, that is, $ac \equiv 1 \mod 8$. This implies that $ab \not\equiv 1 \mod 8$ and hence (d) holds. □

Acknowledgements The first author extends his thanks to the Department of Mathematics in University of Florida for their hospitality during his visit for the Focused Week on Integral Lattices in February 2010. Both authors thank Maria Ines Icaza for her helpful comments and discussion. They also thank the referee for his/her patience and helpful comments.

References

1. W.K. Chan and B.-K. Oh, *Almost universal ternary sums of triangular numbers*, Proc. Amer. Math. Soc. **137** (2009), 3553–3562.
2. W. Duke and R. Schulze-Pillot, *Representation of integers by positive ternary quadratic forms and equidistribution of lattice points on ellipsoids*, Invent. Math. **99** (1990), no. 1, 49–57.
3. A.G. Earnest and J.S. Hsia, *Spinor norms of local integral rotations II*, Pacific J. Math. **61** (1975), no. 1, 71–86.
4. A.G. Earnest and J.S. Hsia, *Spinor genera under field extensions II: 2 unramfied in the bottom field*, Amer. J. Math. **100** (1978), no. 3, 523–538.
5. A.G. Earnest, J.S. Hsia and D.C. Hung, *Primitive representations by spinor genera of ternary quadratic forms*, J. London Math. Soc. (2) **50** (1994), no. 2, 222–230.
6. J.S. Hsia, *Spinor norms of local integral rotations I*, Pacific J. Math. **57** (1975), no. 1, 199–206.
7. B. Kane and Z.W. Sun, *On almost universal mixed sums of squares and triangular numbers*, Trans. Amer. Math. Soc. **362** (2010), 6425–6455.
8. O.T. O'Meara, *Introduction to quadratic forms*, Springer-Verlag, New York, 1963.
9. R. Schulze-Pillot, *Exceptional integers for genera of integral ternary positive definite quadratic forms*, Duke Math. J. **102** (2000), no. 2, 351–357.

Weighted Generating Functions for Type II Lattices and Codes

Noam D. Elkies and Scott Duke Kominers

Abstract We give a new structural development of harmonic polynomials on Hamming space, and harmonic weight enumerators of binary linear codes, that parallels one approach to harmonic polynomials on Euclidean space and weighted theta functions of Euclidean lattices. Namely, we use the finite-dimensional representation theory of $\mathfrak{sl}_2$ to derive a decomposition theorem for the spaces of discrete homogeneous polynomials in terms of the spaces of discrete harmonic polynomials, and prove a generalized MacWilliams identity for harmonic weight enumerators. We then present several applications of harmonic weight enumerators, corresponding to some uses of weighted theta functions: an equivalent characterization of t-designs, the Assmus–Mattson Theorem in the case of extremal Type II codes, and configuration results for extremal Type II codes of lengths 8, 24, 32, 48, 56, 72, and 96.

Key words and Phrases Harmonic polynomial • Weight enumerator • Binary code • Extremal code • Theta function • Lattice • Design • Configuration result

Mathematics Subject Classification (2010): Primary: 94B05; Secondary: 05B05, 11H71, 33C50, 33C55

N.D. Elkies (✉)
Department of Mathematics, Harvard University, One Oxford Street, Cambridge,
MA 02138, USA
e-mail: elkies@math.harvard.edu

S.D. Kominers
Society of Fellows, Harvard University, and Becker Friedman Institute for Research in
Economics, University of Chicago, 78 Mount Auburn street, Cambridge, MA 02138, USA
e-mail: kominers@fas.harvard.edu; skominers@uchicago.edu

K. Alladi et al. (eds.), *Quadratic and Higher Degree Forms*, Developments
in Mathematics 31, DOI 10.1007/978-1-4614-7488-3_4,

1 Introduction

A well-known and fruitful analogy relates lattices L in Euclidean space $\mathbb{R}^n$ with linear codes C in binary Hamming space $\mathbb{F}_2^n$. (See for instance [Ebe02], [Elk00], and [CS99, 3.2].) Under this analogy the theta function

$$\Theta_L(q) = \sum_{v \in L} q^{\langle v,v\rangle/2} = \sum_{k \geq 0} \left(\sum_{\langle v,v\rangle = 2k} 1 \right) q^k, \tag{1.1}$$

a generating function that counts vectors $v \in L$ in spheres $\{v : \langle v,v\rangle = 2k\}$ about the origin, corresponds to the weight enumerator

$$W_C(x,y) = \sum_{c \in C} x^{n-\mathrm{wt}(c)} y^{\mathrm{wt}(c)} = \sum_{w=0}^{n} \left(\sum_{\mathrm{wt}(c)=w} 1 \right) x^{n-w} y^w, \tag{1.2}$$

a generating function counting words $c \in C$ in Hamming spheres $\{c : \mathrm{wt}(c) = w\}$ about the origin. This paper concerns a generalization of Θ_L and W_C that can be used not only to count lattice or code elements in each sphere, by summing the constant function 1 as in (1.1) and (1.2), but also to measure their distribution, by summing a suitable nonconstant function. In the lattice case, we sum a harmonic polynomial P on $\mathbb{R}^n$, yielding the weighted theta function

$$\Theta_{L,P}(q) = \sum_{v \in L} P(v) q^{\langle v,v\rangle/2} = \sum_{k \geq 0} \left(\sum_{\langle v,v\rangle = 2k} P(v) \right) q^k. \tag{1.3}$$

In the code case, we sum a discrete harmonic polynomial Q on $\mathbb{F}_2^n$, yielding the harmonic weight enumerator[1]

$$W_{C,Q}(x,y) = \sum_{c \in C} Q(c) x^{n-\mathrm{wt}(c)} y^{\mathrm{wt}(c)} = \sum_{w=0}^{n} \left(\sum_{\mathrm{wt}(c)=w} Q(c) \right) x^{n-w} y^w. \tag{1.4}$$

Weighted theta functions have been used extensively to study the configurations of lattice vectors. But discrete harmonic polynomials and harmonic weight enumerators are relatively unknown and rarely used. Moreover, the known construction of discrete harmonic polynomials Q, and the known proofs of the basic properties of

[1] While the analogy between $\Theta_{L,P}$ and $W_{C,Q}$ suggests calling $W_{C,Q}$ a "weighted weight enumerator", the comical juxtaposition of the two senses of "weight" dissuades us from using that phrase. Since Bachoc [Bac99] had already introduced the term "harmonic weight enumerator" that avoids this juxtaposition, we happily follow her usage.

these Q and of the associated $W_{C,Q}$ (see [Del78, Bac99]), involve manipulations of intricate combinatorial sums that are considerably harder than, and look nothing like, the developments of their Euclidean counterparts.

Here we give a structural development of discrete harmonic polynomials and harmonic weight enumerators that parallels the more familiar theory of harmonic polynomials on $\mathbb{R}^n$ and weighted theta functions. In each case we use an action of the Lie algebra $\mathfrak{sl}_2$ on spaces of functions on $\mathbb{R}^n$ (for lattices) or on $\mathbb{F}_2^n$ (for codes). While the two cases are not completely parallel, the remaining distinctions are inherent in the structure of Euclidean and Hamming space; for instance, homogeneous polynomials on $\mathbb{F}_2^n$ cannot be defined by $Q(cv) = c^d Q(v)$, and since Hamming space is finite all the representations of $\mathfrak{sl}_2$ that figure in the discrete theory are finite-dimensional. Once we have established the new approach to discrete harmonic polynomials and harmonic weight enumerators, we use it to give cleaner derivations of the Assmus–Mattson theorem [AM69] and of the Koch condition [Koc87] on the tetrad systems of Type II codes of length 24.[2] Finally we outline some further applications to the configurations of minimal-weight words in extremal Type II codes that parallel recent configuration results for extremal Type II lattices.

The rest of the paper is organized as follows. We first outline the $\mathfrak{sl}_2$ approach to harmonic polynomials on $\mathbb{R}^n$ and to the construction and basic properties of weighted theta functions, and the connection with design properties of Type II lattices. In the next section we review the MacWilliams identity for weight enumerators and Gleason's theorem for the weight enumerator of a Type II code. In the following three sections we use the $\mathfrak{sl}_2$ theory to develop the theory of discrete harmonic polynomials Q, prove the MacWilliams identity for harmonic weight enumerators $W_{C,Q}$, and study the important special case where Q is a "zonal harmonic polynomial" (discrete harmonic polynomial invariant under a subgroup $S_w \times S_{n-w}$ of the group S_n of coordinate permutations of $\mathbb{F}_2^n$). The next two sections relate these polynomials with t-designs and recover the Assmus–Mattson theorem for extremal Type II codes and the Koch condition for Type II codes of length 24. Finally we use these techniques to show for several values of n that any extremal Type II code of length n is generated by its words of minimal weight, again in analogy with known results for extremal Type II lattices. In an Appendix, we give a direct proof of Gleason's theorems for self-dual codes of Types I and II; certain polynomials needed to describe harmonic weight enumerators occur naturally in the course of this proof.

While the present paper considers codes only over $\mathbb{F}_2$, discrete harmonic polynomials and harmonic weight enumerators generalize to linear codes over arbitrary finite fields $\mathbb{F}_q$ (see [Bac01]). Our development of these notions extends to that setting too, using representations of $\mathfrak{sl}_q$ instead of $\mathfrak{sl}_2$. This change introduces enough new complications that we defer the analysis to future work.

[2] The second of these requires only the $W_{C,Q}$ for Q of degree 1, which coincide with Ott's "local weight enumerators" [Ott99].

2 Weighted Theta Functions and Configurations of Type II Lattices

2.1 Lattice-Theoretic Preliminaries

By a *lattice* in Euclidean space $\mathbb{R}^n$ we mean a discrete subgroup $L \subset \mathbb{R}^n$ of rank n; equivalently, L is the $\mathbb{Z}$-span of the columns of an invertible $n \times n$ real matrix, say M (which does not depend uniquely on L: two such matrices M, M' yield the same L if and only if $M^{-1}M'$ has integer entries and determinant ± 1). The *covolume* $\mathrm{Vol}(\mathbb{R}^n/L)$ of the lattice is then $|\det M|$. The *dual lattice* is defined by

$$L^* = \{v^* \in \mathbb{R}^n : \forall v \in L, \langle v, v^* \rangle \in \mathbb{Z}\}. \tag{2.1}$$

If L is the $\mathbb{Z}$-span of the columns of the invertible matrix M then L^* is the $\mathbb{Z}$-span of the columns of the transpose of M^{-1}; in particular $\mathrm{Vol}(\mathbb{R}^n/L^*) = \mathrm{Vol}(\mathbb{R}^n/L)^{-1}$.

If $L = L^*$ then L is *self-dual*. Then $\langle v, v' \rangle \in \mathbb{Z}$ for all $v, v' \in L$, and the norm map $L \to \mathbb{Z}$, $v \mapsto \langle v, v \rangle$ reduces modulo 2 to a group homomorphism $L \to \mathbb{Z}/2\mathbb{Z}$. The lattice is said to be *even* or *of Type II* if this homomorphism is trivial, that is, if $\langle v, v \rangle \in 2\mathbb{Z}$ for all $v \in L$; otherwise L is said to be *odd* or *of Type I*.

Examples. For each $n \geq 1$ the lattice $\mathbb{Z}^n \subset \mathbb{R}^n$ is of Type I. It is the unique Type I lattice in $\mathbb{R}^n$ for $n = 1$, and unique up to isomorphism for $n \leq 8$, but not unique for any $n \geq 9$; there are finitely many isomorphism classes of Type I lattices in $\mathbb{R}^n$, but the number of classes grows rapidly as $n \to \infty$ (see for instance [CS99, p. 403]).

If $\mathbb{R}^n$ contains a Type II lattice then $n \equiv 0 \bmod 8$ (see [Ser73, Chap. V]). Such a lattice may be constructed as follows. For any n let D_n be the sublattice of $\mathbb{Z}^n$ consisting of all $(x_1, \ldots, x_n)$ such that $\sum_{j=1}^n x_j \equiv 0 \bmod 2$, and let D_n^+ be the union of D_n and the translate of D_n by the half-integer vector $(1/2, 1/2, \ldots, 1/2)$. Then D_n^+ is:

- a lattice if and only if $2 \mid n$,
- self-dual if and only if $4 \mid n$, and
- of Type II if and only if $8 \mid n$.

For $n = 8$, this lattice D_8^+ coincides with the Gosset root lattice E_8, which is known to be the unique Type II lattice in $\mathbb{R}^8$ up to isomorphism; we give one proof of its uniqueness at the end of this section.[3] There are two Type II lattices for $n = 16$ (namely $E_8 \oplus E_8$ and D_{16}^+), and 24 for $n = 24$ (the Niemeier lattices [Nie73]); for large $n \equiv 0 \bmod 8$ the number is again always finite but grows rapidly as $n \to \infty$ (see for instance [CS99, p. 50]).

[3] Serre [Ser73, Chap. VII] uses the notation E_n for our D_n^+ for all $n \equiv 0 \bmod 8$, but this notation has not been widely adopted. For $n \equiv 4 \bmod 8$ the Type I lattice D_n^+ is isomorphic with $\mathbb{Z}^n$ if and only if $n = 4$.

2.2 Poisson Summation

The *Poisson summation formula* is a remarkable identity relating the sum of a function f over a lattice and the sum of the Fourier transform of f over the dual lattice. We review this formula in the case of Schwartz functions, which is all that we need. Recall that a *Schwartz function* is a C^∞ function $f : \mathbb{R}^n \to \mathbb{C}$ such that f and all its partial derivatives decay as $o(\langle x, x\rangle^k)$ for all k as $\langle x, x\rangle \to \infty$. We define the *Fourier transform* $\hat{f} : \mathbb{R}^n \to \mathbb{C}$ by

$$\hat{f}(y) = \int_{x\in\mathbb{R}^n} f(x)\, e^{2\pi i\langle x,y\rangle}\, d\mu(x); \tag{2.2}$$

$\hat{f}$ is a Schwartz function if f is.

Theorem 2.1 (Poisson Summation Formula). *Let L be any lattice in $\mathbb{R}^n$. Then*

$$\sum_{x\in L} f(x) = \frac{1}{\mathrm{Vol}(\mathbb{R}^n/L)} \sum_{y\in L^*} \hat{f}(y) \tag{2.3}$$

for all Schwartz functions $f : \mathbb{R}^n \to \mathbb{C}$.

Proof. Define $F : \mathbb{R}^n \to \mathbb{C}$ by

$$F(z) = \sum_{x\in L} f(x+z).$$

Because f is Schwartz, the sum converges absolutely to a C^∞ function, whose value at $z = 0$ is the left-hand side of (2.3). Since $F(z) = F(x+z)$ for all $z \in \mathbb{R}^n$ and $x \in L$, the function descends to a C^∞ function on $\mathbb{R}^n/L$, and thus has a Fourier expansion

$$F(z) = \sum_{y\in L^*} \hat{F}(-y)\, e^{2\pi i\langle y,z\rangle}, \tag{2.4}$$

where

$$\hat{F}(y) = \frac{1}{\mathrm{Vol}(\mathbb{R}^n/L)} \int_{z\in\mathbb{R}^n/L} F(z)\, e^{2\pi i\langle z,y\rangle}\, d\mu(z).$$

Note that the vectors $y \in L^*$ are exactly those for which $e^{2\pi i\langle x,y\rangle}$ is well-defined on $\mathbb{R}^n/L$. Now choose a fundamental domain R for $\mathbb{R}^n/L$; for instance, let $v_1, \ldots, v_n$ be generators of L and set $R = \{a_1v_1 + \cdots + a_nv_n : 0 \le a_i < 1\}$. Then we have

$$\begin{aligned}\mathrm{Vol}(\mathbb{R}^n/L)\hat{F}(y) &= \int_{z\in R} F(z)\, e^{2\pi i\langle y,z\rangle}\, d\mu(z)\\ &= \int_{z\in R} \sum_{x\in L} f(x+z)\, e^{2\pi i\langle y,z\rangle}\, d\mu(z)\end{aligned}$$

$$= \sum_{x\in L} \int_{z\in R+x} f(z)\, e^{2\pi i\langle y,z\rangle}\, d\mu(z)$$

$$= \int_{z\in\mathbb{R}^n} f(z)\, e^{2\pi i\langle y,z\rangle}\, d\mu(z) \;=\; \hat{f}(y),$$

where we use in the next-to-last step the fact that $\mathbb{R}^n$ is the disjoint union of the translates $R+x$ of R by lattice vectors. Thus (2.4) becomes

$$F(z) = \frac{1}{\mathrm{Vol}(\mathbb{R}^n/L)} \sum_{y\in L^*} \hat{f}(-y)\, e^{2\pi i\langle y,z\rangle}. \tag{2.5}$$

Taking $z=0$ we obtain (2.3). □

2.3 Theta Functions

Suppose now that q is a real number with $0 < q < 1$. We may then take $f(x) = q^{\langle x,x\rangle/2}$ and recognize the left-hand side of (2.3) as the sum $\Theta_L(q)$ of (1.1). The Poisson summation formula then yields the following functional equation for theta functions.

Proposition 2.2. *Let L be any lattice in* $\mathbb{R}^n$*. Then*

$$\Theta_{L^*}(e^{-2\pi t}) = \mathrm{Vol}(\mathbb{R}^n/L)t^{-n/2}\Theta_L(e^{-2\pi/t}) \tag{2.6}$$

for all $t>0$.

Proof. Let $f(x) = \exp(-\pi\langle x,x\rangle/t)$ in (2.3). We claim that

$$\hat{f}(y) = t^{n/2}\exp(-\pi\langle y,y\rangle t). \tag{2.7}$$

Indeed, choosing any orthonormal coordinates $(x_1,\ldots,x_n)$ for $\mathbb{R}^n$, we see that the integral (2.2) defining $\hat{f}(y)$ factors as

$$\prod_{j=1}^{n} \int_{-\infty}^{\infty} e^{-\pi x_j^2/t} e^{2\pi i x_j y_j}\, dx_j,$$

which reduces our claim to the case $n=1$, which is the familiar definite integral

$$\int_{-\infty}^{\infty} e^{-\pi x^2/t}\, e^{2\pi i x y}\, dx = t^{1/2} e^{-\pi t y^2}$$

(see for instance [Rud76, Example 9.43, pp. 237–238] or [Kör90, Lemma 50.2(i), pp. 246–247]). Using these f and $\hat{f}$ in the Poisson summation formula (2.3) we deduce the functional equation (2.6). □

Now suppose L is a Type II lattice. Then $L^* = L$, so the functional equation relates Θ_L to itself, and $\mathrm{Vol}(\mathbb{R}^n/L) = 1$. Moreover, each of the exponents $\langle v, v\rangle/2$ occurring in the formula (1.1) is an integer, so $\Theta_L(q)$ is a power series in q and extends to a function on the unit disc $|q| < 1$ in $\mathbb{C}$. Thus by analytic continuation the identity $\Theta_L(e^{-2\pi t}) = t^{-n/2}\Theta_L(e^{-2\pi/t})$ holds for all $t \in \mathbb{C}$ of positive real part. But $\Theta_L(e^{-2\pi t})$, being a power series in $e^{-2\pi t}$, is also invariant under $t \mapsto t + i$. This leads us to define the function

$$\theta_L(\tau) := \Theta_L(e^{2\pi i\tau}) = \sum_{v\in L} e^{\pi\langle v,v\rangle i\tau} \tag{2.8}$$

for τ in the Poincaré upper half-plane

$$\mathcal{H} := \{\tau \in \mathbb{C} : \mathrm{Im}(\tau) > 0\}.$$

Then $\theta_L(\tau) = \theta_L(\tau + 1)$, and the Poisson identity gives $\theta_L(\tau) = \tau^{-n/2}\theta_L(-1/\tau)$: the expected factor of $i^{n/2}$ disappears because $n \equiv 0 \bmod 8$ for all Type II lattices. It follows that

$$\theta_L(\tau) = (c\tau + d)^{-n/2}\,\theta_L\left(\frac{a\tau + b}{c\tau + d}\right) \tag{2.9}$$

for all $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in the subgroup of $\mathrm{SL}_2(\mathbb{R})$ generated by $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$. This subgroup is the full modular group $\mathrm{SL}_2(\mathbb{Z})$ of integer matrices of determinant 1. (See [Ser73, Chap. VII] for this and the remaining results noted in this paragraph.) The identity (2.9) for all such $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, together with the fact that $\theta_L(\tau)$ remains bounded as $\mathrm{Im}(\tau) \to \infty$ (because then $q \to 0$), then shows that θ_L is a modular form of weight $n/2$ for $\mathrm{SL}_2(\mathbb{Z})$. Since $n/2 \equiv 0 \bmod 4$, this means that θ_L is a polynomial in the normalized Eisenstein series

$$\mathcal{E}_4 = \theta_{E_8}(\tau) = 1 + 240\sum_{n=1}^{\infty}\frac{n^3q^n}{1-q^n} = 1 + 240q + 2160q^2 + 6720q^3 + \cdots$$

of weight 4 (where again $q = e^{2\pi i\tau}$) and the cusp form[4]

$$\Delta(\tau) = q\prod_{n=1}^{\infty}(1-q^n)^{24} = q - 24q^2 + 252q^3 - 1472q^4 \cdots$$

of weight 12. Moreover the coefficient of $\mathcal{E}_4^{n/8}$ in this polynomial equals 1 because that coefficient is the constant coefficient in the q-expansion, which is the number of lattice vectors of norm zero.

[4] That is, a modular form that vanishes at all the cusps; for $\mathrm{SL}_2(\mathbb{Z})$ there is only one cusp, at $\mathrm{Im}(\tau) \to \infty$, so a modular form in $\mathrm{SL}_2(\mathbb{Z})$ is a cusp form if and only if its expansion as a power series in q has constant coefficient zero. Note that the notation of [Ser73] diverges from the usual practice that we follow: our $\mathcal{E}_4$, $\mathcal{E}_6$, and Δ are what Serre calls E_2, E_3 and $(2\pi)^{-12}\Delta$. (We use "$\mathcal{E}$" rather than "E" to avoid confusion with the E_8 lattice.)

It follows for example that if $n = 8$ or $n = 16$ then $\theta_L = \mathcal{E}_4^{n/8}$, while if $n = 8m$ with $m = 3$, 4, or 5 and L contains no vectors v with $\langle v, v\rangle = 2$ then $\theta_L = \mathcal{E}_4^m - 240m\theta_{E_8}^{m-3}\Delta$ (so for example the q^2 coefficient is $720m(211 - 40m) > 0$ and L has that many vectors v with $\langle v, v\rangle = 4$). It is known that such L are unique up to isomorphism for $n = 8$ and $n = 24$ (the E_8 and Leech lattices respectively), but there are two choices for $n = 16$, and literally millions for $n = 32$ (see [Kin03]) and many more for $n = 40$, all with the same number of vectors of norm $2k$ for each k.

More generally, given any $n = 8m$ the theta series of any Type II lattice L can be written uniquely as $\mathcal{E}_4^m + \sum_{k=1}^{\lfloor m/3\rfloor} a_k\Delta^k\mathcal{E}_4^{m-3k}$ for some a_k. If L contains no vectors v with $0 < \langle v, v\rangle \leq 2\lfloor m/3\rfloor$ then the a_k are uniquely determined by induction, and thus all such lattices have the same theta series. Such lattices L are known as *extremal lattices*, and their common theta function θ_L is the *extremal theta function*. Siegel [Sie69] proved that the $q^{\lfloor m/3\rfloor+1}$ coefficient of θ_L is positive, from which Mallows, Odlyzko, and Sloane [MOS75] deduced that a Type II lattice $L \subset \mathbb{R}^n$ has minimal norm at most $2(\lfloor m/3\rfloor + 1)$, with equality if and only if L is extremal.

2.4 The Spaces of Harmonic Polynomials

Let $\mathscr{P}$ be the $\mathbb{C}$-vector space of polynomials on $\mathbb{R}^n$, and $\mathscr{P}_d$ $(d = 0, 1, 2, \ldots)$ its subspace of homogeneous polynomials of degree d, so that $\mathscr{P} = \bigoplus_{d=0}^\infty \mathscr{P}_d$. The *Laplacian* is the differential operator defined by[5]

$$\Delta = \sum_{j=1}^n \frac{\partial^2}{\partial x_j^2} : \mathrm{C}^\infty(\mathbb{R}^n) \to \mathrm{C}^\infty(\mathbb{R}^n), \quad \mathscr{P} \to \mathscr{P}, \quad \mathscr{P}_d \to \mathscr{P}_{d-2}. \tag{2.10}$$

Here $x_1, \ldots, x_n$ are any orthonormal coordinates on $\mathbb{R}^n$, and $\mathscr{P}_d$ is taken to be $\{0\}$ for $d < 0$. The space of *harmonic polynomials* of degree d is then

$$\mathscr{P}_d^0 := \ker(\Delta : \mathscr{P}_d \to \mathscr{P}_{d-2}); \tag{2.11}$$

this is the degree-d homogeneous part of

$$\mathscr{P}^0 := \bigoplus_{d=0}^\infty \mathscr{P}_d^0 = \ker(\Delta : \mathscr{P} \to \mathscr{P}). \tag{2.12}$$

[5] The use of Δ for this operator and Δ for the modular form $\eta^{24} = q\prod_{n=1}^\infty(1 - q^n)^{24}$ may be unfortunate, but should not cause confusion, despite the similarity between the two symbols, because they never appear together outside this footnote. The alternative notation L for the Laplacian would be much worse, as we regularly use L for a lattice.

For example, $\mathscr{P}_0^0$ and $\mathscr{P}_1^0$ are the spaces of constant and linear functions respectively, of dimensions 1 and n; and a quadratic polynomial $P = \sum_{1 \le j \le k \le n} a_{jk} x_j x_k$ is harmonic if and only if $\sum_{j=1}^n a_{jj} = 0$, because ΔP is the constant polynomial $2 \sum_{j=1}^n a_{jj}$.

It is well known, and we shall soon demonstrate, that $\Delta : \mathscr{P}_d \to \mathscr{P}_{d-2}$ is surjective, whence

$$\dim(\mathscr{P}_d^0) = \dim(\mathscr{P}_d) - \dim(\mathscr{P}_{d-2}) = \binom{n+d-1}{d} - \binom{n+d-3}{d}. \quad (2.13)$$

We shall use two further operators on $\mathrm{C}^\infty(\mathbb{R}^n)$ and on its subspace $\mathscr{P}$. The first is

$$\mathsf{E} := x \cdot \nabla = \sum_{j=1}^n x_j \frac{\partial}{\partial x_j}. \quad (2.14)$$

Euler proved that if $P \in \mathrm{C}^\infty(\mathbb{R}^n)$ is homogeneous of degree d then $\mathsf{E}P = d \cdot P$; in particular $\mathscr{P}_d$ is the d-eigenspace of $\mathsf{E}|_{\mathscr{P}}$. The second operator is multiplication by the norm:

$$\mathsf{F} := \langle x, x \rangle = \sum_{j=1}^n x_j^2 : \; P \mapsto \langle x, x \rangle P. \quad (2.15)$$

Clearly F injects each $\mathscr{P}_d$ into $\mathscr{P}_{d+2}$. Thus $\mathscr{P}_d^0 = \ker(\mathsf{F}\Delta : \mathscr{P}_d \to \mathscr{P}_d)$; that is, $\mathscr{P}_d^0$ is the zero eigenspace of the operator $\mathsf{F}\Delta$ on $\mathscr{P}_d$. We next show that the other eigenspaces are $\mathsf{F}^k \mathscr{P}_{d-2k}^0$ for $k = 1, 2, \ldots, \lfloor d/2 \rfloor$, and that $\mathscr{P}_d$ is the direct sum of these eigenspaces, from which the surjectivity of $\Delta : \mathscr{P}_d \to \mathscr{P}_{d-2}$ soon follows.

We begin by finding the commutators of $\Delta, \mathsf{E}, \mathsf{F}$. Recall that the *commutator* of any two operators A, B on some vector space is

$$[A, B] = AB - BA = -[B, A].$$

For example, $[x_j, x_k] = [\partial/\partial x_j, \partial/\partial x_k] = 0$ for all j, k, while $[\partial/\partial x_j, x_k] = \delta_{jk}$ (Kronecker delta). Applying these formulas repeatedly, we obtain the commutation relations

$$[\Delta, \mathsf{F}] = 4\mathsf{E} + 2n, \quad [\mathsf{E}, \Delta] = -2\Delta, \quad [\mathsf{E}, \mathsf{F}] = 2\mathsf{F}. \quad (2.16)$$

This suggests the commutation relations

$$[\mathsf{X}, \mathsf{Y}] = \mathsf{H}, \quad [\mathsf{H}, \mathsf{X}] = 2\mathsf{X}, \quad [\mathsf{H}, \mathsf{Y}] = -2\mathsf{Y} \quad (2.17)$$

satisfied by the standard basis

$$(\mathsf{X}, \mathsf{H}, \mathsf{Y}) = \left(\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right) \right) \quad (2.18)$$

of $\mathfrak{sl}_2$. Indeed (2.16) is tantamount to an isomorphism of Lie algebras from $\mathfrak{sl}_2$ to the span of $\{\Delta, \mathsf{E} + \frac{n}{2}, \mathsf{F}\}$ that takes $(\mathsf{X}, \mathsf{H}, \mathsf{Y})$ to $(\frac{1}{2\varpi}\Delta, -(\mathsf{E} + \frac{n}{2}), -\frac{\varpi}{2}\mathsf{F})$ for some nonzero ϖ (all choices of ϖ are equivalent via conjugation by diagonal matrices;

later the choice $\varpi = 2\pi$ will be most natural for us). Some steps in the following analysis are familiar from the representation theory of $\mathfrak{sl}_2$, though here only infinite-dimensional representations arise.

Now suppose $P \in \mathscr{P}_d$ is in the λ-eigenspace of $\mathsf{F}\Delta$ for some λ. Then $\langle x,x\rangle P = \mathsf{F}P$ is in the $(\lambda+4d+2n)$-eigenspace of $\mathsf{F}\Delta$ acting on $\mathscr{P}_{d+2}$, because

$$\mathsf{F}\Delta\mathsf{F}P = \mathsf{F}(\mathsf{F}\Delta + [\Delta,\mathsf{F}])P = \mathsf{F}(\mathsf{F}\Delta + 4\mathsf{E} + 2n)P = \mathsf{F}(\lambda + 4d + 2n)P.$$

By induction on $k = 0,1,2,\ldots$ it follows that $\mathsf{F}^k P$ is an eigenvector of $\mathsf{F}\Delta|_{\mathscr{P}_{d+2k}}$ with eigenvalue

$$\lambda + \sum_{j=0}^{k-1}\bigl(4(d+2j)+2n\bigr) = \lambda + k\bigl(4(d+k-1)+2n\bigr).$$

Replacing d by $d-2k$ and taking $\lambda = 0$, we see that if $P \in \mathscr{P}^0_{d-2k}$ then $\mathsf{F}^k P$ is an eigenvector of $\mathsf{F}\Delta|_{\mathscr{P}_d}$ with eigenvalue

$$\lambda_d(k) := k\bigl(4(d-k-1)+2n\bigr).$$

We next prove that this accounts for all the eigenspaces of $\mathsf{F}\Delta|_{\mathscr{P}_d}$.

Lemma 2.3. *Fix $d \geq 0$. For integers k, k' such that $0 \leq k < k' \leq d/2$ we have $\lambda_d(k) < \lambda_d(k')$.*

Proof. By induction it is enough to check this for $k' = k+1$. We compute

$$\lambda_d(k+1) - \lambda_d(k) = 2n + 4(d-2k') \geq 2n > 0,$$

as claimed. □

Corollary 2.4. *The sum over $k = 0,1,\ldots,\lfloor d/2\rfloor$ of the subspaces $\mathsf{F}^k\mathscr{P}^0_{d-2k}$ of $\mathscr{P}_d$ is direct.*

Proof. By Lemma 2.3, the $\lambda_d(k)$ are strictly increasing, and thus distinct. Our claim follows because $\mathsf{F}^k\mathscr{P}^0_{d-2k}$ is a subspace of the $\lambda_d(k)$ eigenspace of $\mathsf{F}\Delta$. □

Proposition 2.5. *For $k = 0,1,\ldots,\lfloor d/2\rfloor$, let $\mathscr{P}^k_d = \mathsf{F}^k\mathscr{P}^0_{d-2k}$. Then:*

(1) *The map $\Delta : \mathscr{P}_d \to \mathscr{P}_{d-2}$ is surjective.*
(2) *$\mathscr{P}_d = \bigoplus_{k=0}^{\lfloor d/2\rfloor}\mathscr{P}^k_d = \mathscr{P}^0_d \oplus \mathsf{F}\mathscr{P}_{d-2}$, and $\mathscr{P} = \bigoplus_{k=0}^{\infty}\mathsf{F}^k\mathscr{P}^0$.*
(3) *$\mathscr{P}^k_d$ is the entire $\lambda_d(k)$ eigenspace of $\mathsf{F}\Delta|_{\mathscr{P}_d}$, and $\mathsf{F}\Delta|_{\mathscr{P}_d}$ has no eigenvalues other than the $\lambda_d(k)$ for $k = 0,1,\ldots,\lfloor d/2\rfloor$.*
(4) *$\dim(\mathscr{P}^0_d) = \dim(\mathscr{P}_d) - \dim(\mathscr{P}_{d-2})$ as claimed in (2.13).*

Proof. The sum $\bigoplus_{k=0}^{\lfloor d/2\rfloor}\mathscr{P}^k_d$ is direct by Corollary 2.4. We prove that it equals $\mathscr{P}_d$ by comparing dimensions. Since F is injective we have $\dim(\mathscr{P}^k_d) = \dim(\mathscr{P}^0_{d-2k})$; moreover

$$\dim(\mathscr{P}^0_{d-2k}) \geq \dim(\mathscr{P}_{d-2k}) - \dim(\mathscr{P}_{d-2k-2}),$$

with equality if and only if $\Delta : \mathscr{P}_{d-2k} \to \mathscr{P}_{d-2k-2}$ is surjectve, because $\mathscr{P}^0_{d-2k}$ is the kernel of $\Delta : \mathscr{P}_{d-2k} \to \mathscr{P}_{d-2k-2}$. Hence $\dim\bigl(\bigoplus_{k=0}^{\lfloor d/2 \rfloor} \mathscr{P}^k_d\bigr)$ is

$$\sum_{k=0}^{\lfloor d/2 \rfloor} \dim(\mathscr{P}^k_d) = \sum_{k=0}^{\lfloor d/2 \rfloor} \dim(\mathscr{P}^0_{d-2k}) \geq \sum_{k=0}^{\lfloor d/2 \rfloor} \bigl(\dim(\mathscr{P}_{d-2k}) - \dim(\mathscr{P}_{d-2k-2})\bigr), \tag{2.19}$$

and the last sum telescopes to $\dim(\mathscr{P}_d)$. Thus equality holds termwise in the last step of (2.19) and $\dim\bigl(\bigoplus_{k=0}^{\lfloor d/2 \rfloor} \mathscr{P}^k_d\bigr) = \dim(\mathscr{P}_d)$. The first of these proves part (1) (using the $k = 0$ term). The second yields

$$\mathscr{P}_d = \bigoplus_{k=0}^{\lfloor d/2 \rfloor} \mathscr{P}^k_d, \tag{2.20}$$

as claimed in part (2); taking the direct sum over d yields $\mathscr{P} = \bigoplus_{k=0}^{\infty} \mathsf{F}^k \mathscr{P}^0$, also claimed in part (2). To complete the proof of part (2) we compare the decompositions (2.20) of $\mathscr{P}_d$ and $\mathscr{P}_{d-2}$ and note that $\mathscr{P}^k_d = \mathsf{F}\mathscr{P}^{k-1}_{d-2}$ for each $k > 0$. Part (3) follows because the decomposition (2.20) diagonalizes $\mathsf{F}\Delta|_{\mathscr{P}_d}$. Finally part (4) is again the equality of the $k = 0$ terms in (2.19). □

Remarks. Part (2) of Proposition 2.5 says in effect that $\mathscr{P} = \bigoplus_{d=0}^{\infty} \bigl(\mathscr{P}^0_d \otimes U_{\frac{n}{2}+d}\bigr)$, where for any real $m > 0$ we write U_m for the infinite-dimensional irreducible representation of $\mathfrak{sl}_2$ with basis $\{\mathsf{Y}^k v\}_{k=0}^{\infty}$ where $\mathsf{X}v = 0$ and $\mathsf{H}v = -mv$. These U_m come from representations in the "discrete series" of unitary representations of the Lie group $\mathrm{SL}_2(\mathbb{R})$ when n is even (see [Lan75, Chap. IX]); when n is odd, they come from discrete-series representations of the "metaplectic" double cover of $\mathrm{SL}_2(\mathbb{R})$ that do not descend to $\mathrm{SL}_2(\mathbb{R})$.

It also follows from part (2) that $\mathscr{P}^0_d \cap \mathsf{F}\mathscr{P}_{d-2} = \{0\}$, and thus that $\mathscr{P}^0$ contains no nonzero multiple of $\langle x, x \rangle$. Proving this was set as problem B-5 on the 2005 Putnam exam, which was the hardest of the 12 problems that year, solved by only 5 of the top 200 scorers (see [KAL06, pp. 736 and 741]). The solution printed in [KAL06, p. 742] uses some of the ingredients used here to prove Proposition 2.5.

2.5 Weighted Theta Functions

The functional equation (2.6) for theta functions of lattices extends to theta functions weighted by a harmonic polynomial.

Theorem 2.6. *Let L be any lattice in $\mathbb{R}^n$, and $P : \mathbb{R}^n \to \mathbb{C}$ any harmonic polynomial of degree d. Then*

$$\Theta_{L^*,P}(e^{-2\pi t}) = i^d \operatorname{Vol}(\mathbb{R}^n/L) t^{-(n/2)-d} \Theta_{L,P}(e^{-2\pi/t}) \tag{2.21}$$

for all $t > 0$.

By the Poisson summation formula, this will follow from the following generalization of (2.7):

Theorem 2.7. *Suppose that $t > 0$ and $P : \mathbb{R}^n \to \mathbb{C}$ is a harmonic polynomial on $\mathbb{R}^n$ of degree d. Define a function $f : \mathbb{R}^n \to \mathbb{R}$ by*

$$f(x) = P(x)\, e^{-\pi\langle x,x\rangle t}. \tag{2.22}$$

Then the Fourier transform of f is

$$\hat{f}(y) = i^d t^{-(\frac{n}{2}+d)} P(y)\, e^{-\pi\langle y,y\rangle/t}. \tag{2.23}$$

Proof. For $t \in \mathbb{C}$ define an operator

$$\mathsf{G}_t : \mathrm{C}^\infty(\mathbb{R}^n) \to \mathrm{C}^\infty(\mathbb{R}^n), \quad g \mapsto e^{-\pi t\langle x,x\rangle} g \tag{2.24}$$

that multiplies every C^∞ function by the Gaussian $e^{-\pi t\langle x,x\rangle}$; these operators constitute a one-parameter group: $\mathsf{G}_t\mathsf{G}_{t'} = \mathsf{G}_{t+t'}$ for all t, t'. We are then interested in $f = \mathsf{G}_t P$ for $P \in \mathscr{P}$ in the intersection of the kernel of Δ with an eigenspace of E. If $P \in \mathscr{P}_d$ then

$$d \cdot f = \mathsf{G}_t(d \cdot P) = \mathsf{G}_t\mathsf{E}P = (\mathsf{G}_t\mathsf{E}\mathsf{G}_{-t})\mathsf{G}_t P = (\mathsf{G}_t\mathsf{E}\mathsf{G}_{-t}) f,$$

so f is in the d-eigenspace of $\mathsf{G}_t\mathsf{E}\mathsf{G}_{-t}$; likewise $f \in \ker \mathsf{G}_t\Delta\mathsf{G}_{-t}$. Since our one-parameter group $\{\mathsf{G}_t\}$ has infinitesimal generator $-\pi\mathsf{F}$, we expect that conjugation by G_t will take Δ, E to some linear combination of $\Delta, \mathsf{E}, \mathsf{F}$. Indeed we find the following relations.[6]

Lemma 2.8 (Conjugation of $\Delta, \mathsf{E}, \mathsf{F}$ by G_t). *The operators G_t commute with F, and we have*

$$\mathsf{G}_t\mathsf{E}\mathsf{G}_{-t} = \mathsf{E} + 2\pi t\mathsf{F}, \quad \mathsf{G}_t\Delta\mathsf{G}_{-t} = \Delta + \pi t(4\mathsf{E} + 2n) + (2\pi t)^2\mathsf{F}. \tag{2.25}$$

Proof. As with the commutation relations (2.16), this comes down to an exercise in differential calculus. Here we start from the fact that G_t commutes with each x_j while $\mathsf{G}_t(\partial/\partial x_j)\mathsf{G}_{-t} = 2\pi t x_j + (\partial/\partial x_j)$, whence the first formula in (2.25) quickly follows, while $\mathsf{G}_t\mathsf{F} = \mathsf{F}\mathsf{G}_t$ is immediate. A somewhat longer computation establishes the second formula. □

Corollary 2.9. *The operators $\Delta, \mathsf{E}, \mathsf{F}$ act on $\mathsf{G}_t\mathscr{P}$, and the subspace $\mathsf{G}_t\mathscr{P}_d^0$ is the intersection of* $\ker(\Delta + \pi t(4\mathsf{E} + 2n) + (2\pi t)^2\mathsf{F})$ *with the d-eigenspace of* $\mathsf{E} + 2\pi t\mathsf{F}$ *in $\mathsf{G}_t\mathscr{P}$.*

[6] This is where it becomes natural to use $\varpi = 2\pi$ when choosing the images of the generators (2.18) of $\mathfrak{sl}_2$: conjugation by G_t then takes $(\mathsf{X}, \mathsf{H}, \mathsf{Y})$ to $(\mathsf{X} - t\mathsf{H} - t^2\mathsf{Y}, \mathsf{H} + 2t\mathsf{Y}, \mathsf{Y})$; other choices would produce more complicated coefficients.

We next relate the Fourier transform of a Schwartz function f with the Fourier transforms of its images under $\Delta, \mathsf{E}, \mathsf{F}$.

Lemma 2.10 (Conjugation of $\Delta, \mathsf{E}, \mathsf{F}$ by the Fourier Transform). *Let $f : \mathbb{R}^n \to \mathbb{C}$ be any Schwartz function. Then:*

(1) *For each $j = 1, \ldots, n$, the Fourier transform of $x_j f$ is $(2\pi i)^{-1}\partial \hat{f}/\partial y_j$, and the Fourier transform of $\partial f/\partial x_j$ is $-2\pi i y_j \hat{f}$.*

(2) *The Fourier transforms of Δf, $(2\mathsf{E}+n)f$, and $\mathsf{F}f$ are respectively $-(2\pi)^2\mathsf{F}\hat{f}$, $-(2\mathsf{E}+n)\hat{f}$, and $-(2\pi)^{-2}\Delta\hat{f}$.*

Proof. Again this is a calculus exercise, here with definite integrals. The formula for the Fourier transform of $\partial f/\partial x_j$ is obtained by integrating by parts with respect to x_j. The Fourier transform of $x_j f$ can be obtained from this using Fourier inversion, or directly by differentiation with respect to y_j of the integral (2.2) that defines $\hat{f}(y)$. We then obtain part (2) by iterating the formulas in part (1) to find the Fourier transform of $\partial^2 f/\partial x_j^2$, $x_j \partial f/\partial x_j$, or $x_j^2 f$, and summing over j. The case of $\mathsf{E}f$ can be explained by writing the operator $2\mathsf{E}+n$ as

$$\sum_{j=1}^{n}\bigl(x_j(\partial/\partial x_j) + (\partial/\partial x_j)\circ x_j\bigr).$$

□

We use this to show that if $f \in \mathsf{G}_t\mathscr{P}$ then $\hat{f} \in \mathsf{G}_{1/t}\mathscr{P}$, that is, that $\hat{f}$ is *some* polynomial multiplied by $e^{-\pi\langle y,y\rangle/t}$. More precisely:

Proposition 2.11. *Let $t \in \mathbb{C}$ with $\mathrm{Re}(t) > 0$. If $f = \mathsf{G}_t P$ for some $P \in \mathscr{P}_d$ then $\hat{f} = \mathsf{G}_{1/t}\widehat{P}$ for some $\widehat{P} = \sum_{d'=0}^{d}\widehat{P}_{d'}$ with each $\widehat{P}_{d'} \in \mathscr{P}_{d'}$ and $\widehat{P}_d = i^d t^{-(\frac{n}{2}+d)}P$. As before $t^{-(\frac{n}{2}+d)}$ denotes the $-(n+2d)$ power of the principal square root of t.*

Proof. We use induction on d. The base case $d = 0$ is the fact that the Fourier transform of $e^{-\pi t\langle x,x\rangle}$ is $t^{-n/2}e^{-\pi\langle y,y\rangle/t}$, which we showed already. Suppose we have established the claim for $P \in \mathscr{P}_d$. By linearity and the fact that $\mathscr{P}_{d+1}$ is spanned by its subspaces $x_j\mathscr{P}_d$, it is enough to prove the proposition with P replaced by x_jP. By the first part of Lemma 2.10, the Fourier transform of $\mathsf{G}_t x_j P = x_j \mathsf{G}_t P$ is

$$\frac{1}{2\pi i}\frac{\partial}{\partial y_j}\bigl(\mathsf{G}_{1/t}\widehat{P}\bigr) = \frac{1}{2\pi i}\mathsf{G}_{1/t}\Bigl(\frac{\partial\widehat{P}}{\partial y_j} - \frac{2\pi}{t}y_j\widehat{P}\Bigr). \tag{2.26}$$

By the inductive hypothesis $\widehat{P}$ has degree d and leading part $\widehat{P}_d = i^d t^{-(\frac{n}{2}+d)}P$. Therefore the right-hand side of (2.26) has degree $d+1$ and leading part

$$\frac{-2\pi t^{-1}}{2\pi i}\widehat{P}_d = \frac{i}{t}\widehat{P}_d = i^{d+1}t^{-(\frac{n}{2}+d+1)}y_jP.$$

This completes the inductive step and the proof. □

To finish the proof of Theorem 2.7, suppose that $P \in \mathscr{P}_d^0$ and that $f(x) = P(x)\, e^{-\pi\langle x,x\rangle t} = \mathsf{G}_t P$. By Corollary 2.9,

$$(\Delta + \pi t(4\mathsf{E} + 2n) + (2\pi t)^2\mathsf{F})f = 0, \qquad (\mathsf{E} + 2\pi t\mathsf{F})f = d \cdot f.$$

Taking the Fourier transform and applying the second part of Lemma 2.10, we deduce

$$(-(2\pi)^2\mathsf{F} - \pi t(4\mathsf{E} + 2n) - t^2\Delta)\hat{f} = 0, \quad -\Big(\mathsf{E} + n + \frac{t}{2\pi}\Delta\Big)\hat{f} = d \cdot \hat{f}.$$

Eliminating $\Delta\hat{f}$, we find $d \cdot \hat{f} = (\mathsf{E} + \frac{2\pi}{t}\mathsf{F})\hat{f}$; that is, $\hat{f}$ is in the d-eigenspace of $\mathsf{E} + 2\pi t^{-1}\mathsf{F}$. By Proposition 2.11, we know that $\hat{f}{=}\mathsf{G}_{1/t}\widehat{P}$ for some $\widehat{P} \in \mathscr{P}$. By Lemma 2.8, then, $\widehat{P}$ is in the d-eigenspace of E; that is, $\widehat{P} \in \mathscr{P}_d$. By Proposition 2.11, we conclude that $\widehat{P} = i^d t^{-(\frac{n}{2}+d)} P$. □

We have now proven the functional equation (2.21) for weighted theta functions $\Theta_{L,P}$ (Theorem 2.6). This identity is trivial when $d = \deg(P)$ is odd, because then $\Theta_{L,P}$ is identically zero (by cancellation of the v and $-v$ terms), but it gives new information when d is even and positive.

Again we consider the special case of a Type II lattice. Generalizing (2.8), we define

$$\theta_{L,P}(\tau) := \Theta_{L,P}(e^{2\pi i\tau}) = \sum_{v\in L} P(v) e^{\pi\langle v,v\rangle i\tau} \tag{2.27}$$

for $\tau \in \mathcal{H}$. Then $\theta_{L,P}(\tau) = \theta_{L,P}(\tau + 1)$, and Theorem 2.6 gives $\theta_{L,P}(\tau) = t^{-(\frac{n}{2}+d)}\theta_L(-1/\tau)$, with the factor i^d absorbed by the change of variable $\tau = it$ because d is even. It follows as before that

$$\theta_{L,P}(\tau) = (c\tau + d)^{-(\frac{n}{2}+d)}\, \theta_{L,P}\left(\frac{a\tau + b}{c\tau + d}\right) \tag{2.28}$$

for all $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$, so $\theta_{L,P}$ is a modular form of weight $\frac{n}{2} + d$ for $\mathrm{SL}_2(\mathbb{Z})$. Hence $\theta_{L,P}$ is a polynomial in $\mathcal{E}_4$ and the weight-6 Eisenstein series

$$\mathcal{E}_6 = 1 - 504\sum_{n=1}^{\infty} \frac{n^5 q^n}{1 - q^n} = 1 - 504q - 16632q^2 - 122976q^3 - \cdots .$$

Moreover $\theta_{L,P}$ is a cusp form once $d > 0$: the constant coefficient is $P(0)$, which vanishes for nonconstant homogeneous P. Hence once d is positive the polynomial giving $\theta_{L,P}$ in terms of $\mathcal{E}_4$ and $\mathcal{E}_6$ is divisible by $\Delta = 12^{-3}(\mathcal{E}_4^3 - \mathcal{E}_6^2)$. (See again [Ser73, Chap. VII].)

In particular $\theta_{L,P} = 0$ when $\frac{n}{2} + d \in \{2, 4, 6, 8, 10, 14\}$ because in those weights the only cusp form is the zero form.[7] Likewise we have the following observation.

Lemma 2.12. *Suppose $n = 8m$ and $L \subset \mathbb{R}^n$ is an extremal lattice. Then $\theta_{L,P} = 0$ for every nonconstant harmonic polynomial P on $\mathbb{R}^n$ whose degree d satisfies $4(m - 3\lfloor m/3 \rfloor) + d \in \{2, 4, 6, 8, 10, 14\}$. If $L \subset \mathbb{R}^n$ is a Type II lattice of minimal norm $n/12$ then $\theta_{L,P} = 0$ for every harmonic polynomial P on $\mathbb{R}^n$ of degree 2.*

Proof. We saw already that $\theta_{L,P}$ is a cusp form. If L is extremal, the q^k coefficient of $\theta_{L,P}$ vanishes for each $k \le \lfloor m/3 \rfloor$. Hence $\Delta^{-\lfloor m/3 \rfloor}\theta_{L,P}$ is a cusp form of weight $4(m - 3\lfloor m/3 \rfloor) + d$, and thus vanishes when $4(m - 3\lfloor m/3 \rfloor) + d \in \{2, 4, 6, 8, 10, 14\}$. Likewise if L has minimal norm $n/12$ and P is a quadratic harmonic polynomial then $\Delta^{1-(n/24)}\theta_{L,P}$ is a cusp form of weight 14, so again $\theta_{L,P} = 0$. □

If L is extremal then Lemma 2.12 applies to 6, 4, or 2 values of d for $n \equiv 0$, 8, or 16 mod 24 respectively. We exploit these vanishing results in the next section.

2.6 Spherical t-Designs, the Venkov Condition on Niemeier Lattices, and Extremal Type II Lattices

For real $\nu > 0$ let $A_\nu : \mathrm{C}^\infty(\mathbb{R}^n) \to \mathbb{C}$ be the functional that takes any function to its average on the sphere $\Sigma_\nu = \{x \in \mathbb{R}^n : \langle x, x \rangle = \nu\}$ with respect to the probability measure on Σ_ν invariant under the orthogonal group. For any positive integer t, a (possibly empty[8]) finite set $D \subset \mathbb{R}^n$ of nonzero vectors of equal norm ν is said to be a *(spherical) t-design* if and only if

$$\sum_{v \in D} P(v) = |D| \cdot A_\nu(P) \tag{2.29}$$

for all $P \in \mathscr{P}$ with $\deg P \le t$.[9] By linearity it is enough to check this condition for $P \in \mathscr{P}_d$ for each $d \le t$, and we may assume $d > 0$ because in the case $d = 0$ of

[7] In this setting $\frac{n}{2} + d$ cannot be as small as 2 because $n \ge 8$, but the possibility of weight 2 arises in the proof of Lemma 2.12.

[8] With this definition $\emptyset$ is a t-design for all t. For most applications only nonempty designs are of interest; for instance it is only when D is nonempty that we can divide both sides of (2.29) by $|D|$ to get the equivalent condition that the average of any polynomial of degree at most t over Σ_ν can be computed by averaging it over $|D|$. But we allow empty designs here, and also later in the coding-theoretic setting, because this simplifies the statements of the results relating lattices with spherical designs.

[9] See [Del78] for explanation of the term "r-design" for this property. For $D \ne \emptyset$, the t-design property is one way to make precise the idea that D is "well distributed" in Σ_ν, and better distributed as t grows. One application, and the original one according to [CS99, pp. 89–90], is numerical integration on Σ_ν, using the right-hand side of (2.29) as an approximation to the

a constant polynomial the condition (2.29) is satisfied automatically. We next prove that it is enough to check (2.29) for *harmonic* polynomials of positive degree. We begin by showing that all such polynomials are in $\ker(A_\nu)$.

Lemma 2.13. *If P is a nonconstant harmonic polynomial then $A_\nu(P) = 0$.*

Proof. Choose any $s > 0$. Since P is homogeneous, $A_\nu(P)$ is a positive multiple of the integral of $\mathsf{G}_s P$ over all of $\mathbb{R}^n$. But this integral is the value of the Fourier transform of $\mathsf{G}_s P$ at the origin. By Theorem 2.7 this value is some multiple of $P(0)$. Since $d > 0$ we have $P(0) = 0$, so $A_\nu(P) = 0$ as claimed. □

Thus our design criterion can be stated as follows.

Lemma 2.14. *A finite subset $D \subset \Sigma_\nu$ is a t-design if and only if $\sum_{v \in D} P(v) = 0$ for all nonconstant harmonic polynomials P of degree at most t.*

Proof. The "only if" direction is immediate from Lemma 2.13. We prove the "if" implication. By the second part of Proposition 2.5 any polynomial of degree $d \leq t$ can be written as $\sum_{k=0}^{\lfloor d/2 \rfloor} \mathsf{F}^k P_k$ with each P_k harmonic of degree $d - 2k$. It is thus enough to check (2.29) for each $\mathsf{F}^k P_k$. But by hypothesis, (2.29) holds for each P_k (including $P_{d/2}$ if d is even, because then P_k is constant). Since the restriction of each $\mathsf{F}^k P_k$ to Σ_ν is $\nu^k P_k$, it follows that (2.29) holds for $\mathsf{F}^k P_k$ as well, and we are done. □

Combining this with Lemma 2.12 yields the following theorem of Venkov [Ven01], which asserts that in an extremal or nearly extremal Type II lattice the vectors of each nonzero norm form a spherical design.

Theorem 2.15. *Let $L \subset \mathbb{R}^n$ be a Type II lattice with minimal norm $2k$. Assume $r := 24k - n$ is nonnegative. Set $t = 3$ if $r = 0$ and $t = (r/2) - 1$ if $r > 0$. Then $L \cap \Sigma_\nu$ is a t-design for every $\nu > 0$.*

Proof. Because $L \cap \Sigma_\nu$ is centrally symmetric, we need only check the criterion of Lemma 2.14 for P of even degree. For such P, Lemma 2.12 applies, so $\theta_{L,P} = 0$. The criterion $\sum_{v \in D} P(v)$ then holds because $\sum_{v \in D} P(v)$ is a coefficient of $\theta_{L,P}$. □

Remarks. In general $L \cap \Sigma_\nu$ need not be a $(t+1)$-design: there will be lattice norms ν and harmonic polynomials P of degree $t+1$ whose sum over $L \cap \Sigma_\nu$ is nonzero. However, when $r > 0$ it will be true that the sum over $L \cap \Sigma_\nu$ of any harmonic polynomial of degree $t+3$ vanishes, because there are no nonzero cusp forms of weight 14. Thus each $L \cap \Sigma_\nu$ is what Venkov [Ven01] called a "*$t\frac{1}{2}$-design*": a finite subset $D \subset \Sigma_\nu$ such that $\sum_{v \in D} P(v) = 0$ for all $P \in \mathscr{P}_d^0$ with either $d \leq t$ or $d = t + 3$.

left-hand side even when P is not polynomial but smooth enough to be well approximated by polynomials.

The fact that in each case $L \cap \Sigma_\nu$ is a 2-design already lets us deduce that if $L \cap \Sigma_\nu$ is nonempty then it spans $\mathbb{R}^n$ as a vector space. Indeed if $L \cap \Sigma_\nu$ does not span $\mathbb{R}^n$ then it is contained in a hyperplane $\{x \in \mathbb{R}^n : \langle x, \dot{x} \rangle = 0\}$ for some nonzero $\dot{x} \in \mathbb{R}^n$; then we can take $P(x) = \langle x, \dot{x} \rangle^2$ in (2.29) and observe that each of the terms $P(v)$ in the left-hand side vanishes, while the factor $A_\nu(P)$ of the right-hand side is strictly positive, so the remaining factor $|D|$ must vanish, making $L \cap \Sigma_\nu = \emptyset$ as claimed.

More precise results can often be obtained when ν equals or slightly exceeds the minimal norm, because then any two vectors in $L \cap \Sigma_\nu$ must have integer inner product, and only a few integers can arise, making the condition that $L \cap \Sigma_\nu$ be a t-design or a $t\frac{1}{2}$-design particularly stringent. We give three examples: Venkov's simplification of Niemeier's classification of Type II lattices in $\mathbb{R}^{24}$; configuration results for extremal Type II lattices in several dimensions, including multiples of 24 up to 96, showing that such lattices are generated by their minimal vectors; and a novel proof of the uniqueness of the E_8 lattice.

Niemeier Lattices

Suppose L is a Type II lattice in $\mathbb{R}^{24}$. Then the hypothesis of Theorem 2.15 is satisfied with $r = 0$ or $r = 24$. In either case we find in particular that $L \cap \Sigma_2$ is a 2-design. But the vectors of norm 2 in any even lattice constitute a root system. Venkov [Ven80] used the requirement that this root system be a 2-design to show *a priori* that it must be among the 24 root systems that arise for the Niemeier lattices, and thus to considerably streamline the classification of Type II lattices in $\mathbb{R}^{24}$.

Configuration Results for Extremal Type II Lattices

While a nonempty shell $L \cap \Sigma_\nu$ in an extremal lattice L must generate $\mathbb{R}^n$ as a vector space, it need not generate L over $\mathbb{Z}$: already $(L, \nu) = (D_{16}^+, 2)$ is a counterexample, since the minimal nonzero vectors of D_{16}^+ generate only the index-2 sublattice D_{16}. Still, for some n it can be proved that every extremal lattice is generated by its vectors of minimal norm $2k$. Let L_0 be the sublattice of L generated by the minimal vectors, and assume $[L : L_0] > 1$. Then there are nonlattice vectors $\dot{v} \in L_0^*$, and $\langle v, \dot{v} \rangle \in \mathbb{Z}$ for all $v \in L \cap \Sigma_{2k}$. If $\dot{v}$ has minimal norm in its coset $\bmod L$ then $|\langle v, \dot{v} \rangle| \leq k$ for all such v. This together with the t-design or $t\frac{1}{2}$-design condition on $L \cap \Sigma_{2k}$ yields a contradiction for several values of n, proving that $L_0 = L$ for each of those n. (See [Ven84], [Oze86a], [Oze86b], [Kom09a], and [Elk12].)

The Uniqueness of E_8

Finally, let $n = 8$ and let $L \subset \mathbb{R}^8$ be any Type II lattice. Then

$$\theta_L = \mathcal{E}_4 = 1 + 240q + 2160q^2 + \cdots ,$$

and L is automatically extremal, so in particular $L \cap \Sigma_2$ is a 7-design of size 240. We shall use these facts to prove that $L \cong E_8$. There are 2160 vectors of norm 4 in L; choose one, and call it $\dot{x}$. Let D be the 7-design $L \cap \Sigma_2$. For $j \in \mathbb{Z}$ let N_j be the number of vectors $x \in D$ such that $\langle x, \dot{x} \rangle = j$. If $N_j \neq 0$, then $|j| \leq \sqrt{8}$ (by Cauchy–Schwarz) and $j \in \mathbb{Z}$ (because $\langle v, v' \rangle \in \mathbb{Z}$ for all $v, v' \in L$); hence $j \in \{-2, -1, 0, 1, 2\}$. Therefore

$$\sum_{j=-2}^{2} N_j = |D| = 240. \tag{2.30}$$

Since D is centrally symmetric, $N_{-j} = N_j$ for each j. Finally, since D is a 7-design, (2.29) holds with $P(x) = \langle x, \dot{x} \rangle^d$ for each positive integer $d \leq 7$. This is automatic for d odd, but for $d = 2, 4, 6$ we get linear equations in N_0, N_1, N_2, and already the $d = 2$ and $d = 4$ equations together with (2.30) let us solve for the N_j. We find

$$(N_{-2}, N_{-1}, N_0, N_1, N_2) = (14, 64, 84, 64, 14). \tag{2.31}$$

(See the Remarks at the end of this section for the evaluation of the functional A_ν on even powers of $\langle x, \dot{x} \rangle$.) In particular there are 14 vectors in D, call them v_i for $1 \leq i \leq 14$, whose inner product with $\dot{x}$ is 2.

For each i we obtain a lattice vector $x_i = 2v_i - \dot{x}$ that is orthogonal to $\dot{x}$ and satisfies $\langle x_i, x_i \rangle = 4$ and $x_i \equiv \dot{x} \bmod 2L$. For any i and i' we have

$$\begin{aligned} \langle x_i, x_{i'} \rangle = \langle 2v_i - \dot{x}, 2v_{i'} - \dot{x} \rangle &= 4\langle v_i, v_{i'} \rangle - 2\langle v_i, \dot{x} \rangle - 2\langle \dot{x}, v_{i'} \rangle + \langle \dot{x}, \dot{x} \rangle \\ &= 4\langle v_i, v_{i'} \rangle - 4 - 4 + 4 \\ &= 4\langle v_i, v_{i'} \rangle - 4 \\ &\equiv 0 \bmod 4. \end{aligned}$$

Thus the vectors x_i for $1 \leq i \leq 14$, together with $\dot{x}$ and $-\dot{x}$, are 16 vectors of norm 4, any two of which are equal, opposite, or orthogonal. Hence the x_i together with $\pm\dot{x}$ are the minimal vectors of an isometric copy of $2\mathbb{Z}^8$ in L. Moreover L also contains $v_i = (\dot{x} + x_i)/2$, and thus contains the $\mathbb{Z}$-span of $\dot{x}$ and the v_i, which is isometric with D_8. But L is self-dual, so $D_8^* \subset L \subset D_8$. Of the three lattices satisfying this condition, one is $\mathbb{Z}^8$, which is of Type I, and the other two are isomorphic with E_8. Therefore $L \cong E_8$, as claimed.

Remarks. A related proof, parallel to the beginning of Conway's proof [Con69] of the uniqueness of the Leech lattice, starts from the observation that each of the 2^8 cosets of $2L$ in L intersects $\{v \in L : \langle v, v \rangle \leq 4\}$ in either $\{0\}$, a pair of minimal vectors, or at most 8 orthogonal pairs of vectors of norm 4. This accounts for at least $1 + 240/2 + 2160/16 = 256 = 2^8$ cosets. Hence equality holds throughout, and any of the nonzero cosets that does not meet Σ_2 gives us a copy of D_8 in L. This approach uses only the modularity of θ_L, not of the more general $\theta_{L,P}$, though it applies in fewer cases. Either technique also yields the number of automorphisms of E_8: there are 2160 choices of $\dot{x}$, and $2^7 7!$ automorphisms of D_8 that fix $\dot{x}$, half of which send E_8 to itself, so $|\mathrm{Aut}(E_8)| = 2160 \cdot 2^6 7! = 696729600$.

For even $d \geq 0$, and a given vector $\dot{x}$ of norm $\dot{\nu} > 0$, the average over Σ_ν of $\langle x, \dot{x} \rangle^d$ is computed as a quotient of Beta integrals. We find that if $P(x) = \langle x, \dot{x} \rangle^d$ then

$$A_\nu(P) = (\nu\dot{\nu})^{d/2} \frac{\int_0^1 u^d (1-u^2)^{(n-3)/2}\, du}{\int_0^1 (1-u^2)^{(n-3)/2}\, du} = (\nu\dot{\nu})^{d/2} \frac{\mathrm{B}\big((d+1)/2, (n-1)/2\big)}{\mathrm{B}\big(1/2, (n-1)/2\big)}, \tag{2.32}$$

where u is the normalized projection $(\nu\dot{\nu})^{-1/2} |\langle x, \dot{x} \rangle|$. Thus

$$A_\nu(P) = (\nu\dot{\nu})^{d/2} \, \frac{1}{n} \, \frac{3}{n+2} \, \frac{5}{n+4} \cdots \frac{d-1}{n+d-2}\,. \tag{2.33}$$

In our case $\nu\dot{\nu} = 2 \cdot 4 = 8$, so $A_\nu(P) = 1, 12/5, 8$ for $d = 2, 4, 6$.

Alternatively we could apply Lemma 2.14 to the *zonal spherical harmonics*, which are harmonic polynomials that depend only on $\langle x, \dot{x} \rangle$. For each degree d there is a one-dimensional space of zonal spherical harmonics, proportional to a Gegenbauer orthogonal polynomial $C_m^{((n-2)/2)}(u)$ with $u = (\nu\dot{\nu})^{-1/2} \langle x, \dot{x} \rangle$. This is equivalent to using (2.32) and (2.33) for t-designs, but for a $t\frac{1}{2}$-design we need the zonal spherical harmonics to exploit the vanishing of $\sum_{v \in D} P(v)$ for $P \in \mathscr{P}^0_{t+3}$. This, too, has an analog in the setting of discrete harmonic polynomials, as in the proof of Theorem 9.2 at the end of this paper.

3 Weight Enumerators of Binary Linear Codes

3.1 Coding-Theoretic Preliminaries

By a *(binary linear) code of length* n we mean a vector subspace of the $\mathbb{F}_2$-vector space $\mathbb{F}_2^n$. In this context, vectors of length n over $\mathbb{F}_2$ are often called (binary) "words" of length n. The *(Hamming) weight* of a word $w \in \mathbb{F}_2^n$, denoted by $\mathrm{wt}(w)$, is the number of nonzero coordinates of w, and the *(Hamming) distance* between two words $w, w' \in \mathbb{F}_2^n$ is $\mathrm{wt}(w' - w)$. We denote by $(\cdot, \cdot)$ the usual bilinear pairing on $\mathbb{F}_2^n$, defined by $(v, w) = \sum_{j=1}^n v_j w_j$. For a linear code $C \subseteq \mathbb{F}_2^n$, the *dual code* is the annihilator $C^\perp$ of C with respect to this pairing; thus $\dim(C) + \dim(C^\perp) = n$ and $C^{\perp\perp} = C$ for every linear code $C \subseteq \mathbb{F}_2^n$.

If $C = C^\perp$ then C is *self-dual*. Then $(c, c') = 0$ for all $c, c' \in C$, and in particular $\mathrm{wt}(c)$ is even for all $c \in C$ because $0 = (c, c)$ is the reduction of $\mathrm{wt}(c)$ mod 2. The map $\mathrm{wt} : C \to \mathbb{Z}$ then reduces mod 4 to a group homomorphism $C \to 2\mathbb{Z}/4\mathbb{Z}$. The code C is said to be *doubly even* or *of Type II* if this homomorphism is trivial, that is, if $(c, c) \in 4\mathbb{Z}$ for all $c \in C$; otherwise C is said to be *singly even* or *of Type*

I. This notation reflects the analogy between binary linear codes and lattices. It also respects the following construction ("Construction A" of [LS71]; see also [CS99, pp. 182–183]) that associates a lattice $L_C \subset \mathbb{R}^n$ to any linear code $C \subseteq \mathbb{F}_2^n$:

$$L_C := \{2^{-1/2}v : v \in \mathbb{Z}^n,\ v \bmod 2 \in C\}. \tag{3.1}$$

Indeed $L_C^* = L_{C^\perp}$, so L_C is self-dual if and only if C is, in which case L_C is of Type I or Type II according as C is of Type I or Type II, respectively.

Examples. If $C = C^\perp$ then $\dim(C) = n/2$, so n is even. For each positive even integer n there is a Type I code of length n consisting of all c such that $c_{2j-1} = c_{2j}$ for each $j \leq n/2$. This is the unique Type I code for $n = 2$, and is unique up to isomorphism (i.e., up to coordinate permutation) for $n \leq 8$, but not unique for any $n \geq 10$; and as with lattices the number of isomorphism classes grows rapidly with n.

If $\mathbb{F}_2^n$ contains a Type II code then $n \equiv 0 \bmod 8$. (This follows via Construction A from the corresponding theorem for lattices, but can also be proven directly.[10]) An example is the *extended Hamming code* in $\mathbb{F}_2^8$: if we identify $\mathbb{F}_2^8$ with the space of $\mathbb{F}_2$-valued functions on $\mathbb{F}_2^3$, the extended Hamming code can be constructed as the subspace of affine-linear functions on $\mathbb{F}_2^3$. The extended Hamming code is the unique Type II code of length 8; there are two such codes of length 16, nine of length 24, and a rapidly growing number as $n \to \infty$ through multiples of 8.

3.2 Discrete Poisson Summation

We define the *discrete Fourier transform* (or *Hadamard transform*) $\hat{f}$ of a function $f : \mathbb{F}_2^n \to \mathbb{C}$ as the function on $\mathbb{F}_2^n$ given by

$$\hat{f}(u) = \sum_{v \in \mathbb{F}_2^n} (-1)^{(u,v)} f(v). \tag{3.2}$$

[10] Suppose C is a self-dual code of length n. Then C contains the all-1s vector $\mathbf{1}$, because $(v, v) = (v, \mathbf{1})$ for all $v \in \mathbb{F}_2^n$, so $C \subseteq C^\perp$ implies $\mathbf{1} \in C^\perp$. Thus C descends to a vector space of dimension $(n/2) - 1$ in $\mathrm{V} := \{0, \mathbf{1}\}^\perp/\{0, \mathbf{1}\}$. Since $2 \mid n$, the perfect pairing $(\cdot, \cdot)$ descends to a perfect pairing on V, so a self-dual code is tantamount to a maximal isotropic subspace of V relative to this pairing. If $4 \mid n$ then the map $\{0, \mathbf{1}\}^\perp \to \mathbb{F}_2, v \mapsto (\mathrm{wt}(c)/2) \bmod 2$ descends to a quadratic form $\mathrm{Q} : \mathrm{V} \to \mathbb{F}_2$ consistent with that pairing. A Type II code is then a self-dual code C that is totally isotropic relative to Q. Such C exists if and only if (V, Q) has Arf invariant zero. But the Arf invariant is 0 or 1 according as $\{v \in \mathrm{V} \colon \mathrm{Q}(v) = 0\}$ has size $2^{n-3} + 2^{(n/2)-2}$ or $2^{n-3} - 2^{(n/2)-2}$. But this count is $(1/2)\sum_{j=0}^{n/4} \binom{n}{4j} = (1/8)\sum_{\mu^4=1}(1+\mu)^n = 2^{n-3} + (1/4)\,\mathrm{Re}(1+i)^n$, so the result follows from the observation that $(1+i)^4 = -4$.

We review the *discrete Poisson summation formula*, a discrete analog of the Poisson summation formula for lattices (Theorem 2.1). Like its lattice analog, the discrete Poisson summation formula relates the sum of a function to the sum of the function's discrete Fourier transform. Here, however, instead of considering the sums of the function and its Fourier transform over a lattice $L \subset \mathbb{R}^n$ and its dual L^*, we consider the sums of the function and its discrete Fourier transform over a linear code $C \subset \mathbb{F}_2^n$ and over $C^\perp$, the dual code of C.

Theorem 3.1 (Discrete Poisson Summation Formula). *Let $C \subset \mathbb{F}_2^n$ be a binary linear code of length n, and let f be a function from $\mathbb{F}_2^n$ to $\mathbb{C}$. Then*

$$\sum_{c\in C} f(c) = \frac{1}{|C^\perp|} \sum_{c'\in C^\perp} \hat{f}(c'). \tag{3.3}$$

We briefly recount the standard proof of Theorem 3.1, which is the one presented in [MS83, p. 127].

Proof of Theorem 3.1. By expanding the sum in the right-hand side of (3.3) and rearranging the order of summation, we obtain

$$\sum_{c'\in C^\perp} \hat{f}(c') = \sum_{c'\in C^\perp} \sum_{v\in\mathbb{F}_2^n} (-1)^{(c',v)} f(v) = \sum_{v\in\mathbb{F}_2^n} f(v) \sum_{c'\in C^\perp} (-1)^{(c',v)}. \tag{3.4}$$

Now, whenever $v \in C \subset \mathbb{F}_2^n$ and $c' \in C^\perp$, we have $(c', v) = 0$ by the definition of $C^\perp$. It follows that the inner sum in (3.4) equals $|C^\perp|$ whenever $v \in C$. Furthermore, when $v \notin C$, the inner sum of (3.4) vanishes.[11] The result then follows immediately. □

3.3 The MacWilliams Identity and Gleason's Theorem

In this section, we recall two classical results from coding theory which are closely related to the theory of lattices. The first of these results, the MacWilliams identity (Theorem 3.2, below), expresses the weight enumerator of $C^\perp$ in terms of the weight enumerator of C. The second result (Theorem 3.3, below) is a famous theorem originally due to Gleason [Gle71], which shows that the weight enumerators of Type II codes can be expressed in terms of two particular weight enumerators.

Theorem 3.2 (MacWilliams Identity ([Mac63]; [CS99, p. 78]; [Ebe02, p. 74]; [MS83, p. 126])). *For any binary linear code C of length n, we have*

[11] In this case, (c', v) takes the values 0 and 1 equally often (see [MS83, p. 127]). (This statement is just an instance of the well-known fact that the sum of a nontrivial character on a finite commutative group vanishes.) We could also adapt the technique we used in proving Theorem 2.1, obtaining discrete Poisson summation via the discrete Fourier expansion of the function $z \mapsto \sum_{c\in C} f(c+z)$.

$$W_C(x,y) = \frac{1}{|C^\perp|} W_{C^\perp}(x+y, x-y). \tag{3.5}$$

Proof. Define a function $f : \mathbb{F}_2^n \to \mathbb{C}$ by $f(v) = x^{n-\mathrm{wt}(v)} y^{\mathrm{wt}(v)}$. Then

$$\hat{f}(u) = (x+y)^{n-\mathrm{wt}(u)} (x-y)^{\mathrm{wt}(u)}.$$

Theorem 3.2 therefore follows directly from the discrete Poisson summation formula (Theorem 3.1). □

Theorem 3.3 (Gleason's Theorem ([Gle71]; [Slo77]; [CS99, p. 192]; [Ebe02, p. 75])). *For any Type II code C, the weight enumerator $W_C(x,y)$ is a polynomial in*

$$\varphi_8 := x^8 + 14x^4y^4 + y^8 \quad \textit{and} \quad \xi_{24} := x^4y^4(x^4 - y^4)^4. \tag{3.6}$$

Proof. Since C is of Type II, the exponent of y in each monomial $x^{n-\mathrm{wt}(v)} y^{\mathrm{wt}(v)}$ is a multiple of 4. Thus each monomial is invariant under the substitution of iy for y, whence the sum $W_C(x,y)$ of these monomials also satisfies the identity $W_C(x,y) = W_C(x,iy)$. Since $C = C^\perp$, we also have an identity

$$\begin{aligned} W_C(x,y) &= \frac{1}{|C|} W_C(x+y, x-y) \\ &= 2^{-n/2} W_C(x+y, x-y) \\ &= W_C\bigl(2^{-1/2}(x+y), 2^{-1/2}(x-y)\bigr) : \end{aligned} \tag{3.7}$$

the first step uses Theorem 3.2; for the second, we deduce $|C| = 2^{n/2}$ from $2^n = |C| \cdot |C^\perp| = |C|^2$; and for the last step, we use the fact that W_C is a homogeneous polynomial of degree n. Therefore this homogeneous polynomial is invariant under the group, call it G_{II}, generated by linear substitutions with matrices $\left(\begin{smallmatrix} 1 & 0 \\ 0 & i \end{smallmatrix}\right)$ and $2^{-1/2} \left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$.

It turns out that G_{II} is a complex reflection group, and thus has a polynomial ring of invariants. Namely, G_{II} is #9 in the Shephard-Todd list [ST54], and its invariant degrees are 8 and 24, with φ_8, ξ_{24} as a convenient choice of generators. This result completes the proof of Gleason's theorem for Type II codes. □

In Appendix A we give a direct proof that $\mathbb{C}[x,y]^{G_{\mathrm{II}}} = \mathbb{C}[\varphi_8, \xi_{24}]$. The literature contains several other approaches to the determination of this invariant ring, including Ebeling's proof in [Ebe02] using the theory of modular forms(!). See [CS99, p. 192]. The method we use reaches G_{II} via a suitable tower of reflection groups starting from $\{1\}$, each normal in the next; along the way we also obtain Gleason's theorem for Type I codes, and encounter a polynomial ψ_{12}, invariant under an index-2 subgroup of G_{II}, that will figure in our subsequent development.

4 The Spaces of Discrete Harmonic Polynomials

In this section, we present some useful results in the theory of *discrete harmonic polynomials*. These polynomials were originally introduced by Delsarte [Del78], who gave a combinatorial development. Here, we give a new approach to these polynomials using the finite-dimensional representation theory of $\mathfrak{sl}_2$.

4.1 Basic Definitions and Notation

A function g on $\mathbb{F}_2$ may be interpreted as a 2×1 matrix $g = \left(\begin{smallmatrix} g_0 \\ g_1 \end{smallmatrix}\right)$, where g_v is the value assumed on input $v \in \mathbb{F}_2$. It is easily computed that the discrete Fourier transform $\hat{g}$ of g is the function

$$\hat{g} = \left(\begin{smallmatrix} g_0+g_1 \\ g_0-g_1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)\left(\begin{smallmatrix} g_0 \\ g_1 \end{smallmatrix}\right);$$

the discrete Fourier transform is therefore encoded by the matrix $\mathsf{T} := \left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$. There is a natural action of $\mathfrak{sl}_2$ on these functions g, defined by multiplication from the left by matrices in $\mathfrak{sl}_2$. Thus, we may interpret the space of functions on $\mathbb{F}_2$ as a representation of $\mathfrak{sl}_2$ isomorphic with the 2-dimensional defining representation V_1 of $\mathfrak{sl}_2$.

More generally, the operator

$$\widetilde{\mathsf{T}} := \mathsf{T}^{\otimes n} \tag{4.1}$$

on $V_1^{\otimes n}$ gives the discrete Fourier transform on $v \in \mathbb{F}_2^n$. For a pure tensor

$$g = \left(\begin{smallmatrix} g_{10} \\ g_{11} \end{smallmatrix}\right) \otimes \cdots \otimes \left(\begin{smallmatrix} g_{n0} \\ g_{n1} \end{smallmatrix}\right) \in V_1^{\otimes n},$$

which takes the value $g_{1v_1} \cdots g_{nv_n}$ on $v \in \mathbb{F}_2^n$, we have

$$\widetilde{\mathsf{T}}g = \widehat{\left(\begin{smallmatrix} g_{10} \\ g_{11} \end{smallmatrix}\right)} \otimes \cdots \otimes \widehat{\left(\begin{smallmatrix} g_{n0} \\ g_{n1} \end{smallmatrix}\right)} = \left(\begin{smallmatrix} g_{10}+g_{11} \\ g_{10}-g_{11} \end{smallmatrix}\right) \otimes \cdots \otimes \left(\begin{smallmatrix} g_{n0}+g_{n1} \\ g_{n0}-g_{n1} \end{smallmatrix}\right).$$

For example, the function on $\mathbb{F}_2^n$ that takes $(1, \ldots, 1)$ to 1 and all other $v \in \mathbb{F}_2^n$ to 0 is

$$g_* = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) \otimes \cdots \otimes \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right);$$

its discrete Fourier transform is[12]

$$\widehat{g_*} = \widetilde{\mathsf{T}}g_* = \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right) \otimes \cdots \otimes \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right).$$

[12] Note that this aligns with the expression

$$\widehat{g_*}(u) = \sum_{v \in \mathbb{F}_2^n} (-1)^{(u,v)} g_*(v) = (-1)^{\sum_{j=1}^n u_j},$$

obtained from the more common definition (3.2) of the discrete Fourier transform given earlier.

4.1.1 Polynomials in the Variables $(-1)^{v_j}$ $(1 \leq j \leq n)$

We denote by $\mathscr{D}$ the $\mathbb{C}$-vector space of functions on $\mathbb{F}_2^n$. We can regard any function in $\mathscr{D}$ as a polynomial in the variables ιv_j $(1 \leq j \leq n)$, where $\iota : \mathbb{F}_2 \to \mathbb{C}$ is the map

$$\iota(0) = 0, \quad \iota(1) = 1.$$

Each of these variables satisfies $(\iota v_j)^2 = \iota v_j$. Hence the 2^n monomials in which each ιv_j appears to power 0 or 1 constitute a basis for $\mathscr{D}$. Each of these monomials has degree at most n, and g_* is the unique degree-n monomial among them.

Instead of working with polynomials in the variables ιv_j $(1 \leq j \leq n)$, we work with the discrete Fourier transforms $(-1)^{v_j}$ $(1 \leq j \leq n)$ of these variables.[13] Thus we consider $\mathscr{D}$ as the $\mathbb{C}$-vector space of polynomial functions Q in the variables

$$(-1)^{v_1}, \ldots, (-1)^{v_n},$$

where $v \in \mathbb{F}_2^n$. We denote by $\mathscr{D}_d$ the subspace of $\mathscr{D}$ consisting of degree-d homogeneous polynomials in the $(-1)^{v_j}$ $(1 \leq j \leq n)$ with each variable $(-1)^{v_j}$ in each term appearing to degree 0 or 1. We adopt the convention that $\mathscr{D}_d = \{0\}$ for $d < 0$.

The preceding discussion shows that any $Q \in \mathscr{D}$ may be interpreted as an element of $V_1^{\otimes n}$, and that the discrete Fourier transform $\hat{Q}$ of Q is equal to $\widetilde{\mathsf{T}}Q$. The action of $\mathfrak{sl}_2$ defined above gives rise to the following action on $\mathscr{D}$: if $M \in \mathfrak{sl}_2$ and $Q \in \mathscr{D}$, then the action of M on Q is given by

$$\left(\sum \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right) \otimes \cdots \otimes M \otimes \cdots \otimes \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)\right) Q.$$

Here, $\sum \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right) \otimes \cdots \otimes M \otimes \cdots \otimes \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$ denotes the operator equal to

$$(M \otimes \cdots \otimes \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)) + \cdots + (\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right) \otimes \cdots \otimes M \otimes \cdots \otimes \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)) + \cdots + (\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right) \otimes \cdots \otimes M),$$

the sum of n tensors, the j-th of which acts as M on the j-th factor and as the identity matrix $\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$ on the other factors.

4.1.2 Conjugation of $\widetilde{\mathsf{X}}$, $\widetilde{\mathsf{H}}$, and $\widetilde{\mathsf{Y}}$ by the Discrete Fourier Transform

Recall that we denote by $(\widetilde{\mathsf{X}}, \widetilde{\mathsf{H}}, \widetilde{\mathsf{Y}})$ the standard basis for $\mathfrak{sl}_2$, exhibited in (2.18). We define the operators X', H', and Y' to be the conjugates of $\widetilde{\mathsf{X}}$, $\widetilde{\mathsf{H}}$, and $\widetilde{\mathsf{Y}}$ by the discrete Fourier transform:

[13] Delsarte [Del78] uses the ιv_j basis, rather than the $(-1)^{v_j}$ basis. We depart from Delsarte's notation because the use of the $(-1)^{v_j}$ basis greatly simplifies our development.

$$\begin{aligned}\mathsf{X}' &:= \mathsf{T}^{-1}\mathsf{X}\mathsf{T} = \frac{1}{2}\left(\mathsf{H} - \mathsf{X} + \mathsf{Y}\right), \\ \mathsf{H}' &:= \mathsf{T}^{-1}\mathsf{H}\mathsf{T} = \mathsf{X} + \mathsf{Y}, \\ \mathsf{Y}' &:= \mathsf{T}^{-1}\mathsf{Y}\mathsf{T} = \frac{1}{2}\left(\mathsf{H} + \mathsf{X} - \mathsf{Y}\right). \end{aligned} \tag{4.2}$$

Conjugation by the Fourier transform operator T induces an isomorphism of Lie algebras

$$\mathsf{X} \longleftrightarrow \mathsf{X}', \quad \mathsf{H} \longleftrightarrow \mathsf{H}', \quad \mathsf{Y} \longleftrightarrow \mathsf{Y}'; \tag{4.3}$$

hence these operators $\mathsf{X}', \mathsf{H}', \mathsf{Y}'$ satisfy the commutation relations of (2.17), namely

$$[\mathsf{X}', \mathsf{Y}'] = \mathsf{H}', \quad [\mathsf{H}', \mathsf{X}'] = 2\mathsf{X}', \quad [\mathsf{H}', \mathsf{Y}'] = -2\mathsf{Y}'. \tag{4.4}$$

We write $\tilde{\mathsf{X}'}$, $\tilde{\mathsf{H}'}$, and $\tilde{\mathsf{Y}'}$ for operators

$$\begin{aligned}\tilde{\mathsf{X}'} &:= \sum \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right) \otimes \cdots \otimes \mathsf{X}' \otimes \cdots \otimes \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right) = \tilde{\mathsf{T}}^{-1}\left(\sum \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right) \otimes \cdots \otimes \mathsf{X} \otimes \cdots \otimes \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)\right)\tilde{\mathsf{T}}, \\ \tilde{\mathsf{H}'} &:= \sum \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right) \otimes \cdots \otimes \mathsf{H}' \otimes \cdots \otimes \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right) = \tilde{\mathsf{T}}^{-1}\left(\sum \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right) \otimes \cdots \otimes \mathsf{H} \otimes \cdots \otimes \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)\right)\tilde{\mathsf{T}}, \\ \tilde{\mathsf{Y}'} &:= \sum \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right) \otimes \cdots \otimes \mathsf{Y}' \otimes \cdots \otimes \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right) = \tilde{\mathsf{T}}^{-1}\left(\sum \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right) \otimes \cdots \otimes \mathsf{Y} \otimes \cdots \otimes \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)\right)\tilde{\mathsf{T}}, \end{aligned} \tag{4.5}$$

which represent the actions of X', H' and Y' on elements of $V_1^{\otimes n}$. The commutation relations of (4.4) extend to these operators, as well:

$$\left[\tilde{\mathsf{X}'}, \tilde{\mathsf{Y}'}\right] = \tilde{\mathsf{H}'}, \quad \left[\tilde{\mathsf{H}'}, \tilde{\mathsf{X}'}\right] = 2\tilde{\mathsf{X}'}, \quad \left[\tilde{\mathsf{H}'}, \tilde{\mathsf{Y}'}\right] = -2\tilde{\mathsf{Y}'}. \tag{4.6}$$

The relations (4.6) induce an isomorphism between $\mathfrak{sl}_2$ and the algebra generated by $\tilde{\mathsf{X}'}$, $\tilde{\mathsf{H}'}$, and $\tilde{\mathsf{Y}'}$.

Now, we have the following result immediately from the definition of $\tilde{\mathsf{H}'}$.

Lemma 4.1. *If $Q \in \mathscr{D}_d$, then $\tilde{\mathsf{H}'}Q = (n - 2d)Q$.*

Proof. The result follows directly, because the 1-eigenspace of H' is the span of $\{\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)\}$ and the (-1)-eigenspace of H' is the span of $\left\{\left(\begin{smallmatrix}1\\-1\end{smallmatrix}\right)\right\}$. □

For $Q \in \mathscr{D}_d$, we observe that $(\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right) \otimes \cdots \otimes \mathsf{X}' \otimes \cdots \otimes \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right))\,Q \in \mathscr{D}_{d-1}$, as we have

$$\mathsf{X}'\left(\begin{smallmatrix}1\\-1\end{smallmatrix}\right) = \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right) \quad \text{and} \quad \mathsf{X}'\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right) = \left(\begin{smallmatrix}0\\0\end{smallmatrix}\right).$$

Thus, $\tilde{\mathsf{X}'}Q = (\sum \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right) \otimes \cdots \otimes \mathsf{X}' \otimes \cdots \otimes \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right))\,Q \in \mathscr{D}_{d-1}$. We define the *space of degree-d discrete harmonic polynomials* by

$$\mathscr{D}_d^0 := \ker\left(\tilde{\mathsf{X}'} : \mathscr{D}_d \to \mathscr{D}_{d-1}\right). \tag{4.7}$$

We then define the *space of discrete harmonic polynomials*, denoted $\mathscr{D}^0$, to be the direct sum

$$\mathscr{D}^0 := \bigoplus_{d=0}^{n} \mathscr{D}_d^0 = \ker\left(\tilde{\mathsf{X}}' : \mathscr{D} \to \mathscr{D}\right). \tag{4.8}$$

4.2 Decomposition of Degree-d Discrete Homogeneous Polynomials

It is immediate from (4.6) that the operator $\tilde{\mathsf{H}}'$ maps $\mathscr{D}^0$ to itself, since if $Q \in \mathscr{D}^0$ then

$$\tilde{\mathsf{X}}'\tilde{\mathsf{H}}'Q = \left(\tilde{\mathsf{H}}'\tilde{\mathsf{X}}' - \left[\tilde{\mathsf{H}}', \tilde{\mathsf{X}}'\right]\right)Q = \left(\tilde{\mathsf{H}}'\tilde{\mathsf{X}}' - 2\tilde{\mathsf{X}}'\right)Q = 0.$$

The next lemma substantially refines this observation. Recall [Ser87, p. 18, Definition 1] that an element e of an $\mathfrak{sl}_2$ module is said to be *primitive of weight* λ if $e \neq 0$, $\mathsf{X}e = 0$, and $\mathsf{H}e = \lambda e$.

Lemma 4.2. *If $Q \in \mathscr{D}_d^0$, then Q is either zero or primitive of weight $n - 2d$ with respect to the representation of $\mathfrak{sl}_2$ induced by the action of $\tilde{\mathsf{X}}'$, $\tilde{\mathsf{H}}'$, and $\tilde{\mathsf{Y}}'$.*

Proof. The result is a direct consequence of Lemma 4.1 because all $Q \in \mathscr{D}^0$ satisfy $\tilde{\mathsf{X}}'Q = 0$. □

Corollary 4.3. *If $d > n/2$ then $\mathscr{D}_d^0 = \{0\}$.*

Proof. Since $\mathscr{D}$ is finite-dimensional, a primitive vector must have nonnegative weight. □

For $d \leq n/2$ and $k = 0, 1, \ldots, d$, we define $\mathscr{D}_d^k := (\tilde{\mathsf{Y}}')^k \mathscr{D}_{d-k}^0$.[14] Combining Lemma 4.2 with the representation theory of $\mathfrak{sl}_2$, we now obtain a decomposition result for $\mathscr{D}_d$ similar to that obtained for $\mathscr{P}_d$ in Proposition 2.5.

Proposition 4.4. *For any $d \leq n/2$, we have the following results.*

(1) *The map $\tilde{\mathsf{X}}' : \mathscr{D}_d \to \mathscr{D}_{d-1}$ is surjective.*
(2) *We have the direct sum decomposition $\mathscr{D}_d = \bigoplus_{k=0}^{d} \mathscr{D}_d^k = \mathscr{D}_d^0 \oplus \tilde{\mathsf{Y}}'\mathscr{D}_{d-1}$.*
(3) *For any nonzero $Q \in \mathscr{D}_d$, the space spanned by $\left\{(\tilde{\mathsf{Y}}')^j Q\right\}_{j=0}^{n-2d}$ is an irreducible $\mathfrak{sl}_2$-module isomorphic to $V_{n-2d} := \mathrm{Sym}^{n-2d}(V_1)$.*
(4) $\dim(\mathscr{D}_d^0) = \dim(\mathscr{D}_d) - \dim(\mathscr{D}_{d-1}) = \binom{n}{d} - \binom{n}{d-1}$.

[14] The notation $\mathscr{D}_d^k$ is consistent with the notation $\mathscr{D}_d^0$ for the space of degree-d discrete harmonic polynomials.

Proof. This follows quickly from Lemma 4.2 together with the finite-dimensional representation theory of $\mathfrak{sl}_2$; see for instance [Ser87, Chap. IV]. The first and second parts follow from the decomposition of any finite-dimensional $\mathfrak{sl}_2$-module as a direct sum of irreducible modules, together with the explicit action of $\mathfrak{sl}_2$ on each of its finite-dimensional irreducible modules [Ser87, Chap. IV, Theorems 2 and 3]. The third part follows from the structure of the irreducible representation generated by a primitive element of given weight [Ser87, Chap. IV, Corollary 2 of Theorem 1]. The fourth part follows from the first part. □

It also follows that $\tilde{\mathsf{X}}' : \mathscr{D}_d \to \mathscr{D}_{d-1}$ is *in*jective if $d-1 \geq n/2$, and thus an isomorphism if $n = 2d-1$; more generally, if $d \geq n/2$ then $\tilde{\mathsf{X}}'^{\,2d-n} : \mathscr{D}_d \to \mathscr{D}_{n-d}$ is an isomorphism.

5 The Generalized MacWilliams Identity for Harmonic Weight Enumerators

For a length-n binary linear code $C \subset \mathbb{F}_2^n$ and a discrete harmonic polynomial Q, the harmonic weight enumerator $W_{C,Q}(x,y)$ is defined by

$$W_{C,Q}(x,y) := \sum_{c \in C} Q(c) x^{n-\mathrm{wt}(c)} y^{\mathrm{wt}(c)}. \tag{5.1}$$

This function encodes the weights and distribution of the codewords of C, as the weighted theta functions of a lattice L encode the norms and distribution of the vectors of L.

We now derive a generalized MacWilliams identity for harmonic weight enumerators.

Theorem 5.1. *For any binary linear code $C \subset \mathbb{F}_2^n$ and $Q \in \mathscr{D}_d^0$, the harmonic weight enumerator $W_{C,Q}(x,y) = \sum_{c \in C} Q(c) x^{n-\mathrm{wt}(c)} y^{\mathrm{wt}(c)}$ satisfies the identity*

$$W_{C,Q}(x,y) = \left(-\frac{xy}{x^2-y^2}\right)^d \cdot \frac{2^{\frac{n}{2}+d}}{|C^\perp|} \cdot W_{C^\perp,Q}\left(\frac{x+y}{\sqrt{2}}, \frac{x-y}{\sqrt{2}}\right). \tag{5.2}$$

Theorem 5.1 was first proven by Bachoc [Bac99], via a purely combinatorial argument. Here, we give a new proof of this result in analogy with the proof of Theorem 2.6.

5.1 Derivation of the Identity

For $Q \in \mathscr{D}$, the function $Q(v)x^{n-\mathrm{wt}(v)}y^{\mathrm{wt}(v)}$ corresponds in the tensor representation to the function

$$\left(\left(\begin{smallmatrix} x & 0 \\ 0 & y \end{smallmatrix}\right)^{\otimes n}\right) Q.$$

Therefore, in analogy with the Gaussian operators G_t defined in Sect. 2.5, we introduce the operators

$$\mathsf{W} := \left(\begin{smallmatrix} x & 0 \\ 0 & y \end{smallmatrix}\right), \qquad \tilde{\mathsf{W}} := \mathsf{W}^{\otimes n}, \tag{5.3}$$

$$\mathsf{V} := \left(\begin{smallmatrix} x+y & 0 \\ 0 & x-y \end{smallmatrix}\right), \qquad \tilde{\mathsf{V}} := \mathsf{V}^{\otimes n}. \tag{5.4}$$

The operator $\tilde{\mathsf{W}}$ serves as a sort of "discrete Gaussian" for weight enumerators. Indeed, the weight enumerator $W_C(x,y)$ of a length-n binary linear code is given by

$$W_C(x,y) = \sum_{c \in C} \left(\tilde{\mathsf{W}} \cdot \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)^{\otimes n}\right)(c), \tag{5.5}$$

and the Fourier transform of $\tilde{\mathsf{W}} \cdot \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)^{\otimes n}$ is equal to $\tilde{\mathsf{V}} \cdot \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)^{\otimes n}$.

Lemma 5.2. *If $Q \in \mathscr{D}_d$, then we have*

$$(\tilde{\mathsf{V}}^{-1}\widetilde{\mathsf{T}}\tilde{\mathsf{W}})Q = \hat{Q},$$

where $\hat{Q} = \sum_{d'=0}^{d} \hat{Q}_{d'}$ with $\hat{Q}_{d'} \in \mathscr{D}_d$ for each d' ($0 \le d' \le d$) and

$$\hat{Q}_d = \left(-\frac{2xy}{x^2-y^2}\right)^d Q. \tag{5.6}$$

Proof. We proceed by strong induction on d. The base case $d = 0$ is immediate, so we suppose that the result holds for $Q \in \mathscr{D}_{d_1}$ for each nonnegative $d_1 \le d$, and deduce that the result holds also for $Q \in \mathscr{D}_{d+1}$.

The discrete Fourier transform operator is linear, hence it suffices to prove the result for the polynomials of the form $(-1)^{v_j} \cdot Q$ with $Q \in \mathscr{D}_d$. Now, we compute the value of $\tilde{\mathsf{V}}^{-1}\widetilde{\mathsf{T}}$ times

$$(-1)^{v_j} \cdot Q(v) \cdot x^{n-\mathrm{wt}(v)}y^{\mathrm{wt}(v)} = \tilde{\mathsf{W}} \cdot \left(\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \otimes \cdots \otimes \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) \otimes \cdots \otimes \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)\right) \cdot Q$$

explicitly. We find that

$$\begin{aligned} &\tilde{\mathsf{V}}^{-1}\widetilde{\mathsf{T}}\left(\tilde{\mathsf{W}}\left(\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \otimes \cdots \otimes \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) \otimes \cdots \otimes \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)\right) Q\right) \\ &\quad = \left(\tilde{\mathsf{V}}^{-1}\widetilde{\mathsf{T}}\tilde{\mathsf{W}}\right)\left(\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \otimes \cdots \otimes \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) \otimes \cdots \otimes \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)\right)\left(\tilde{\mathsf{V}}^{-1}\widetilde{\mathsf{T}}\tilde{\mathsf{W}}\right)^{-1}(\tilde{\mathsf{V}}^{-1}\widetilde{\mathsf{T}}\tilde{\mathsf{W}})Q \end{aligned}$$

$$= \left(\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \otimes \cdots \otimes \begin{pmatrix} 0 & \frac{x-y}{x+y} \\ \frac{x+y}{x-y} & 0 \end{pmatrix} \otimes \cdots \otimes \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \right) (\tilde{\mathsf{V}}^{-1}\widetilde{\mathsf{T}}\tilde{\mathsf{W}})Q$$
$$= \left(\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \otimes \cdots \otimes \begin{pmatrix} 0 & \frac{x-y}{x+y} \\ \frac{x+y}{x-y} & 0 \end{pmatrix} \otimes \cdots \otimes \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \right) \hat{Q}, \tag{5.7}$$

where the last equality in (5.7) follows on applying the inductive hypothesis to $(\tilde{\mathsf{V}}^{-1}\widetilde{\mathsf{T}}\tilde{\mathsf{W}})Q$.

It is clear that the right-hand side of (5.7) has maximal degree $d+1$, since $\hat{Q}$ is of degree d and

$$\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \otimes \cdots \otimes \begin{pmatrix} 0 & \frac{x-y}{x+y} \\ \frac{x+y}{x-y} & 0 \end{pmatrix} \otimes \cdots \otimes \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$$

is the identity on all but one coordinate. To finish the proof of the lemma, we compute the degree-$(d+1)$ term of (5.7). Now, since

$$\begin{pmatrix} 0 & \frac{x-y}{x+y} \\ \frac{x+y}{x-y} & 0 \end{pmatrix} \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) = \frac{x^2+y^2}{x^2-y^2} \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) - \frac{2xy}{x^2-y^2} \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right),$$

the degree-$(d+1)$ term of (5.7) must equal $-\frac{2xy}{x^2-y^2}\hat{Q}_d$.[15] The desired expression (5.6) then follows from the inductive hypothesis. □

Lemma 5.3. *If $Q \in \mathscr{D}^0$ and $\tilde{\mathsf{H}}'Q = \lambda \cdot Q$, then*

(1) $(\tilde{\mathsf{V}}^{-1}\widetilde{\mathsf{T}}\tilde{\mathsf{W}})\tilde{\mathsf{X}}'(\tilde{\mathsf{V}}^{-1}\widetilde{\mathsf{T}}\tilde{\mathsf{W}})^{-1}Q = 0$ *and*
(2) $(\tilde{\mathsf{V}}^{-1}\widetilde{\mathsf{T}}\tilde{\mathsf{W}})\tilde{\mathsf{H}}'(\tilde{\mathsf{V}}^{-1}\widetilde{\mathsf{T}}\tilde{\mathsf{W}})^{-1}Q = \lambda \cdot Q$.

Proof. Explicit computation gives

$$(\tilde{\mathsf{V}}^{-1}\widetilde{\mathsf{T}}\tilde{\mathsf{W}})\tilde{\mathsf{X}}'(\tilde{\mathsf{V}}^{-1}\widetilde{\mathsf{T}}\tilde{\mathsf{W}})^{-1} = -\frac{x^2-y^2}{2xy} \cdot \tilde{\mathsf{X}}', \tag{5.8}$$

$$(\tilde{\mathsf{V}}^{-1}\widetilde{\mathsf{T}}\tilde{\mathsf{W}})\tilde{\mathsf{H}}'(\tilde{\mathsf{V}}^{-1}\widetilde{\mathsf{T}}\tilde{\mathsf{W}})^{-1} = \tilde{\mathsf{H}}' + \frac{x^2+y^2}{xy} \cdot \tilde{\mathsf{X}}'. \tag{5.9}$$

The first and second results follow directly from (5.8) and (5.9), respectively, since

$$Q \in \mathscr{D}^0 = \ker(\tilde{\mathsf{X}}'). \qquad \square$$

Corollary 5.4. *The operators $\tilde{\mathsf{X}}'$ and $\tilde{\mathsf{H}}'$ act on $(\tilde{\mathsf{V}}^{-1}\widetilde{\mathsf{T}}\tilde{\mathsf{W}})\mathscr{D}^0$. The subspace $(\tilde{\mathsf{V}}^{-1}\widetilde{\mathsf{T}}\tilde{\mathsf{W}})\mathscr{D}^0_d$ of $\mathscr{D}$ is the intersection of* $\ker(\tilde{\mathsf{X}}')$ *and the $(n-2d)$-eigenspace of $\tilde{\mathsf{H}}' + \frac{x^2+y^2}{xy}\tilde{\mathsf{X}}'$ in $(\tilde{\mathsf{V}}^{-1}\widetilde{\mathsf{T}}\tilde{\mathsf{W}})\mathscr{D}^0_d$.*

[15] Here, $\hat{Q}_d$ is the degree-d term of $\hat{Q}$, as in the lemma statement.

5.1.1 Proof of the Generalized MacWilliams Identity

As a final step en route to Theorem 5.1, we prove an expression analog to Proposition 2.11 for the discrete Fourier transform of the product of $\widetilde{\mathsf{W}}$ and a discrete harmonic polynomial $Q \in \mathscr{D}_d^0$.

Proposition 5.5. *If $Q \in \mathscr{D}_d^0$, then*

$$\left(\tilde{\mathsf{V}}^{-1}\widetilde{\mathsf{T}}\tilde{\mathsf{W}}\right)Q = \left(-\frac{2xy}{x^2-y^2}\right)^d Q. \tag{5.10}$$

Proof. From Corollary 5.4, we see that $\left(\tilde{\mathsf{V}}^{-1}\widetilde{\mathsf{T}}\tilde{\mathsf{W}}\right)Q$ is in both $\mathscr{D}^0$ and (since then $\widetilde{\mathsf{X}}'Q = 0$) the $(n-2d)$-eigenspace of $\widetilde{\mathsf{H}}'$. That is, $\left(\tilde{\mathsf{V}}^{-1}\widetilde{\mathsf{T}}\tilde{\mathsf{W}}\right)Q \in \mathscr{D}_d^0$. The result then follows immediately from Lemma 5.2. □

Finally, we obtain the generalized MacWilliams identity by combining Proposition 5.5 with the discrete Poisson summation formula (Theorem 3.1).

Proof of Theorem 5.1. We obtain the discrete Fourier transform of $\tilde{\mathsf{W}}Q$ from Proposition 5.5:

$$\widetilde{\mathsf{T}}\left(\tilde{\mathsf{W}}Q\right) = \left(\frac{-2xy}{x^2-y^2}\right)^d \tilde{\mathsf{V}}Q = \left(\frac{-2xy}{x^2-y^2}\right)^d \cdot 2^{n/2} \cdot \left(\begin{pmatrix} \frac{x+y}{\sqrt{2}} & 0 \\ 0 & \frac{x-y}{\sqrt{2}} \end{pmatrix}^{\otimes n}\right) \cdot Q. \tag{5.11}$$

The desired formula (5.2) then follows directly from (5.11), upon applying Theorem 3.1. □

Remark. One interesting consequence of Theorem 5.1 is that for any $Q \in \mathscr{D}_d^0$ the harmonic weight enumerator $W_{C,Q}x, y$ is a multiple of $x(xy)^d$

Corollary 5.6. *For C a binary linear code and $Q \in \mathscr{D}_d^0$,*

$$\frac{W_{C,Q}(x,y)}{(xy)^d}$$

is a polynomial in the variables x, y.

Proof. By Theorem 5.1,

$$\frac{W_{C,Q}(x,y)}{(xy)^d} = \left(-\frac{1}{x^2-y^2}\right)^d \cdot \frac{2^{\frac{n}{2}+d}}{|C^\perp|} \cdot W_{C^\perp,Q}\left(\frac{x+y}{\sqrt{2}}, \frac{x-y}{\sqrt{2}}\right). \tag{5.12}$$

The left-hand side is a rational function in x and y whose denominator divides $(xy)^d$, and the right-hand side is a rational function whose denominator divides $(x^2-y^2)^d$. Since $(xy)^d$ and $(x^2-y^2)^d$ are relatively prime, (5.12) is an identity between polynomials in x and y. □

As we see at the end of Sect. 7, Corollary 5.6 also follows directly from the $\mathfrak{sl}_2$ development of discrete harmonic polynomials.

5.2 A Generalization of Gleason's Theorem

In addition to the generalized MacWilliams identity, Bachoc [Bac99] obtained a harmonic weight enumerator generalization of Gleason's theorem. As we use this result in Sect. 7, we state it here.

Theorem 5.7 (Bachoc [Bac99]). *Let C be a Type II code of length n and let $Q \in \mathscr{D}_d^0$. Then, the harmonic weight enumerator $W_{C,Q}(x,y)$ is an element of the principal module $\mathbb{C}[\varphi_8, \xi_{24}]\psi_d$ for the polynomial algebra $\mathbb{C}[\varphi_8, \xi_{24}]$, whose generator is given by*

$$\psi_d := \begin{cases} 1 & d \equiv 0 \bmod 4, \\ x^3y^3(x^4-y^4)^2(x^8-y^8)(x^8-34x^4y^4+y^8) & d \equiv 1 \bmod 4, \\ x^2y^2(x^4-y^4)^2 & d \equiv 2 \bmod 4, \\ xy(x^8-y^8)(x^8-34x^4y^4+y^8) & d \equiv 3 \bmod 4. \end{cases} \tag{5.13}$$

The degree-12 polynomial ψ_2 is a square root of ξ_{24}; thus the harmonic enumerators that can arise for even d are elements of the polynomial ring $\mathbb{C}[\varphi_8, \psi_2]$, which is the ring of invariants for a complex reflection group contained with index 2 in G_{II} (see the Appendix). For odd d, the polynomials ψ_d are more complicated covariants of G_{II}; we have $\psi_1 = \psi_2\psi_3$ and $\psi_3^2 = \psi_2(\varphi_8^3 - 108\xi_{24})$.

6 Zonal Harmonic Polynomials

We now introduce the *zonal harmonic polynomials*, a class $\mathscr{Z}\mathscr{D}^0$ of discrete harmonic polynomials analog to the zonal spherical harmonics mentioned at the end of Sect. 2. Specifically, we fix some $\dot{v} \in \mathbb{F}_2^n$ and some d with $0 \leq d \leq \mathrm{wt}(\dot{v})$, and determine the space $\mathscr{Z}\mathscr{D}_d^0 \subset \mathscr{D}_d^0$ of degree-d discrete harmonic polynomials invariant under coordinate permutations fixing $\dot{v}$.

6.1 Preliminaries

Throughout, we fix $\dot{v} \in \mathbb{F}_2^n$. We denote by $\mathscr{Z}\mathscr{D}_d \subset \mathscr{D}_d$ the space of degree-d discrete homogeneous polynomials invariant under the group of coordinate permutations fixing $\dot{v}$, and set $\mathscr{Z}\mathscr{D}_d^0 := \mathscr{Z}\mathscr{D}_d \cap \mathscr{D}_d^0$. We say that a polynomial in $\mathscr{Z}\mathscr{D}_d^0$ is a *zonal* *harmonic polynomial of degree* d, and we define the space $\mathscr{Z}\mathscr{D}^0$ of *zonal harmonic polynomials* by

$$\mathscr{Z}\mathscr{D}^0 := \bigoplus_{d=0}^{\mathrm{wt}(\dot{v})} \mathscr{Z}\mathscr{D}_d^0. \tag{6.1}$$

6.1.1 Generators of $\mathscr{Z}\mathscr{D}_d$

We now fix some d with $0 \leq d \leq \mathrm{wt}(\dot{v})$ and let

$$\mathrm{C}_{1;\dot{v}} := \{j : \dot{v}_j = 1\}, \quad \mathrm{C}_{0;\dot{v}} := \{j : \dot{v}_j = 0\}.$$

Now, we denote by $Q_{d,k;\dot{v}}(v)$ the degree-d discrete polynomial

$$\begin{aligned} Q_{d,k;\dot{v}}(v) &:= \sum_{\substack{\{j_1,\dots,j_k\}\subseteq \mathrm{C}_{1;\dot{v}} \\ \{j_{k+1},\dots,j_d\}\subseteq \mathrm{C}_{0;\dot{v}}}} (-1)^{(v_{j_1}+\cdots+v_{j_k})+(v_{j_{k+1}}+\cdots+v_{j_d})} \\ &= \sum_{\substack{\{j_1,\dots,j_k\}\subseteq \mathrm{C}_{1;\dot{v}} \\ \{j_{k+1},\dots,j_d\}\subseteq \mathrm{C}_{0;\dot{v}}}} (-1)^{v_{j_1}}\cdots(-1)^{v_{j_k}}\cdot(-1)^{v_{j_{k+1}}}\cdots(-1)^{v_{j_d}} \in \mathscr{D}_d. \end{aligned} \tag{6.2}$$

The sum is nonempty for all d ($0 \leq d \leq \mathrm{wt}(\dot{v})$) since $|\mathrm{C}_{1;\dot{v}}| = \mathrm{wt}(\dot{v})$ and $|\mathrm{C}_{0;\dot{v}}| = n - \mathrm{wt}(\dot{v})$.

By construction, it is clear that $Q_{d,k;\dot{v}} \in \mathscr{Z}\mathscr{D}_d$. Conversely, we have the following lemma.

Lemma 6.1. *The polynomials* $\{Q_{d,k;\dot{v}}\}_{k=0}^{d}$ *generate* $\mathscr{Z}\mathscr{D}_d$.

Proof. The result follows immediately from the requirement that any $Q \in \mathscr{Z}\mathscr{D}_d$ be invariant under all permutations simultaneously permuting the $\mathrm{wt}(\dot{v})$ nonzero coordinates of $\dot{v}$ and the $n-\mathrm{wt}(\dot{v})$ vanishing coordinates in $\dot{v}$, together with the fact that the multilinear monomials in the variables $(-1)^{v_j}$ are a basis for $\mathscr{D}$. □

Additionally, we have a combinatorial formula for $Q_{d,k;\dot{v}}(v)$.

Proposition 6.2. *We have*

$$Q_{d,k;\dot{v}}(v) = \left(\sum_{i=0}^{k}(-1)^i\binom{\mathrm{wt}(v\cap\dot{v})}{i}\binom{\mathrm{wt}(\dot{v})-\mathrm{wt}(v\cap\dot{v})}{k-i}\right)\times$$
$$\left(\sum_{i=0}^{d-k}(-1)^i\binom{\mathrm{wt}(v)-\mathrm{wt}(v\cap\dot{v})}{i}\binom{(n-\mathrm{wt}(\dot{v}))-(\mathrm{wt}(v)-\mathrm{wt}(v\cap\dot{v}))}{d-k-i}\right). \tag{6.3}$$

The proof of Proposition 6.2 is immediately obtained from evaluation of the expression (6.2) for $Q_{d,k;\dot{v}}$.

6.1.2 The Action of $\tilde{\mathsf{X}}'$ on $Q_{d,k;\dot{v}}$

Now, we determine the action of $\tilde{\mathsf{X}}'$ on the polynomials $\{Q_{d,k;\dot{v}}\}_{k=0}^{\mathrm{wt}(\dot{v})}$.

Lemma 6.3. *We have*

$$\tilde{\mathsf{X}}'Q_{d,k;\dot{v}} = \big((n-\mathrm{wt}(\dot{v}))-(d-k-1)\big)Q_{d-1,k;\dot{v}} + \big(\mathrm{wt}(\dot{v})-(k-1)\big)Q_{d-1,k-1;\dot{v}}. \tag{6.4}$$

Proof. First, we observe that

$$\tilde{\mathsf{X}}' \cdot \left((-1)^{v_{j_1}+\cdots+v_{j_d}}\right) = \sum_{\ell=1}^{d}(-1)^{v_{j_0}+v_{j_1}+\cdots+v_{j_{\ell-1}}+v_{j_{\ell+1}}+\cdots+v_{j_d}+v_{j_{d+1}}}, \tag{6.5}$$

where we have used the convention that $v_{j_0} = 0 = v_{j_{d+1}}$.[16] It then follows from (6.5) that

$$\tilde{\mathsf{X}}'Q_{d,k;\dot{v}} = b_k \cdot Q_{d-1,k;\dot{v}} + b_{k-1}\cdot Q_{d-1,k-1;\dot{v}}$$

for constants $b_k, b_{k-1} \in \mathbb{Z}$. To see that

$$b_{k-1} = \mathrm{wt}(\dot{v}) - (k-1),$$

we observe that each monomial term in $Q_{d-1,k;\dot{v}}$ can arise from $\mathrm{wt}(\dot{v})-(k-1)$ different monomial terms in $Q_{d,k;\dot{v}}$. Likewise, we obtain

$$b_k = (n-\mathrm{wt}(\dot{v})) - (d-k-1). \qquad \square$$

6.2 *Determination of the Zonal Harmonic Polynomials*

We now combine Lemma 6.1 and Lemma 6.3 to characterize $\mathscr{Z}\mathscr{D}_d^0$.

Proposition 6.4. *If $Q \in \mathscr{Z}\mathscr{D}_d^0$, then $Q = b_0 \cdot Q_{d;\dot{v}}$ for some constant $b_0 \in \mathbb{C}$, where*

$$Q_{d;\dot{v}}(v) := \sum_{k=0}^{d}(-1)^k\left(\prod_{\ell=0}^{k-1}\frac{(n-\mathrm{wt}(\dot{v}))-(d-\ell-1)}{\mathrm{wt}(\dot{v})-\ell}\right)Q_{d,k;\dot{v}}(v). \tag{6.6}$$

[16] To avoid having to adopt this convention, we could have used the slightly more standard notation $\sum_{\ell=1}^{d}(-1)^{v_{j_1}+\cdots+\widehat{v_{j_\ell}}+\cdots+v_{j_d}}$. We opt not to use this notation because it conflicts with our usage of $\hat{\cdot}$ for the discrete Fourier transform.

Proof. We consider some $Q \in \mathscr{X}\mathscr{D}_d^0 = \mathscr{X}\mathscr{D}_d \cap \mathscr{D}_d^0$. By Lemma 6.1, there exist constants $\{b_k\}_{k=0}^{\mathrm{wt}(\dot{v})} \subset \mathbb{C}$ such that

$$Q = \sum_{k=0}^{\mathrm{wt}(\dot{v})} b_k \cdot Q_{d,k;\dot{v}}.$$

Since $Q \in \mathscr{D}_d^0$, we have

$$\begin{aligned} 0 = \tilde{\mathsf{X}}' Q = \tilde{\mathsf{X}}' \left(\sum_{k=0}^{\mathrm{wt}(\dot{v})} b_k \cdot Q_{d,k;\dot{v}} \right) = \sum_{k=0}^{\mathrm{wt}(\dot{v})} b_k \cdot \tilde{\mathsf{X}}' Q_{d,k;\dot{v}} \\ = \sum_{k=0}^{\mathrm{wt}(\dot{v})} b_k \cdot \big(((n - \mathrm{wt}(\dot{v})) - (d-k-1)\big) Q_{d-1,k;\dot{v}} + \big(\mathrm{wt}(\dot{v}) - (k-1)) Q_{d-1,k-1;\dot{v}} \big) \\ = \sum_{k=0}^{\mathrm{wt}(\dot{v})} \Big(b_k \big((n - \mathrm{wt}(\dot{v})) - (d-k-1) \big) + b_{k+1} (\mathrm{wt}(\dot{v}) - (k)) \Big) Q_{d-1,k;\dot{v}}. \end{aligned}$$

(The penultimate equality follows from Lemma 6.3.) By comparing coefficients, we then obtain

$$b_{k+1} = -\frac{(n - \mathrm{wt}(\dot{v})) - (d-k-1)}{\mathrm{wt}(\dot{v}) - k}\, b_k$$

for each k ($0 \le k \le \mathrm{wt}(\dot{v}) - 1$); the result follows. □

Corollary 6.5. *For each d ($0 \le d \le \mathrm{wt}(\dot{v})$), we have* $\dim(\mathscr{X}\mathscr{D}_d^0) = 1$.

7 *t*-Designs and Extremal Type II Codes

A *t-(n, w, λ)-design* is a (possibly empty[17]) collection D of distinct w-element subsets of $\{1, \ldots, n\}$ with the property that $|\{S' \in D : S \subseteq S'\}| = \lambda$ for every $S \subset \{1, \ldots, n\}$ with $|S| = t$. This generalizes the notion of a *Steiner system*, which is a t-$(n, w, 1)$ design. For example, the codewords of weight 4 in the extended Hamming code form a 3-$(8, 4, 1)$-design, and the codewords of weight 12 in the extended binary Golay code form a 5-$(24, 12, 48)$ design. We shall see that these are special cases of behavior common to all extremal Type II codes. When n, w, and λ are undetermined or clear from context, we omit the qualifier "(n, w, λ)" and simply refer to a t-(n, w, λ)-design as a *t-design*. (See [CvL91] for more about t-designs, their uses and their relations with error-correcting codes.)

[17] Again we allow $D = \emptyset$, which is a t-$(n, w, 0)$-design for all t and w. As with spherical designs, for most applications only nonempty D are of interest, but allowing empty designs simplifies the statements of the results relating codes with combinatorial designs.

7.1 An Equivalent Characterization of t-Designs

Each $S' \in D$ may be represented by its *indicator vector* $(c_1, \ldots, c_n)$, in which $c_j = 1$ if and only if $j \in S'$. Thus, a t-(n, w, λ)-design D corresponds to a subset of the *Hamming sphere of radius* w,

$$\sigma_w := \{v \in \mathbb{F}_2^n : \mathrm{wt}(v) = w\}.$$

We henceforth treat this representation of D as completely equivalent to the setwise representation of D, using the relevant terminology interchangeably.

We now introduce the following equivalent characterization of t-designs.

Proposition 7.1. *A set $D \subseteq \sigma_w$ is a t-design if and only if*

$$\sum_{v \in D} Q(v) = 0$$

for all $Q \in \bigcup_{d=1}^{t} \mathscr{D}_d^0$.

Proposition 7.1 is equivalent to Theorem 7 of Delsarte [Del78]. Our development of $\mathscr{D}^0$ leads to a new proof of this result, which we present below. In Sect. 7.2, we apply Proposition 7.1 to prove a special case of the Assmus–Mattson theorem [AM69].

Throughout this section, we write χ_X for the characteristic function of the set X, and recall that $\widetilde{\mathsf{H}}$ denotes the action of H on $V_1^{\otimes n}$,

$$\widetilde{\mathsf{H}} := \sum \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \otimes \cdots \otimes \mathsf{H} \otimes \cdots \otimes \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right). \tag{7.1}$$

We begin with a lemma regarding projections of functions $Q \in \mathscr{D}$ to the Hamming sphere σ_w.

Lemma 7.2. *For $Q \in \mathscr{D}$, we have $\chi_{\sigma_w} Q = \pi_{n-2w}(Q)$, where $\pi_{n-2w}(Q)$ is the projection of Q to the $n - 2w$ eigenspace of the action of $\widetilde{\mathsf{H}}$ on $V_1^{\otimes n}$.*

Proof. This is immediate because the 1- and (-1)-eigenspaces of H are respectively spanned by $\{\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)\}$ and $\{\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)\}$. □

We now demonstrate Proposition 7.1.

Proof of Proposition 7.1. We denote by $\mathscr{O}$ the subset of $V_1^{\otimes n}$ consisting of tensor products of t copies of $\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ and $n - t$ copies of $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$. It is clear that $\mathscr{O}$ spans $\mathscr{D}_d$ for any d ($0 \leq d \leq t$). Now, the set D is a t-design if and only if, for all $R \in \mathscr{O}$,

$$|\sigma_w| \cdot (\chi_D, R) = |D| \cdot (\chi_{\sigma_w}, R),$$

where $|\cdot|$ is the cardinality function and $(\cdot,\cdot)$ is the inner product. It therefore suffices to show that the set $\{\chi_{\sigma_w} R : R \in \mathscr{O}\}$ is spanned by

$$\bigcup_{d=0}^{t} \{\chi_{\sigma_w} Q : Q \in \mathscr{D}_d^0\}.$$

By the second part of Proposition 4.4, any $R \in \mathscr{O}$ may be written in the form

$$R = \sum_{j=0}^{t} (\tilde{\mathsf{Y}}')^j Q_j,$$

with $Q_j \in \bigoplus_{d=0}^{t-j} \mathscr{D}_d^0$. By Lemma 7.2 and the hypothesis, it then only remains to demonstrate that $\pi_{n-2w}((\tilde{\mathsf{Y}}')^j Q_j)$ and $\pi_{n-2w}(Q_j)$ are related by a constant factor i.e., that for each $j = 0, \ldots, t$, we have

$$\pi_{n-2w}((\tilde{\mathsf{Y}}')^j Q_j) = b \cdot \pi_{n-2w}(Q_j) \tag{7.2}$$

for some constant b depending on both j and t.

Now, given any $Q \in \mathscr{D}_d^0$, we see by the third part of Proposition 4.4 that the polynomials $(\tilde{\mathsf{Y}}')^k Q$ $(0 \le k \le n-2d)$ span an irreducible representation of $\mathfrak{sl}_2$ isomorphic to V_{n-2d}. We may regard this representation as degree $n-2d$ homogeneous part of the polynomial algebra $\mathbb{C}[u_0, u_1]$ with generators u_0, u_1 and with actions of $\mathsf{X}', \mathsf{H}', \mathsf{Y}'$ respectively given by

$$u_0' \frac{\partial}{\partial u_1'}, \quad \left(u_0' \frac{\partial}{\partial u_0'} - u_1' \frac{\partial}{\partial u_1'}\right), \quad u_1' \frac{\partial}{\partial u_0'}, \tag{7.3}$$

where $u_0' = u_0 + u_1$ and $u_1' = u_0 - u_1$. With this identification, we may take $Q = (u_0')^{n-2d}$, as

$$Q \in \ker\left(\tilde{\mathsf{X}}' : \mathscr{D}_d^0 \to \mathscr{D}_{d-1}^0\right).$$

We now show that $\pi_{n-2w}((\tilde{\mathsf{Y}}')^k Q)$ and $\pi_{n-2w}(Q)$ are related by a constant factor for each k $(0 \le k \le n-2d)$; the desired expression (7.2) follows. We observe that H acts as

$$u_0 \frac{\partial}{\partial u_0} - u_1 \frac{\partial}{\partial u_1}. \tag{7.4}$$

Therefore, $\pi_{n-2w}(Q) = \pi_{n-2w}((u_0+u_1)^{n-2d})$ equals $\binom{n-2d}{w-d} u_0^{n-(d+w)} u_1^{w-d}$. To see this, note that $\pi_{n-2w}((u_0+u_1)^{n-2d}) = \binom{n-2d}{b_1} u_0^{b_0} u_1^{b_1}$ with $b_0 + b_1 = n-2d$ and $b_0 - b_1 = n-2w$. (The latter statement follows from the definition of $\pi_{n-2w}(\cdot)$.) Likewise,

$$\pi_{n-2w}((\tilde{\mathsf{Y}}')^k Q) = \pi_{n-2w}((\tilde{\mathsf{Y}}')^k (u_0+u_1)^{n-2d})$$

is the $u_0^{n-(d+w)} u_1^{w-d}$ component of $(\tilde{\mathsf{Y}}')^k (u_0+u_1)^{n-2d}$. Since this component is equal to

$$u_0^{n-(d+w)} u_1^{w-d} = \pi_{n-2w}(Q)$$

up to a constant factor, we are done. □

Remark. The constant relating $\pi_{n-2w}((\tilde{\mathrm{Y}}')^k Q)$ and $\pi_{n-2w}(Q)$ in the proof of Proposition 7.1 is obtained directly from the identification of $\left\{(\tilde{\mathrm{Y}}')^k Q\right\}_{k=0}^{n-2d}$ with V_{n-2d}. Consequently, this constant is independent of the choice of $Q \in \mathscr{D}_d^0$.

Proposition 7.1 leads to another equivalent characterization of t-designs which makes the analogy between t-designs and spherical t-designs explicit. We have the following corollary, which is equivalent to Theorem 6 of Delsarte [Del78].

Corollary 7.3. *A set $D \subseteq \sigma_w$ is a t-design if and only if*

$$\sum_{v \in D} Q(v) = \frac{|D|}{|\sigma_w|} \sum_{v \in \sigma_w} Q(v) \tag{7.5}$$

for all $Q \in \bigcup_{d=0}^{t} \mathscr{D}_d$.

Proof. As (7.5) is immediate when Q is constant, the result follows directly from Proposition 7.1 and the second part of Proposition 4.4. □

Finally, we note that the proof of Proposition 7.1 shows that each $Q \in \mathscr{D}_d^0$ is supported on $\bigcup_{w=d}^{n-d} \sigma_w$. This fact leads to a second proof of Corollary 5.6.

Alternate Proof of Corollary 5.6. As $Q \in \mathscr{D}_d^0$ is supported on $\bigcup_{w=d}^{n-d} \sigma_w$, we know that

$$W_{C,Q}(x,y) = \sum_{w=0}^{n} \left(\sum_{\substack{c \in C \\ \mathrm{wt}(c)=w}} Q(c) \right) x^{n-w} y^w = \sum_{w=d}^{n-d} \left(\sum_{\substack{c \in C \\ \mathrm{wt}(c)=w}} Q(c) \right) x^{n-w} y^w.$$

The result then follows immediately. □

7.2 The Extremal Type II Code Case of the Assmus–Mattson Theorem

To illustrate the power of Proposition 7.1, we now prove the Assmus–Mattson theorem [AM69] in the important special case of an *extremal Type II code*, that is, a binary linear code C whose minimal (nonzero) weight

$$\min(C) := \min\{\mathrm{wt}(c) : c \in C,\ c \neq 0\}$$

attains the upper bound $4\lfloor n/24 \rfloor + 4$ derived by Mallows and Sloane [MS73] from Gleason's theorem for Type II codes.

For any code C and integer w, we define C_w to be the subset of C consisting of codewords of weight w. For $n \equiv 0 \bmod 8$, we define $\mathrm{t}(n)$ by

$$\mathrm{t}(n) := \begin{cases} 5 & n \equiv 0 \bmod 24, \\ 3 & n \equiv 8 \bmod 24, \\ 1 & n \equiv 16 \bmod 24. \end{cases} \tag{7.6}$$

Theorem 7.4. *If C is an extremal Type II code of length n, then C_w is a t-design for each $t \leq \mathrm{t}(n)$ and any w.*

By Proposition 7.1, this theorem follows quickly from the following result, which is slightly more general and is a coding-theoretic analog of the $r > 0$ part of Theorem 2.15.

Proposition 7.5. *If C is an extremal Type II code of length n, then for any w and any choices of $d \in \{1, \ldots, \mathrm{t}(n)\} \cup \{\mathrm{t}(n) + 2\}$ and $Q \in \mathscr{D}_d^0$, we have*

$$\sum_{c \in C_w} Q(c) = 0.$$

Proposition 7.5 was originally proven by Calderbank and Delsarte [CD93]. Here, we demonstrate how Proposition 7.5 follows quickly from Theorem 5.7. This approach is due to Bachoc [Bac99]. Our exposition of this argument slightly expands that of Bachoc [Bac99], which demonstrates only four cases of the result.

Proof of Proposition 7.5. We let $d \in \{1, \ldots, \mathrm{t}(n)\} \cup \{\mathrm{t}(n) + 2\}$ and $Q \in \mathscr{D}_d^0$. Then, we consider the harmonic weight enumerator $W_{C,Q}(x, y)$. By Theorems 5.1 and 5.7, we see that $W_{C,Q}(x, y)/(xy)^d$ is of the form $\xi_{24}^{(\min(C)-d-b_d)/4} \cdot f$, where b_d equals the valuation at y of ψ_d. This factor arises because the valuation at y of $W_{C,Q}(x, y)$ is at least $\min(C)$.

We see that if $W_{C,Q}(x, y)$ is nonzero, then it has degree equal to

$$(n \bmod 24) + 4d - 24 \tag{7.7}$$

if $d \equiv 0 \bmod 2$. Similarly, f has degree

$$(n \bmod 24) + 4d - 36 \tag{7.8}$$

if $d \equiv 1 \bmod 2$. Since (7.7) and (7.8) are always negative for $d \in \{1, \ldots, \mathrm{t}(n)\} \cup \{\mathrm{t}(n) + 2\}$, we must have $f \equiv 0$, whence

$$\sum_{w=0}^{n} \left(\sum_{c \in C_w} Q(c) \right) x^{n-w} y^w = W_{C,Q}(x, y) \equiv 0. \qquad \square$$

We note the following special case of Proposition 7.1 which is relevant to our proofs of configuration results in Sect. 9.

Corollary 7.6. *If C is an extremal Type II code of length n and $w > 0$, then we have*

$$\sum_{c \in C_w} Q_{t;\dot{v}}(c) = 0$$

for any $t \in \{1, \ldots, \mathrm{t}(n)\} \cup \{\mathrm{t}(n) + 2\}$.

Remarks. As Bachoc [Bac99] illustrates, it is possible to prove the full Assmus–Mattson theorem with a harmonic weight enumerator argument similar to that used in the proof of Proposition 7.5. We have focused on the case of an extremal Type II code because the full force of Corollary 7.6 is required in Sect. 9.

8 The Koch Condition on Type II Codes of Length 24

8.1 Tetrad Systems

For any code C and integer w, we define $\mathcal{C}_w(C)$ to be the linear subcode of C generated by the set $C_w \subseteq C$ consisting of codewords of weight w. (This notation is analogous to that of Ozeki [Oze86b] for lattices.)

For a doubly even code $C \subset \mathbb{F}_2^n$, the set C_4 is called the *tetrad system* of C. In analogy with the theory of root systems for lattices, the code $\mathcal{C}_4(C)$ generated by C_4 is called the *tetrad subcode* of C, and if $\mathcal{C}_4(C) = C$ then C is called a *tetrad code*. The irreducible tetrad codes are exactly

- the codes d_{2k} ($k \geq 2$), consisting of all words $c \in \mathbb{F}_2^{2k}$ of doubly even weight such that $c_{2j-1} = c_{2j}$ for each $j = 1, 2, \ldots, k$;
- the $[7, 3, 4]$ dual Hamming code, called e_7 in this context; and
- the $[8, 4, 4]$ extended Hamming code, here called e_8

(see [Koc87]). We use the names d_{2k}, e_7, e_8 because the Construction A lattices $L_{d_{2k}}$, L_{e_7}, and L_{e_8} are isomorphic with the root lattices D_{2k}, E_7, and E_8 respectively.

Analogous to the Coxeter number of an irreducible root system, we define the *tetrad number* $\eta(C)$ of an irreducible tetrad code C of length m to be $|C_4|/m$. A quick computation shows that each of the m coordinates of C takes the value 1 on exactly $4\eta(C)$ words in C_4, and that $\eta(d_{2k}) = (k-1)/4$ for each k, while $\eta(e_7) = 1$ and $\eta(e_8) = 7/4$.

8.2 Koch's Tetrad System Condition

Through appeal to the condition of Venkov [Ven80] restricting the possible root systems of Type II lattices of rank 24, Koch [Koc87] obtained a condition on the tetrad systems of Type II codes of length 24. Specifically, he showed the following result.

Proposition 8.1. *If C is a Type II code of length 24, then C has one of the following nine tetrad systems:*

$$\emptyset, \quad 6d_4, \quad 4d_6, \quad 3d_8, \quad 2d_{12}, \quad d_{24}, \quad 2e_7 + d_{10}, \quad 3e_8, \quad e_8 + d_{16}. \tag{8.1}$$

Koch recovered this condition from the Niemeier [Nie73] classification of Type II lattices of rank 24 via Construction A. The condition is also a consequence of the classification of Type II codes of length 24 given by Pless and Sloane [PS75].

8.3 A Purely Coding-Theoretic Proof of Koch's Condition

Here, we present our proof [EK10] of Proposition 8.1 using the theory of harmonic weight enumerators. This argument is closely analogous to that of Venkov [Ven80] for the corresponding criterion on root systems of Type II lattices of rank 24. We thus begin with a coding-theoretic analog of [Ven80, Proposition 1].

Lemma 8.2. *If C is a Type II code of length 24, then*

- *either $C_4 = \emptyset$ or for each j $(1 \le j \le 24)$ there exists $c \in C_4$ such that $c_j = 1$, and*
- *each irreducible component of $\mathcal{C}_4(C)$ has tetrad number equal to $|C_4|/24$.*

Proof of Lemma 8.2. For each j $(1 \le j \le n)$, we denote by $Q_{1,j,n}$ the discrete harmonic polynomial defined by

$$Q_{1,j,n}(v) := n \cdot (-1)^{v_j} - \sum_{k=1}^{n} (-1)^{v_k} \in \mathscr{D}_1^0. \tag{8.2}$$

As in the proof of Proposition 7.5, we see that the harmonic weight enumerator

$$W_{C,Q_{1,j,24}}(x,y) = \sum_{w=0}^{24} \left(\sum_{c \in C_w} Q_{1,j,24}(c) \right) x^{24-w} y^w \tag{8.3}$$

vanishes for each j $(1 \le j \le 24)$. We then obtain

$$\sum_{c \in C_4} (8 - 48c_j) = 0 \tag{8.4}$$

for each j $(1 \le j \le 24)$, since the left-hand side of (8.4) is the $x^{20}y^4$ coefficient of the right-hand side of (8.3). Reorganizing (8.4) shows that

$$|\{c \in C_4 : c_j = 1\}| = |C_4|/6. \tag{8.5}$$

The first part of the lemma then follows. In the case that $C_4 \neq \emptyset$, we also obtain from (8.5) that each irreducible component of $\mathcal{C}_4(C)$ has tetrad number $\frac{1}{4}|C_4|/6 = |C_4|/24$. □

Remark. Since the discrete harmonic polynomial $Q_{1,j,n}$ has degree 1 and is invariant under the coordinate permutations that fix j, it is proportional to the zonal harmonic polynomial $Q_{1;\dot{v}}$ where $\dot{v}$ is the j-th unit vector.

Proof of Proposition 8.1. As noted in Sect. 8.1, there is at most one tetrad system with tetrad number η for each $\eta \notin \{1, 7/4\}$, while for each $\eta \in \{1, 7/4\}$ there are exactly two tetrad systems with tetrad number η, with $\eta(d_{10}) = \eta(e_7) = 1$ and $\eta(d_{16}) = \eta(e_8) = 7/4$.

Now, Lemma 8.2 implies that if $C_4 \neq \emptyset$, then either C_4 consists of μ copies of the tetrad system d_{2k} for some μ and $k > 1$ such that $\mu \cdot 2k = 24$, or it has one of the following two forms:

- $\delta_{10} d_{10} + \varepsilon_7 e_7$, with $\varepsilon_7 > 0$ and $10\delta_{10} + 7\varepsilon_7 = 24$, or
- $\delta_{16} d_{16} + \varepsilon_8 e_8$, with $\varepsilon_8 > 0$ and $16\delta_{16} + 8\varepsilon_8 = 24$.

The resulting tetrad systems are precisely the eight nonempty systems listed in (8.1). □

9 Configurations of Extremal Type II Codes

Let C be an extremal Type II code of length $n = 8, 24, 32, 48, 56, 72$, or 96. Set $w_0 = \min(C)$, so that $w_0 = 4, 8, 8, 12, 12, 16$, or 20 respectively. We prove that C is generated by C_{w_0}. Our approach uses the harmonic weight enumerator machinery developed in Sect. 5, following the approach used for lattices in [Ven84], [Oze86a], [Oze86b], and [Kom09a].

First, we present a few preliminaries. For any $\dot{v} \in \mathbb{F}_2^n$ and any j ($0 \leq j \leq n$), we denote by $N_j(C; \dot{v})$ the value

$$N_j(C; \dot{v}) := |\{c \in C_{w_0} : \mathrm{wt}(c \cap \dot{v}) = j\}| \,. \tag{9.1}$$

Lemma 9.1. *If $\dot{c}$ is a minimal-weight representative of the class $[\dot{c}] \in C/\mathcal{C}_{w_0}(C)$ and $c \in C_{w_0}$, we have the inequality*

$$\mathrm{wt}(c \cap \dot{c}) \leq \frac{w_0}{2}.$$

Proof. This follows quickly, because if $\mathrm{wt}(c \cap \dot{c}) > w_0/2$, then $[\dot{c}]$ contains a codeword $c + \dot{c}$ of weight

$$\mathrm{wt}(c + \dot{c}) = \mathrm{wt}(c) + \mathrm{wt}(\dot{c}) - 2\,\mathrm{wt}(c \cap \dot{c}) < \mathrm{wt}(\dot{c}).$$

This contradicts the minimality of $\dot{c}$ in $[\dot{c}]$. □

We now prove our configuration result for Type II codes of lengths $n = 48$ and 72. The corresponding results for the remaining values of n are presented in [EK12] and [Kom09b].

Theorem 9.2. *If C is an extremal Type II code of length $n = 48$ or 72, then*

$$C = \mathcal{C}_{w_0}(C).$$

Proof. We consider the equivalence classes of $C/\mathcal{C}_{w_0}(C)$ and assume for the sake of contradiction that there is some class $[\dot{c}] \in C/\mathcal{C}_{w_0}(C)$ with minimal-weight representative $\dot{c}$ for which $\mathrm{wt}(\dot{c}) = s > w_0$.

As C is self-dual, we have $N_j(C;c) = 0$ for all odd j. Additionally, by Lemma 9.1, we must have $N_{2j'}(C;\dot{c}) = 0$ for $j' > w_0/4$. We now develop a system of equations in the

$$\frac{w_0}{4} + 1$$

variables $N_0(C;\dot{c}), N_2(C;\dot{c}), \ldots, N_{w_0/2}(C;\dot{c})$. One such equation is

$$N_0(C;\dot{c}) + N_2(C;\dot{c}) + \cdots + N_{w_0/2}(C;\dot{c}) = |C_{w_0}|; \tag{9.2}$$

Corollary 7.6 with $\dot{v} = \dot{c}$ yields $\mathrm{t}(n) + 1$ more. This yields a system of

$$\mathrm{t}(n) + 2 > \frac{w_0}{4} + 1$$

equations in the variables $N_{2j'}(C;\dot{c})$ $(0 \leq j' \leq w_0/4)$.

For $n = 48, 72$, the (extended) determinants of these inhomogeneous systems are

$$2^{26}3^5 5^2 7^1 11^2 23^2 43^1 47^1 \left(\frac{11s^3 - 396s^2 + 4906s - 20736}{(s-3)(s-2)^2(s-1)^3 s^3} \right), \tag{9.3}$$

$$2^{42}3^5 5^2 7^2 11^2 13^1 17^3 23^2 67^2 71^1 \left(\frac{39s^4 - 2600s^3 + 67410s^2 - 800440s + 3650496}{(s-4)(s-3)^2(s-2)^3(s-1)^4 s^4} \right), \tag{9.4}$$

respectively[18]; these determinants must vanish, as they are derived from overdetermined systems. Since equations (9.3) and (9.4) have no integer roots s, we have reached a contradiction. □

Acknowledgements The authors thank Zachary Abel, Henry Cohn, John H. Conway, John F. Duncan, Benedict H. Gross, Abhinav Kumar, Barry Mazur, Gabriele Nebe, Ken Ono, Vera Pless, Eric M. Rains, Shrenik Shah, and the anonymous referee for helpful comments and suggestions.

[18] These determinants were computed using the formula of Proposition 6.2. We omit the equations obtained from the zonal spherical harmonic polynomials of the largest degrees when there are more than $\frac{w_0}{4} + 2$ equations obtained by this method.

During parts of this research, Elkies was supported by NSF grants DMS-0501029 and DMS-1100511, and Kominers was supported by the Harvard College Program for Research in Science and Engineering (PRISE), a Harvard Mathematics Department Highbridge Fellowship, an NSF Graduate Research Fellowship, a Yahoo! Key Scientific Challenges Program Fellowship, a Terence M. Considine Fellowship funded by the John M. Olin Center, and an AMS-Simons Travel Grant.

This work includes a part of the second author's undergraduate thesis [Kom09b].

Appendix A. Proof of Gleason's Theorems for Binary Codes

Let $G_{\rm I}$ be the subgroup of $\mathrm{GL}_2(\mathbb{C})$ generated by $\left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right)$ and $2^{-1/2}\left(\begin{smallmatrix}1&1\\1&-1\end{smallmatrix}\right)$, and let $G_{\rm II}$ be the subgroup of $\mathrm{GL}_2(\mathbb{C})$ generated by $\left(\begin{smallmatrix}1&0\\0&i\end{smallmatrix}\right)$ and $2^{-1/2}\left(\begin{smallmatrix}1&1\\1&-1\end{smallmatrix}\right)$. We have seen, using (3.7) for the second generator, that if C is a binary code of Type I (respectively Type II) then its weight enumerator W_C is contained in the subring of $\mathbb{C}[x,y]$ invariant under linear substitutions with matrices in $G_{\rm I}$ (resp. $G_{\rm II}$). Here we show that the $G_{\rm I}$ invariants are generated by x^2+y^2 and $\delta_8 := x^2y^2(x^2-y^2)^2$, and the $G_{\rm II}$ invariants are generated by $\varphi_8 = W_{e_8}(x,y) = x^8+14x^4y^4+y^8$ and $\xi_{24} = x^4y^4(x^4-y^4)^4$. Note that these are consistent with $G_{\rm I} \subset G_{\rm II}$ because $\varphi_8 = (x^2+y^2)^4 - 4\delta_8$.

We first show that $G_{\rm I}$, and thus also $G_{\rm II}$, contains the signed permutation subgroup of $\mathrm{GL}_2(\mathbb{C})$, which is isomorphic with the eight-element dihedral group and is generated by $\left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)$. Indeed[19] $G_{\rm I} \ni \left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right) = \left(\begin{smallmatrix}1&0\\0&i\end{smallmatrix}\right)^2$, and we calculate that $\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)$ is the conjugate of $\left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right)$ by $2^{-1/2}\left(\begin{smallmatrix}1&1\\1&-1\end{smallmatrix}\right)$. Clearly a polynomial in x,y is invariant under the four matrices $\left(\begin{smallmatrix}\pm1&0\\0&\pm1\end{smallmatrix}\right)$ if and only if it is a polynomial in x^2 and y^2. To be invariant also under $\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)$ it must be a symmetric polynomial in x^2 and y^2. Thus the invariants under this dihedral group are the polynomials in x^2+y^2 and x^2y^2.

We can already find the $G_{\rm I}$-subring. Since the involution $2^{-1/2}\left(\begin{smallmatrix}1&1\\1&-1\end{smallmatrix}\right)$ fixes x^2+y^2 and takes x^2y^2 to $(x^2-y^2)^2/4$, it follows that the weight enumerator of a Type I code is a polynomial in x^2+y^2, $x^2y^2+(x^2-y^2)^2/4$, and $x^2y^2(x^2-y^2)^2/4$. Using the identity $x^2y^2+(x^2-y^2)^2/4 = (x^2+y^2)^2/4$, we dispense with the second of those three generators, and recover Gleason's theorem for self-dual binary codes C (whether of Type I or Type II): the weight enumerator of such a code is a polynomial in x^2+y^2 and δ_8.

To find instead the $G_{\rm II}$ invariants, we next adjoin the matrix $i\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$. We first show that this matrix is contained in the scalar subgroup of $G_{\rm II}$. We claim that the scalars in $G_{\rm I}$ are the 8-th roots of unity. Any scalar matrix $\mu\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$ has determinant

[19] In the coding context we could also show directly that W_C is invariant under $\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)$, that is, that $W_C(x,y) = W_C(y,x)$. Any binary linear code C satisfies $W_C(x,y) = W_C(y,x)$ if and only if C contains the all-1s vector $\mathbf{1}$: in the forward direction, the number of weight n codewords is $W_C(0,1)$, while $W_C(1,0) = 1$ always; in the reverse direction, translation by $\mathbf{1}$ gives for each w a bijection between the codewords of weight w and the codewords of weight $n-w$. But we noted already that a self-dual code, whether of Type I or Type II, contains $\mathbf{1}$.

μ^2, and our generators of $G_{\rm II}$ have determinants i and -1, so $\mu\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right) \in G_{\rm II}$ implies $\mu^8 = 1$. All such μ appear because $G_{\rm II}$ contains

$$\left(2^{-1/2}\left(\begin{smallmatrix}1&1\\1&-1\end{smallmatrix}\right)\left(\begin{smallmatrix}1&0\\0&i\end{smallmatrix}\right)\right)^3 = 2^{-3/2}(2+2i)\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right) = e^{\pi i/4}\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right). \tag{A.1}$$

(The invariance of W_C under $e^{\pi i/4}\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$ already shows that $8 \mid n$.) In particular $G_{\rm II}$ contains $i\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$. This transformation fixes x^2y^2 and takes x^2+y^2 to $-(x^2+y^2)$. Hence the polynomials invariant under the signed permutation group and $i\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$ are precisely the polynomials in x^2y^2 and $(x^2+y^2)^2$.

Let $Q_1 = (x^2+y^2)^2$, $Q_2 = -4x^2y^2$, and $Q_3 = -(Q_1+Q_2) = -(x^2-y^2)^2$. We next find elements of $G_{\rm II}$ that permute the Q_j. One is $\varsigma := e^{-3\pi i/4}\cdot 2^{-1/2}\left(\begin{smallmatrix}1&1\\1&-1\end{smallmatrix}\right)\cdot\left(\begin{smallmatrix}1&0\\0&i\end{smallmatrix}\right)$, which is a 3-cycle contained in $G_{\rm II}$ by (A.1). We calculate that ς permutes the Q_j cyclically. The other is the diagonal matrix $e^{\pi i/4}\left(\begin{smallmatrix}1&0\\0&i\end{smallmatrix}\right)$, which takes Q_2 to itself and Q_1, Q_3 to each other. Thus the subring of $\mathbb{C}[x,y]$ invariant under the subgroup of $G_{\rm II}$ generated by signed permutations, $i\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$, ς, and $e^{\pi i/4}\left(\begin{smallmatrix}1&0\\0&i\end{smallmatrix}\right)$ consists of the polynomials in Q_1, Q_2, Q_3 invariant under arbitrary permutations. Since the three Q_j are independent but for the relation $Q_1+Q_2+Q_3 = 0$, the invariant subring is generated by their elementary symmetric functions of degrees 2 and 3. We calculate that these are

$$Q_1Q_2 + Q_3Q_1 + Q_2Q_3 = -\varphi_8 \quad \text{and} \quad Q_1Q_2Q_3 = 4\psi_2,$$

where $\psi_2 := x^2y^2(x^4-y^4)^2$ is the degree-12 invariant of (5.13). Thus the invariant subring is $\mathbb{C}[\varphi_8, \psi_2]$. Finally the scalar $e^{\pi i/4}$ fixes φ_8 and takes ψ_2 to $-\psi_2$, so the subring of $\mathbb{C}[\varphi_8, \psi_2]$ invariant under $e^{\pi i/4}$ is $\mathbb{C}[\varphi_8, \psi_2^2]$. Since $\psi_2^2 = \xi_{24}$, this proves that any $G_{\rm II}$-invariant polynomial is contained in is $\mathbb{C}[\varphi_8, \xi_{24}]$.

While we proved only that $\mathbb{C}[\varphi_8, \xi_{24}]$ contains the invariant subring $\mathbb{C}[x,y]^{G_{\rm II}}$, we readily conclude that $\mathbb{C}[\varphi_8, \xi_{24}] = \mathbb{C}[x,y]^{G_{\rm II}}$ by verifying that both φ_8 and ξ_{24} are invariant under $G_{\rm II}$. This can be checked either by direct computation of the action of our generators $2^{-1/2}\left(\begin{smallmatrix}1&1\\1&-1\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}1&0\\0&i\end{smallmatrix}\right)$, or by finding Type II codes C_n of length $n = 8$ and $n = 24$ such that $W_{C_8} = \varphi_8$ and $W_{C_{24}} = \varphi_8^3 + \alpha\xi_{24}$ for some $\alpha \neq 0$. We take for C_8 the extended Hamming code, and for C_{24} the extended Golay code or any of the other indecomposable Type II codes of length 24.

References

[AM69] E. F. Assmus and H. F. Mattson, *New 5-designs*, Journal of Combinatorial Theory **6** (1969), 122–151.

[Bac99] C. Bachoc, *On harmonic weight enumerators of binary codes*, Designs, Codes and Cryptography **18** (1999), 11–28.

[Bac01] ———, *Harmonic weight enumerators of non-binary codes and MacWilliams identities*, Codes and Association Schemes (A. Barg and S. Litsyn, eds.), DIMACS Series in Discrete Mathematics and Theoretical Computer Science, vol. 56, American Mathematical Society, 2001, pp. 1–24.

[CD93] A. R. Calderbank and P. Delsarte, *On error-correcting codes and invariant linear forms*, SIAM Journal on Discrete Mathematics **6** (1993), 1–23.

[Con69] J. H. Conway, *A characterization of Leech's lattice*, Inventiones Mathematicæ **7** (1969), 137–142.

[CS99] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*, 3rd ed., Springer-Verlag, 1999.

[CvL91] P. J. Cameron and J. H. van Lint, *Designs, Graphs, Codes and their Links*, Cambridge University Press, 1991 (London Math. Society Student Texts **22**).

[Del78] Ph. Delsarte, *Hahn polynomials, discrete harmonics, and t-designs*, SIAM Journal on Applied Mathematics **34** (1978), 157–166.

[Ebe02] W. Ebeling, *Lattices and codes: A course partially based on lectures by F. Hirzebruch*, 2nd ed., Vieweg, 2002.

[EK10] N. D. Elkies and S. D. Kominers, *On the classification of Type II codes of length* 24, SIAM Journal on Discrete Mathematics **23** (2010), 2173–2177.

[EK12] ______, *Configurations of extremal Type II codes*, in preparation, 2013.

[Elk00] N. D. Elkies, *Lattices, Linear Codes, and Invariants I, II*, Notices of the American Mathematical Society **47** (2000), 1238–1245 and 1382–1391.

[Elk12] ______, *On the quotient of an extremal Type II lattice of rank* 40*,* 80*, or* 120 *by the span of its minimal vectors*, in preparation, 2013.

[Gle71] A. M. Gleason, *Weight polynomials of self-dual codes and the MacWilliams identities*, Actes, Congrés International de Mathématiques (Nice, 1970), vol. 3, Gauthiers-Villars, 1971, pp. 211–215.

[KAL06] L. F. Klosinski, G. L. Alexanderson, and L. C. Larson, *The Sixty-Sixth William Lowell Putnam Mathematical Competition*, American Mathematical Monthly **113** (2006), 733–743.

[Kin03] O. D. King, *A mass formula for unimodular lattices with no roots*, Mathematics of Computation **72** (2003), 839–863.

[Koc87] H. Koch, *Unimodular lattices and self-dual codes*, Proceedings of the International Congress of Mathematicians (Berkeley, Calif., 1986), American Mathematical Society, 1987, pp. 457–465.

[Kom09a] S. D. Kominers, *Configurations of extremal even unimodular lattices*, International Journal of Number Theory **5** (2009), 457–464.

[Kom09b] ______, *Weighted generating functions and configuration results for Type II lattices and codes*, Undergraduate Thesis, Harvard University, 2009, http://www.scottkom.com/articles/kominers_thesis.pdf.

[Kör90] T. W. Körner, *Fourier Analysis*, Cambridge University Press, 1990.

[Lan75] S. Lang, $\mathrm{SL}_2(\mathbb{R})$, Addison-Wesley, 1975.

[LS71] J. Leech and N. J. A. Sloane, *Sphere packing and error-correcting codes*, Canadian Journal of Mathematics **23** (1971), 718–745.

[Mac63] F. J. MacWilliams, *A theorem on the distribution of weights in a systematic code*, Bell System Technical Journal **42** (1963), 79–84.

[MOS75] C. L. Mallows, A. M. Odlyzko, and N. J. A. Sloane, *Upper bounds for modular forms, lattices and codes*, Journal of Algebra **36** (1975), 68–76.

[MS73] C. L. Mallows and N. J. A. Sloane, *An upper bound for self-dual codes*, Information and Control **22** (1973), 188–200.

[MS83] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, 3rd ed., North-Holland Mathematical Library, vol. 16, North-Holland, 1983.

[Nie73] H.-V. Niemeier, *Definite quadratische Formen der Dimension* 24 *und Diskriminante* 1, Journal of Number Theory **5** (1973), 142–178 (German).

[Ott99] U. Ott, *Local weight enumerators for binary self-dual codes*, Journal of Combinatorial Theory, Series A **86** (1999), 362–381.

[Oze86a] M. Ozeki, *On even unimodular positive definite quadratic lattices of rank* 32, Mathematische Zeitschrift **191** (1986), 283–291.

[Oze86b] ———, *On the configurations of even unimodular lattices of rank* 48, Archiv der Mathematik **46** (1986), 54–61.

[PS75] V. Pless and N. J. A. Sloane, *On the classification and enumeration of self-dual codes*, Journal of Combinatorial Theory, Series A **18** (1975), 313–335.

[Rud76] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed., McGraw-Hill, 1976.

[Ser73] J.-P. Serre, *A Course in Arithmetic*, Springer-Verlag, 1973.

[Ser87] ———, *Complex Semisimple Lie Algebras*, Springer-Verlag, 1987.

[Sie69] C. L. Siegel, *Berechnung von Zetafunktionen an ganzzahligen Stellen*, Nachrichten der Akademie der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse, II **1969** (1969), 87–102 [pp. 82–97 in *Gesammelte Abhandlungen IV*, Berlin: Springer 1979].

[Slo77] N. J. A. Sloane, *Error-Correcting Codes and Invariant Theory: New Applications of a Nineteenth-Century Technique*, American Mathematical Monthly **84** (1977), 82–107.

[ST54] G. C. Shephard and J. A. Todd, *Finite unitary reflection groups*, Canadian Journal of Mathematics **6** (1954), 274–304.

[Ven80] B. B. Venkov, *On the classification of integral even unimodular* 24*-dimensional quadratic forms*, Proceedings of the Steklov Institute of Mathematics **148** (1980), 63–74, $\cong$ [CS99, Chapter 18].

[Ven84] ———, *Even unimodular Euclidean lattices in dimension* 32, Journal of Mathematical Sciences **26** (1984), 1860–1867.

[Ven01] ———, *Réseaux et designs sphériques*, Réseaux Euclidiens, Designs Sphériques et Formes Modulaires, Monographies de L'Enseignement Mathématique, vol. 37, Enseignement Mathematique, Genève, 2001, pp. 10–86 (French).

Quadratic Forms and Automorphic Forms

Jonathan Hanke

Abstract These notes give a friendly four-part introduction to various aspects of the arithmetic and analytic theories of quadratic forms, aimed at a graduate-level audience. The main themes discussed are: geometry and local-global methods, theta functions and Siegel's theorem, Clifford algebras and spin groups, and adelic theta liftings via the Weil representation.

Key words Quadratic forms • Automorphic forms • Local densities • Mass formulas • Theta series • Neighbor method • Clifford algebra • Spin group • Theta lifting

MSC Classification(s): 11E08,11E12,11E41,11E45,11E88,11F12,11F27.

Introduction

These notes are an extension of the rough notes provided for my four lecture graduate level course on "Quadratic Forms and Automorphic Forms" at the March 2009 Arizona Winter School on Quadratic Forms. They are meant to give a survey of some aspects of the classical theory of quadratic forms over number fields and their rings of integers (e.g. over $\mathbb{Q}$ and $\mathbb{Z}$), and their connection with modular and automorphic forms.

Originally I had hoped to expand these notes to include many other interesting topics related to Clifford algebras and various automorphic "liftings" that are a natural outgrowth of this chain of ideas. Due to practical time deadlines these notes

J. Hanke (✉)
One PalmerSquare, Suite 441, Princeton, NJ 08542, USA
e-mail: jonhanke@gmail.com

K. Alladi et al. (eds.), *Quadratic and Higher Degree Forms*, Developments in Mathematics 31, DOI 10.1007/978-1-4614-7488-3_5,

are essentially a written version of the talks which have been filled out to include precise references for all theorems and details of less well-known arguments with the hope of enabling an eager graduate student to gain a working knowledge of the basic ideas and arguments for each of the topics covered.

I would like to thank the organizers of the 2009 Arizona Winter School (David Savitt, Fernando Rodriguez-Villegas, Matt Papanikolas, William Stein, and Dinesh Thakur) for the opportunity to give these lectures, as well as all of the students who worked very hard in the evenings to make progress on the research projects associated to these lectures. Special thanks go to John Voight for his many hours helping these students, and also for all of the work he put into helping to write up notes for Professor Conway's lectures. Also several helpful comments made by Pete L. Clark, Danny Krashen, Rishikesh, Robert Varley and the anonymous referee when proofreading of these notes. Finally, a special thank you to Krishna Alladi for his multi-year efforts to organize this volume and bring it to publication in its present form.

I am grateful to MSRI for their hospitality and 24-hour library access during the Spring 2011 semester, during which the final version of these notes were written and many of the references were added. I also acknowledge the support of NSF Grant DMS-0903401 during the Winter School and while these notes were being written.

For a more lively (but perhaps less precise) introduction to this material, the reader is encouraged to view videos of the lectures online at the Arizona Winter School webpage

http://swc.math.arizona.edu/aws/09/

Any further information relating to the published version of these notes will be posted on my website

http://jonhanke.com

Dedication

I would like to extend a special thanks to my advisor Goro Shimura, without whom I would not have become involved with this beautiful subject, and these notes would have not been possible. His dedication to careful exposition and referencing have been a major influence on these notes, and hopefully this attention to detail will make it easier for the reader seeking to learn this material. Having said this, despite my best efforts I am sure that these notes contain mistakes, and any corrections are very welcome. Please feel free to send them to jonhanke@gmail.com.

I dedicate these notes to the many excellent expositors whose efforts have helped me to learn new areas of mathematics, including (but not limited to) Tony Knapp, Steve Gelbart, Bill Duke, Henryk Iwaniec, and Serge Lang.

1 Background on Quadratic Forms

1.1 Notation and Conventions

We let $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ denote the usual integers, rational numbers, real numbers, and complex numbers, and also denote the natural numbers as $\mathbb{N} := \mathbb{Z}_{>0} := \{1, 2, \cdots\}$. We say that an $n \times n$ matrix $A = (a_{ij}) \in M_n(R)$ over a ring R is **even** if $a_{ii} \in 2R$, and **symmetric** if $a_{ij} = a_{ji}$ for all $1 \leq i, j \leq n$. The symmetric matrices in $M_n(R)$ are denoted by $\mathrm{Sym}_n(R)$. We denote the trivial (mod 1) Dirichlet character sending all integers to 1 by $\mathbf{1}$. In analytic estimates, it is common to use the notation $X \gg Y$ to mean that $X > C \cdot Y$ for some (implied) constant $C \in \mathbb{R} > 0$.

Suppose R is an integral domain and V is a (finite dimensional) vectorspace over its field of fractions F. By a **lattice** or **R-lattice** in V we will mean a finitely generated R-module over R that spans V. In particular, notice that we will always assume that our lattices have full rank in V.

If F is a number field (i.e. a finite field extension of $\mathbb{Q}$), then we define a **place** or **normalized valuation** of F to be an equivalence class of metrics $|\cdot|_v$ on F that induce the same topology on F. We implicitly identify v with the distinguished metric in each class agreeing with the usual absolute value when $F_v = \mathbb{R}$ or $\mathbb{C}$, and giving $|\mathfrak{p}|_v = |\mathcal{O}_v/\mathfrak{p}\mathcal{O}_v|$ when v is non-archimedean and F_v has valuation ring $\mathcal{O}_v$ with maximal ideal $\mathfrak{p}$. If $p \in \mathbb{N}$ with $(p) = \mathfrak{p} \cap \mathbb{Z}$ then we have an associated **valuation** on $F_v^\times$ given by $\mathrm{ord}_v(\cdot) := -\log_p(\cdot) = \mathrm{ord}_\mathfrak{p}(\cdot)$. We define a **(non-zero) squareclass** of a field K to be an element of $K^\times/(K^\times)^2$, and when K is a non-archimedean local field then the valuation ord_v descends to give a $(\mathbb{Z}/2\mathbb{Z})$-valued valuation on squareclasses.

We say that an R-valued quadratic form over a (commutative) ring R is **primitive** if the ideal generated by its values $(Q(R))$ is R and we say that Q **represents** m if $m \in Q(R)$. Given a quadratic space (V, Q) of dimension n over a field F, we define the orthogonal group $O(V) := O_Q(V)$ to be the set of invertible linear transformations $L : V \to V$ so that $Q(L(\vec{v})) = Q(\vec{v})$ for all $\vec{v} \in V$. Given a basis of V, $O(V)$ can be realized as a subset of $\mathrm{GL}_n(F)$. We also define the **special orthogonal group** $SO(V) := O^+(V)$ as the (orientation-preserving) subgroup of $O(V)$ having determinant 1.

In these notes we will study aspects of the theory of quadratic forms over rings R and fields F of characteristic $\mathrm{char}(\cdot) \neq 2$ (i.e. where $1 + 1 \neq 0$). While one can discuss quadratic forms in characteristic 2, we can no longer equate them with symmetric bilinear forms, so the theory there is more complicated. For references valid in characteristic 2, we refer the reader to [Kap03, Knu91, EKM08, Bak81, Sah60].

Our main interest is in quadratic forms over the field $\mathbb{Q}$, the ring $\mathbb{Z}$, and their completions, though we may consider a more general setting (e.g. a number field and its ring of integers) when there are no additional complications in doing so.

1.2 Definitions of Quadratic Forms

In this section we give some basic definitions and ideas used to understand quadratic forms and the numbers they represent. We define a **quadratic form** $Q(\vec{x})$ **over a ring** R to be a degree 2 homogeneous polynomial

$$Q(\vec{x}) := Q(x_1, \cdots, x_n) := \sum_{1 \leq i \leq j \leq n} c_{ij} x_i x_j$$

in n variables with coefficients c_{ij} in R.

When division by 2 is allowed (either in R or in some ring containing R) we can also consider the quadratic form $Q(\vec{x})$ as coming from the symmetric **Gram bilinear form**

$$B(\vec{x}, \vec{y}) := \sum_{1 \leq i,j \leq n} b_{ij} x_i y_j = {}^t\vec{x}\, B\, \vec{y} \tag{1.1}$$

via the formula $Q(\vec{x}) = B(\vec{x}, \vec{x}) = {}^t\vec{x}\, B\, \vec{x}$, where the matrix $B := (b_{ij})$ and $b_{ij} = \frac{1}{2}(c_{ij} + c_{ji})$ (with the convention that $c_{ij} = 0$ if $i > j$). We refer to the symmetric matrix $B = (b_{ij})$ as the **Gram matrix** of Q. It is common to relate a quadratic form $Q(\vec{x})$ to its Gram bilinear form $B(\vec{x}, \vec{y})$ by the **polarization identity**

$$\begin{aligned} Q(\vec{x} + \vec{y}) &= B(\vec{x} + \vec{y}, \vec{x} + \vec{y}) \\ &= B(\vec{x}, \vec{x}) + 2B(\vec{x}, \vec{y}) + B(\vec{y}, \vec{y}) \\ &= Q(\vec{x}) + 2B(\vec{x}, \vec{y}) + Q(\vec{y}). \end{aligned} \tag{1.2}$$

From either of these formulas for $B(\vec{x}, \vec{y})$, we see that the matrix $B \in \frac{1}{2}\mathrm{Sym}_n(R)$.

Since often $\frac{1}{2} \notin R$, it is somewhat unnatural to consider the Gram bilinear form since it is not an object defined over R. However the **Hessian bilinear form** $H(\vec{x}, \vec{y}) := 2B(\vec{x}, \vec{y})$ is defined over R and can be seen to be very naturally associated to the quadratic form $Q(\vec{x})$ by the polarization identity. This definition is motivated by the fact that the matrix $H := 2B \in \mathrm{Sym}_n(R)$ of the Hessian bilinear form is the **Hessian matrix** of second order partial derivatives of $Q(\vec{x})$ (i.e. $H = (H_{ij})$ where $a_{ij} = \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} Q(\vec{x})$). For this reason, it is often preferable to use the Hessian formulation when working over rings, and in particular when $R = \mathbb{Z}$. Notice that the diagonal coefficients h_{ii} of the Hessian matrix are even, so H is actually an even symmetric matrix. In the geometric theory of quadratic forms the Hessian bilinear form is often referred to as the **polar form** of Q because of its close connection with the polarization identity (see [EKM08, p. 39]).

Another perspective on quadratic forms is to think of them as **free quadratic** R**-modules**, by which we mean an R-module $M \cong R^n$ equipped with a "quadratic function"

$$Q : M \to R$$

which is a function satisfying

1. $Q(a\vec{x}) = a^2Q(\vec{x})$ for all $a \in R$, and for which
2. $H(\vec{x}, \vec{y}) := Q(\vec{x} + \vec{y}) - Q(\vec{x}) - Q(\vec{y})$ is a bilinear form.

This perspective is equivalent to that of a quadratic form because one can recover the quadratic form coefficients from $c_{ii} = Q(\vec{e}_i)$, and when $i < j$ we have $c_{ij} = H(\vec{e}_i, \vec{e}_j)$.

In the case when $R = \mathbb{Z}$, this perspective allows us to think of Q as a "quadratic lattice" which naturally sits isometrically in a vector space over $\mathbb{Q}$ that is also equipped with a quadratic form Q. More precisely, we define a **quadratic space** to be a pair (V, Q) where V is a finite-dimensional vector space over F and Q is a quadratic form on V. We say that a module M is a **quadratic lattice** if M is a (full rank) R-lattice in a quadratic space (V, Q) over F, where F is the field of fractions of R. Notice that we can always realize a quadratic form $Q(\vec{x})$ over an integral domain R as being induced from a free quadratic lattice R^n in some quadratic space (V, Q), by thinking of Q as a function on $V = F^n$ where F is the fraction field of R. (*Note: However if R is not a principal ideal domain then there will exist non-free quadratic lattices, which cannot be described as quadratic forms!*)

We say that a quadratic form $Q(\vec{x})$ is R**-valued** if $Q(R^n) \subseteq R$. From the polarization identity we see that $Q(\vec{x})$ is R-valued iff $Q(\vec{x})$ is defined over R (i.e. all coefficients $c_{ij} \in R$) because $c_{ij} = H(\vec{e}_i, \vec{e}_j)$ when $i \neq j$ and $c_{ii} = Q(\vec{e}_i)$.

1.3 Equivalence of Quadratic Forms

Informally, we would like to consider two quadratic forms as being "the same" if we can rewrite one in terms of the other by a change of variables. More precisely, we say that two quadratic forms Q_1 and Q_2 are **equivalent over** R, and write $Q_1 \sim_R Q_2$, if there is an invertible linear change of variables $\phi(\vec{x})$ with coefficients in R so that $Q_2(\vec{x}) = Q_1(\phi(\vec{x}))$. We can represent $\phi(\vec{x})$ (with respect to the generators $\vec{e}_i$ of the free module R^n) as left-multiplication by an invertible matrix M over R – i.e. $M \in M_n(R)$ and

$$\phi(\vec{x}) = M\vec{x}.$$

Expressing this equivalence in terms of the associated Hessian and Gram matrices, we see that composition with ϕ gives the equivalence relations

$$H_1 \sim_R {}^tMH_1M = H_2 \qquad \text{and} \qquad B_1 \sim_R {}^tMB_1M = B_2 \tag{1.3}$$

where

$$Q_i(\vec{x}) = \tfrac{1}{2}{}^t\vec{x}H_i\vec{x} = {}^t\vec{x}B_i\vec{x}.$$

For quadratic lattices, the corresponding notion is to say that two quadratic R-modules are equivalent if there is a R-module isomorphism between the modules commuting with the quadratic functions (i.e. preserving all values of the quadratic function). Because this R-module isomorphism preserves all values of the quadratic function, it is often referred to as an **isometry**. We will see later that this idea of *thinking of equivalent quadratic lattices as isometric lattices in a quadratic space is very fruitful*, and can be used to give a very geometric flavor to questions about the arithmetic of quadratic forms.

1.4 Direct Sums and Scaling

Two important constructions for making new quadratic forms from known ones are the operations of scaling and direct sum. Given $a \in R$ and a quadratic form $Q(\vec{x})$ defined over R, we can define a new **scaled quadratic form** $a \cdot Q(\vec{x})$ which is also defined over R. We can also try to detect if a quadratic form is a scaled version of some other quadratic form by looking at its values, generated either as a bilinear form or as a quadratic form. We therefore define the **(Hessian and Gram) scale** and **norm** of Q by

$$\mathrm{Scale}_H(Q) := \{H(\vec{x},\vec{y}) \mid \vec{x},\vec{y} \in R^n\} \tag{1.4}$$

$$\mathrm{Scale}_G(Q) := \{B(\vec{x},\vec{y}) \mid \vec{x},\vec{y} \in R^n\} \tag{1.5}$$

$$\mathrm{Norm}(Q) := \{Q(\vec{x}) \mid \vec{x} \in R^n\} \tag{1.6}$$

and notice that $\mathrm{Scale}_{H/G}(a \cdot Q) = a \cdot \mathrm{Scale}_{H/G}(Q)$ and $\mathrm{Norm}(a \cdot Q) = a^2 \cdot \mathrm{Norm}(Q)$.

Another useful construction for making new quadratic forms is to take the direct sum of two given quadratic forms. Given $Q_1(\vec{x}_1)$ and $Q_2(\vec{x}_2)$ in n_1 and n_2 distinct variables over R respectively, we define their **(orthogonal) direct sum** $Q_1 \oplus Q_2$ as the quadratic form

$$(Q_1 \oplus Q_2)(\vec{x}) := Q_1(\vec{x}_1) + Q_2(\vec{x}_2)$$

in $n_1 + n_2$ variables, where $\vec{x} := (\vec{x}_1, \vec{x}_2)$.

1.5 The Geometry of Quadratic Spaces

Quadratic forms become much simpler to study when the base ring R is a field (which we denote by F). In that case we know that all finite-dimensional R-modules are free (i.e. every finite-dimensional vector space has a basis), and that there is exactly one of each dimension n. This simplification allows us to replace commutative algebra with linear algebra when studying quadratic forms, and motivates our previous definition of a **quadratic space** as pair (V, Q) consisting of a quadratic form Q on a finite-dimensional vector space V over F. We will often refer to any quadratic form (in n variables) over a field as a quadratic space, with the implicit understanding that we are considering the vector space $V := F^n$. We also often refer to an equivalence of quadratic spaces (over F) as an **isometry**.

We now present some very useful geometric classification theorems about quadratic spaces. To do this, we define the **(Gram) inner product** of $\vec{v}_1, \vec{v}_2 \in V$ as the value of the Gram symmetric bilinear form $B(\vec{v}_1, \vec{v}_2)$.[1] We say that two vectors $\vec{v}_1, \vec{v}_2 \in V$ are **perpendicular** or **orthogonal** and write $\vec{v}_1 \perp \vec{v}_2$ if their inner product $B(\vec{v}_1, \vec{v}_2) = 0$. Similarly we say that a vector $\vec{v}$ is **perpendicular to a subspace** $W \subseteq V$ if $\vec{v} \perp \vec{w}$ for all $\vec{w} \in W$, and we say that two subspaces $W_1, W_2 \subseteq V$ are perpendicular if $\vec{w}_1 \perp \vec{w}_2$ for all $\vec{w}_i \in W_i$.

Our first theorem says we can always find an orthogonal basis for V, which we can see puts the (Gram/Hessian) matrices associated to Q in diagonal form:

Theorem 1.5.1 (Orthogonal splitting/diagonalization). *Every quadratic space V admits an orthogonal basis.*

Proof. This is proved in Cassels's book [Cas78, Lemma 1.4 on pp. 13–14], where he refers to this as a "normal basis". □

Given a quadratic space (V, Q), to any a choice of basis $\mathcal{B} := \{\vec{v}_1, \dots, \vec{v}_n\}$ for V we can associate a quadratic form $Q_{\mathcal{B}}(\vec{x})$ by expressing elements of V in the coordinates of $\mathcal{B}$:

$$\mathcal{B} \quad \longmapsto \quad Q_{\mathcal{B}}(\vec{x}) := Q(x_1\vec{v}_1 + \cdots x_n\vec{v}_n)$$

We can use this association to define the **determinant** $\det(Q)$ of (V, Q) as the determinant $\det(B)$ of the Gram matrix B of the associated quadratic form $Q_{\mathcal{B}}(\vec{x})$ for some basis $\mathcal{B}$. However since changing the basis $\mathcal{B}$ induces the equivalence relation (1.3), we see that $\det(Q)$ is only well-defined up to multiplication by $\det(M)^2 \in K^\times$, and so $\det(Q)$ gives a well-defined square-class in $K/(K^\times)^2$. We say that Q is **degenerate** if $\det(Q) = 0$, otherwise we say that Q is **non-degenerate** (or **regular**). By convention, the zero-dimensional quadratic space has $\det(Q) = 1$ and is non-degenerate.

Lemma 1.5.2. *If a quadratic space (V, Q) is degenerate, then there is some non-zero vector $\vec{v} \in V$ perpendicular to V (i.e., $\vec{v} \perp V$).*

The next theorem states that we can always reduce a degenerate quadratic space to a non-degenerate space by (orthogonally) splitting off a **zero space** (V, Q), which we define to be a vector space V equipped with the identically zero quadratic form $Q(\vec{x}) = 0$. Notice that a zero space has an inner product that is identically zero, so it is always perpendicular to itself. (In the literature a zero space is often referred to as a **totally isotropic space**.) We define **the radical of a quadratic space** (V, Q) as the maximal (quadratic) subspace of V perpendicular to all $\vec{v} \in V$, which is just the set of vectors perpendicular to all of V.

[1] We could also have defined an inner product using the Hessian bilinear form, but this choice is less standard in the literature and the difference will not matter for our purposes in this section.

Lemma 1.5.3 (Radical Splitting). *Every quadratic space can be written as a orthogonal direct sum of a zero space (the radical of the quadratic space) and a non-degenerate space.*

Proof. This is given in Cassels's book [Cas78, Lemma 6.1 on p. 28]. □

We say that a non-zero vector $\vec{v} \in V$ is **isotropic** if $Q(\vec{v}) = 0$ and say that $\vec{v}$ is **anisotropic** otherwise. Extending this definition to subspaces, we say that a subspace $U \subseteq (V, Q)$ is isotropic if it contains an isotropic vector, and anisotropic otherwise. Notice that U is a totally isotropic subspace $\iff$ every non-zero vector in U is isotropic.

For non-degenerate quadratic spaces, isotropic vectors play a key role because of their close relation to the **hyperbolic plane** H_2, which is defined as the two-dimensional quadratic space (say with coordinates x and y) endowed with the quadratic form $Q(\vec{x}) = Q(x, y) = xy$. We also refer to the orthogonal direct sum of r hyperbolic planes as the **hyperbolic space** H_{2r}, which has dimension $2r$.

Theorem 1.5.4 (Totally Isotropic Splitting). *Suppose (V, Q) is a non-degenerate quadratic space. Then for every r-dimensional zero subspace $U \subseteq V$ we can find a complementary r-dimensional subspace $U' \subseteq V$ so that $V = U \oplus U' \oplus W$ as vector spaces, and (as quadratic subspaces) W is non-degenerate and $U \oplus U'$ is equivalent to the hyperbolic space H_{2r}.*

Proof. This is shown in Lam's Book [Lam05, Theorem 3.4(1–2), p. 10], or by repeated application of [Cas78, Corollary 1, p. 15]. □

In the case that W is isotropic, we can repeatedly apply this to split off additional hyperbolic spaces until W is anisotropic. So we could have initially taken U to be a totally isotropic subspace of maximal dimension in (V, Q), called a **maximal totally isotropic subspace**, and then concluded that W was anisotropic.

Another particularly useful result about quadratic spaces is that there is a large group of F-linear isometries of (V, Q), called the **orthogonal group** and denoted as $O_Q(V)$ or $O(V)$, that acts on (V, Q). We will see throughout these lectures that the orthogonal group is very closely connected to the arithmetic of quadratic forms, partly because of the following important structural theorem of Witt which classifies isometric quadratic subspaces within a given quadratic space in terms of the orbits of $O(V)$.

Theorem 1.5.5 (Witt's Theorem). *Suppose that U and U' are non-degenerate isometric (quadratic) subspaces of a quadratic space V. Then any isometry $\alpha : U \to U'$ extends to an isometry $\alpha : V \to V$.*

Proof. This is proved in almost every quadratic forms book, e.g. Cassels's book [Cas78, Theorem 4.1 on p. 21], Shimura's book [Shi10, Theorem 22.2 on pp. 116–117], and Lam's book [Lam05, Theorems 4.2 and 4.7 on pp. 12–15]. □

Notice that Witt's Theorem shows that the dimension of a maximal isotropic subspace of a quadratic space (V, Q) is a well-defined number (independent of the particular maximal isotropic subspace of V that we choose), since otherwise we could find an isometry of V that puts the smaller subspace properly inside the larger one, violating the assumption of maximality.

1.6 Quadratic Forms over Local Fields

We now suppose that (V, Q) is a non-degenerate quadratic space over one of the local fields $F = \mathbb{R}$, $\mathbb{C}$ or $\mathbb{Q}_p$ where p is a positive prime number. In this setting we can successfully classify quadratic spaces in terms of certain invariants associated to them. The major result along these lines is that in addition to the dimension and determinant, at most one additional invariant is needed to classify non-degenerate quadratic spaces up to equivalence.

Theorem 1.6.1. *There is exactly one non-degenerate quadratic space over $\mathbb{C}$ of each dimension n.*

Proof. From the orthogonal splitting Theorem 1.5.1 we see that any such $Q(\vec{x}) \sim_{\mathbb{C}} \sum_{i=1}^{n} c_i x_i^2$ with $c_i \in \mathbb{C}^\times$. However since $\mathbb{C}^\times = (\mathbb{C}^\times)^2$ every c_i can be written as some a_i^2, so we see that $Q(\vec{x}) = \sum_{i=1}^{n} (a_i x_i)^2 \sim_{\mathbb{C}} \sum_{i=1}^{n} x_i^2$. □

Theorem 1.6.2. *The non-degenerate quadratic spaces over $\mathbb{R}$ are in $1-1$ correspondence with the pairs (n, p_1) where $0 \leq p_1 \leq n$.*

Proof. Since $\mathbb{R}^\times$ has two squareclasses $\pm(\mathbb{R}^\times)^2$, we see that the diagonal elements can be chosen to be either 1 or -1. Since the dimension of a maximal totally isotropic subspace is a well-defined isometry invariant, this characterizes the number of $(1, -1)$ pairs on the diagonal, and then the remaining diagonal entries all have the same sign. Its orthogonal complement is anisotropic, and is either 1_{n-2r} or -1_{n-2r} depending on the sign of the values it represents. □

In practice it is more standard to use the **signature invariant** $p := p_1 - p_2$ instead of p_1, however they are equivalent for our purpose of giving a complete set of invariants for quadratic spaces over $\mathbb{R}$.

Theorem 1.6.3. *The non-degenerate quadratic spaces over $\mathbb{Q}_p$ are in $1-1$ correspondence with the triples (n, d, c) where $n = \dim(Q) \in \mathbb{Z} \geq 0$, $d = \det(Q) \in \mathbb{Q}_p^\times/(\mathbb{Q}_p^\times)^2$, and $c \in \{\pm 1\}$ is the Hasse invariant of Q, under the restrictions that*

1. $c = 1$ *if either* $n = 1$ *or* $(n, d) = (2, -1)$, *and also*
2. $(n, d, c) = (0, 1, 1)$ *if* $n = 0$.

Proof. This is proved in Cassels's book [Cas78, Theorem 1.1 on p. 55]. □

Another area in which we can extract more information about quadratic forms over local fields is in terms of the **maximal anisotropic dimension of** K, which is defined to be the largest dimension of an anisotropic subspace of any quadratic space over K. This is sometimes called the u**-invariant** of K.

Theorem 1.6.4. *The maximal anisotropic dimensions of* $\mathbb{C}, \mathbb{R}$ *and* $\mathbb{Q}_p$ *are* $1, \infty$, *and* 4.

Proof. Over $\mathbb{C}$ any form in $n \geq 2$ variables is isotropic, and any non-zero form of dimension 1 is anisotropic. Over $\mathbb{R}$ we see that $Q(\vec{x}) = \sum_{i=1}^{n} x_i^2$ is anisotropic for any $n \in \mathbb{N}$. Over $\mathbb{Q}_p$ any form in $\mathfrak{n} \geq 5$ variables isotropic, and there is always an anisotropic form in four variables arising as the norm form from the unique ramified quaternion algebra over $\mathbb{Q}_p$ (see [Lam05, Theorem 2.12 and Corollary 2.11, p. 158]). □

In particular, over $\mathbb{R}$ says that it is possible for quadratic forms of any dimension to be anisotropic, which means that either Q represents only positive or only negative values (using a non-zero vector), and in these cases we say Q is **positive definite** or **negative definite** respectively. When a non-degenerate Q is isotropic over $\mathbb{R}$ then it must represent both positive and negative values, in which case we say that Q is **indefinite**. For $\mathbb{Q}_p$ there is no notion of positive and negative, but one can concretely understand the u-invariant $u(\mathbb{Q}_p) = 4$ from the existence of certain (non-split) quaternion algebras at every prime p, which will be discussed in more detail in Sect. 3. The norm forms of these quaternion algebras assume all values of F and do not represent zero non-trivially.

1.7 The Geometry of Quadratic Lattices: Dual Lattices

Quadratic lattices also have a kind of geometry associated to them that is a little more subtle than the "perpendicular" geometry of subspaces of quadratic spaces. Given a (full rank) quadratic lattice L (over R) in a quadratic space (V, Q) (over the fraction field F of R), we can consider the elements of V where the linear form $H(\cdot, L)$ is in any fixed ideal I of R.[2] When $R = F$ is a field, the only ideals are F and (0); taking $I = F$ imposes no condition, and taking $I = (0)$ recovers the notion of the orthogonal complement $L^{\perp}$ of L. However when $R \neq F$ then taking $I = R$ gives an interesting integral notion of "orthogonality" which is very useful for making other lattices that are closely related to L.

We define the **(Hessian) dual lattice** $L^{\#}$ of L to be the set of vectors in V that have integral inner product with all $\vec{w} \in L$, i.e.

$$L^{\#} := \{\vec{v} \in V \mid H(\vec{v}, \vec{w}) \in R \text{ for all } \vec{w} \in L\}.$$

Notice that if $H(\vec{x}, \vec{y})$ is R-valued for all $\vec{x}, \vec{y} \in L$ then we have $L \subseteq L^{\#}$. When an R-valued lattice L is free as an R-module, we also know that the matrix of H in any

[2] For any subset $S \subseteq V$ the set $H(S, L)$ is an R-module, and so it is natural to consider maximal subsets of V where $H(S, L)$ is a fixed R-module. From the bilinearity of H, we see that these maximal sets S are also R-modules.

basis of L is symmetric with coefficients in R and even diagonal (i.e. all $a_{ii} \in 2R$), so H is an even symmetric matrix. We then define the **level** of L to be the smallest (non-zero) ideal $\mathfrak{n} \subseteq R$ so that the matrices in $\mathfrak{n}H^{-1}$ are also even. The level is a very useful invariant of L which appears when we take dual lattices (because H^{-1} is the matrix of basis vectors for the dual basis of the given basis $\mathcal{B}$ of L in the coordinates of $\mathcal{B}$), and in particular it plays an important role in the theory of theta functions (see Sect. 2).

In the special case where $R = \mathbb{Z}$ the level $\mathfrak{n}$ can be written as $\mathfrak{n} = (N)$ for some $N \in \mathbb{N}$, and this (minimal) N is what is usually referred to as the level of the quadratic lattice (which is also a quadratic form since all $\mathbb{Z}$-lattices are free).

We say that a quadratic form over a ring R is **(Hessian) unimodular** if its Hessian bilinear form has unit determinant (i.e. $\det(H) \in R^{\times}$). In terms of quadratic lattices, this is equivalent to saying that the associated quadratic R-lattice $(L, Q) = (R^n, Q)$ is **(Hessian) self-dual** (i.e. $L^{\#} = L$).

Remark 1.7.1. It is somewhat more customary for authors to define the dual lattice [O'M71, §82F, p. 230] as the **Gram dual lattice**,

$$L_G^{\#} := \{\vec{v} \in V \mid B(\vec{v}, \vec{w}) \in R \text{ for all } \vec{w} \in L\}.$$

and for the analogous notion of unimodular and self-dual to be **Gram unimodular** (i.e. $\det(B) \in R^{\times}$) and **Gram self-dual** (i.e. $L_G^{\#} = L$). While either definition will suffice for a Jordan splitting theorem (Theorem 1.8.2) in terms of unimodular lattices (because Gram and Hessian unimodular lattices are simply scaled versions of each other), the Hessian definitions are more natural from an arithmetic perspective (e.g. in the definition of level, the proof of Theorem 1.8.2, and our discussion of neighboring lattices in Sect. 1.10). If 2 is invertible in R, then there is no distinction between the Hessian and Gram formulations, so over local (p-adic) rings this only makes a difference over the 2-adic integers $\mathbb{Z}_2$.

1.8 Quadratic Forms over Local (p-adic) Rings of Integers

If we consider quadratic forms over the ring of integers $\mathbb{Z}_p$ of the p-adic field $\mathbb{Q}_p$, then the classification theorem is more involved because the valuation and units will both play a role. The main result along these lines involves the notion of a "Jordan splitting", which breaks Q into a sum of pieces scaled by powers of p which are as simple as possible.

Lemma 1.8.1. *Suppose that R is a principal ideal domain, L is a quadratic R-lattice in the quadratic space (V, Q), and W is a non-degenerate subspace of (V, Q). If* $\mathrm{Scale}(L \cap W) = \mathrm{Scale}(L)$*, then we can write $L = (L \cap W) \oplus (L \cap W^{\perp})$.*

Proof. Since R is a principal ideal domain, we know that all finitely generated R-modules are free [Lan95, Theorem 7.1, p. 146], giving $L \cong R^n$ and $L \cap W \cong R^k$ where V and W have dimensions n and k respectively. The structure theorem for finitely-generated R-modules [Lan95, Theorem 7.8(i), p. 153] says that we can find a set of n vectors $\{\vec{w}_1, \dots, \vec{w}_k, \vec{v}_{k+1}, \dots \vec{v}_n\}$ that generate L whose first k elements generate $L \cap W$. Without loss of generality we can scale the quadratic form Q so that $\mathrm{Scale}_H(L) = R$, in which case the Hessian matrix of $L \cap W$ in this basis is in $\mathrm{GL}_n(R)$. Therefore for each generator $\vec{v}_j$ of L, we can realize the vector $(H(\vec{w}_i, \vec{v}_j))_{1 \leq i \leq k}$ as an R-linear combination of its columns, and this linear combination can be used to adjust each $\vec{v}_j$ to be orthogonal to all $\vec{w}_i$, completing the proof. □

Theorem 1.8.2 (Jordan Decomposition). *A non-degenerate quadratic form over $\mathbb{Z}_p$ can be written as a direct sum*

$$Q(\vec{x}) = \bigoplus_{j \in \mathbb{Z}} p^j Q_j(\vec{x}_j)$$

where the $Q_j(\vec{x}_j)$ are unimodular. More explicitly, if $p > 2$ then each Q_j is a direct sum of quadratic forms $u_i x^2$ for some p-adic units $u_i \in \mathbb{Z}_p^\times$, and if $p = 2$ then each Q_j is a direct sum of some collection of the unimodular quadratic forms $u_i x^2$, xy, and $x^2 + xy + y^2$.

Proof. This follows from successively applying Lemma 1.8.1, and noticing that the minimal scale for a sublattice of L can always be attained by a rank 1 sublattice when $p \neq 2$, and be a rank 2 sublattice when $p = 2$. The explicit statement for $p = 2$ follows from checking equivalences between the rank 2 unimodular lattices.

Most authors state this result in terms of Gram unimodular lattices. When $p > 2$ this is given in Cassels's book [Cas78, Lemma 3.4 on p. 115], while $p = 2$ is given in the very explicit form stated here as [Cas78, Lemma 4.1 on p. 117]. See also [Ger08, Theorem 8.1, p. 162 and Theorem 8.9, p. 168] and [O'M71, Theorem 93:29 on p. 277]. For the general classification of integral quadratic forms over number fields at primes over $p = 2$ see O'Meara's book [O'M71, Theorem 93:28 on pp. 267–276], though there only invariants (and not explicit representatives) are given. □

Remark 1.8.3. As a convention, we consider the ring of integers of $\mathbb{R}$ and $\mathbb{C}$ to be just $\mathbb{R}$ and $\mathbb{C}$ again, so there is nothing new to say in that situation.

1.9 Local-Global Results for Quadratic Forms

A useful idea for studying quadratic forms over either $\mathbb{Q}$ or $\mathbb{Z}$ is to consider them locally over all completions (by thinking of their coefficients in the associated local field $\mathbb{Q}_v$ or ring $\mathbb{Z}_v$), and then try to use information about these "local" quadratic forms to answer questions about the original "global" quadratic form. While it is easy to pass from Q to a local quadratic form Q_v defined over its completion at the valuation v, it is more difficult to reverse this process to glue together a set of local forms Q_v for all v to obtain some "global" quadratic form Q.

We now examine the extent to which this can be done. Our first theorem tells us that for quadratic spaces over Q this “local-global” procedure works flawlessly, and we can check the (rational) equivalence of forms using only local information.

Theorem 1.9.1 (Hasse-Minkowski Theorem). *Given two quadratic forms Q_1 and Q_2 defined over $\mathbb{Q}$, we have*

$$Q_1 \sim_{\mathbb{Q}} Q_2 \iff Q_1 \sim_{\mathbb{Q}_v} Q_2 \textit{ for all places } v \textit{ of } \mathbb{Q}.$$

Proof. This is stated as the “Weak Hasse principle” in Cassels’s book [Cas78, Theorem 1.3 on p. 77], but proved in [Cas78, §6.7, pp. 85–86]. □

Remark 1.9.2. The same result holds if we replace $\mathbb{Q}$ with any number field K, and replace the $\mathbb{Q}_v$ with all of the completions K_v at all places of K.

We denote the $\mathbb{Z}$-equivalence class of Q by $\mathrm{Cls}(Q)$, and refer to it as the **class** of Q. Given two quadratic forms Q_1 and Q_2 over $\mathbb{Z}$, we always have that

$$Q_1 \sim_{\mathbb{Z}} Q_2 \implies Q_1 \sim_{\mathbb{Z}_v} Q_2 \quad \text{for all places } v$$

since the linear transformation giving the $\mathbb{Z}$-equivalence is also defined over each completion $\mathbb{Z}_v$. (Recall that $\mathbb{Z}_\infty := \mathbb{R}$ by convention.) However unlike with quadratic forms over $\mathbb{Q}$, we are not guaranteed that local equivalence over all $\mathbb{Z}_v$ will ensure equivalence over $\mathbb{Z}$. The number of distinct $\mathbb{Z}$-equivalence classes of quadratic forms that are locally $\mathbb{Z}_v$-equivalent to Q at all places is called the **class number** h_Q of the quadratic form Q, and the set of all forms with the same localization as Q is called the **genus of** Q, so $h_Q = |\mathrm{Gen}(Q)|$.

It is a major result of Siegel from the reduction theory of (either definite or indefinite) quadratic forms over $\mathbb{R}$ that $h_Q < \infty$. The class number of an indefinite quadratic form of dimension $n \geq 3$ is particularly simple to compute, and can be found in terms of a few local computations, but the class number of a definite form is considerably more complicated to understand exactly.

Theorem 1.9.3. *The class number h_Q is finite.*

Proof. This follows from the reduction theory of quadratic forms, which shows that every class of quadratic forms over $\mathbb{Z}$ has some representative (of the same determinant) whose coefficients lie in a compact set. This together with the discreteness of the (integer) coefficients of Q gives that there are only finitely many classes of quadratic forms of bounded discriminant. A proof can be found in [Cas78, Theorem 1.1, p. 128 and Lemma 3.1, p. 135]. □

Remark 1.9.4. It is also useful to discuss the **proper class** of Q, denoted $\mathrm{Cls}^+(Q)$ which is the set of all $Q' \in \mathrm{Cls}(Q)$ where $Q'(\vec{x}) = Q(M\vec{x})$ with $\det(M) = 1$. Since $\det(M) \in \{\pm 1\}$, we see that there are at most two proper classes in a class, and so there are also finitely many proper classes in a given genus. The notion of proper classes is only meaningful when n is even (because when n is odd the

$n \times n$ scalar matrix $M = -1_n$ has $\det(-1_n) = -1$, so $\mathrm{Cls}(Q) = \mathrm{Cls}^+(Q)$), and is important for formulating the connection between proper classes of binary quadratic forms and ideal classes in quadratic number fields. This connection is discussed further in the Bhargava's notes [Bha].

There is also a somewhat more geometrical notion of the **class and genus of a quadratic lattice** $L \subset (V, Q)$, by considering the orbit of L under the action of the rational or adelic orthogonal group. In the language of quadratic forms, this says that two (free) quadratic lattices are in the same class or genus iff any associated quadratic forms (by choosing bases for the lattices) are in the same class or genus (respectively). This gives rise to a class number $h(L)$ which is the number of classes in the genus of L (and is again finite), and this agrees with the class number of the associated quadratic form when L is free. This notion of class and genus of a lattice is discussed in [Ger08, Definition 9.7, pp. 180–181] and will be revisited in Sect. 4.5.

Interestingly, while indefinite forms appear more complicated on the surface, their arithmetic is usually *easier* to understand than that of definite forms, as can be seen from the following theorems. The main idea is due to Eichler who discovered that the arithmetic of a certain simply connected algebraic group called the *spin group*, which is a double covering of $\mathrm{SO}(Q)$ and is very easy to understand via a property called "strong approximation". This naturally leads to a notion of **(proper) spinor equivalence**, and we call the orbit of L under this equivalence the **(proper) spinor genus** $\mathrm{Spn}^+(L)$ of L. We will discuss these notions briefly in Sect. 3.5. Some references for further reading about this topic are [Cas78, pp. 186–191], [Shi10, pp. 177–178, 192], [O'M71, pp. 315–321], [PR94, §7.4, pp. 427–433], and [Kne66].

1.10 The Neighbor Method

In this section we describe the method of neighboring lattices due to Kneser, which gives a useful construction for enumerating all classes in a given (spinor) genus of quadratic forms. The idea is that one can perform explicit operations on a given quadratic lattice L to produce different lattices that are obviously locally integrally equivalent to L. By doing this carefully, one can find representatives of all classes in the genus $\mathrm{Gen}(L)$.

Definition 1.10.1. Given two integer-valued quadratic lattices $L, L' \subset (V, Q)$ and some prime $p \in \mathbb{N}$, we say L and L' are p**-neighbors** if $[L : L \cap L'] = [L' : L \cap L'] = p$ and $H(L, L') \not\subseteq \mathbb{Z}$.

1.10.1 Constructing p-Neighbors

Given a quadratic lattice L in a non-degenerate quadratic space (V, Q), we now explain how to construct all of its p-neighboring lattices L' explicitly in terms of certain vectors in L.

Theorem 1.10.2. *If $p \in \mathbb{N}$ is prime, then every p-neighboring lattice L' of a given $\mathbb{Z}$-valued primitive quadratic lattice L with $\mathrm{Scale}_H(L) = \mathbb{Z}$ has the form*

$$L' = \tfrac{1}{p}\vec{w} + L_{\vec{w},p,\perp}$$

where

$$L_{\vec{w},p,\perp} := \{\vec{v} \in L \mid H(\vec{v},\vec{w}) \equiv 0 \pmod{p}\},$$

for some primitive vector $\vec{w} \in L$ with $p^2 \mid Q(\vec{w})$.

Proof. Any index p superlattice L' of L'' must be of the form $L' = L'' + \frac{1}{p}\vec{w}$ for some primitive vector $\vec{w}$ in L'', because by the structure theorem [Lan95, Theorem 7.8(i), p. 153] one can choose a basis for L' starting with some $\vec{w}$ so that replacing $\vec{w}$ by $p\vec{w}$ gives a basis for L''. For such an L' to be $\mathbb{Z}$-valued we must at least have $Q(\frac{1}{p}\vec{w}) \in \mathbb{Z}$, which is equivalent to $p^2 \mid Q(\vec{w})$. Further since every $\vec{x} \in L'$ can be written as $\vec{x} = \vec{y} + \frac{a}{p}\vec{w}$ for some $\vec{y} \in L_{\vec{w},p,\perp}$ and some $a \in \mathbb{Z}$, we have

$$\begin{aligned} Q(\vec{x}) &= Q(\vec{y} + \tfrac{a}{p}\vec{w}) \\ &= Q(\vec{y}) + H(\vec{y}, \tfrac{a}{p}\vec{w}) + Q(\tfrac{a}{p}\vec{w}) \\ &= \underbrace{Q(\vec{y})}_{\in\mathbb{Z}} + \tfrac{a}{p}H(\vec{y},\vec{w}) + \underbrace{aQ(\tfrac{1}{p}\vec{w})}_{\in\mathbb{Z}} \in \mathbb{Z}, \end{aligned}$$

and so we must have $H(\vec{y},\vec{w}) \in p\mathbb{Z}$ for all $\vec{y} \in L''$. However this condition defines an index p sublattice of L, since it is the kernel of the surjective homomorphism $L \to \mathbb{Z}/p\mathbb{Z}$ defined by $\vec{v} \mapsto H(\vec{v},\vec{w})$. By reversing our reasoning, we see that all such L' are p-neighbors of L. □

An important fact about p-neighbors L' of L is that they are all in the same genus $\mathrm{Gen}(L)$. It is interesting to ask how many classes in the genus of L can be created by taking repeated p-neighbors starting from L. If one is allowed to vary the prime p, then this p-neighbor procedure gives all (proper) classes in $\mathrm{Gen}(Q)$.

Theorem 1.10.3. 1. *If L' is a p-neighbor of L then $L' \in \mathrm{Gen}(L)$.*
2. *If $p \nmid 2\det(L)$ and $n \geq 3$, then any $L'' \in \mathrm{Spn}^+(L)$ can be obtained by taking repeated p-neighbors of L.*
3. *If the prime p is allowed to vary, then we can obtain all proper classes $\mathrm{Cls}^+(L)$ in $\mathrm{Gen}(L)$ by taking repeated p-neighbors of L.*

Proof. This definition of p-neighbor and local equivalence of p-neighbors is proved in [Tor05, §3.1, pp. 31–35]. The spanning of the spinor genus is proved in [BH83, Proposition 1, p. 339], and the spanning of the genus is proved in [BH83, Theorem 2, p. 340]. □

In fact, one can make more precise statements about exactly which spinor genera appear by taking p neighbors because the p-neighboring operation can always be realized by an element of spinor norm $p(\mathbb{Q}^\times)^2$. The image of this squareclass in the finite set of $\mathbb{Q}^\times$-squareclasses modulo spinor norms $\mathrm{sn}(O^+_{\mathbb{Q}}(V))$ and modulo the adelic spinor norms of the stabilizer of $O^+_{\mathbb{A}}(L)$ of L determines exactly which of the (at most two) spinor genera can be reached by taking repeated p-neighbors of L.

There is also a nice characterization of the p-neighbors of L in terms of the non-singular points of the associated hypersurface $Q(\vec{x}) = 0$ over $\mathbb{F}_p$.

Theorem 1.10.4. *The p-neighbors of L are in bijective correspondence with the non-singular points of $Q(\vec{x}) = 0$ in $\mathbb{P}^{n-1}(\mathbb{F}_p)$.*

Proof. See [Tor05, Theorem 3.5, p. 34] or [SH98, Proposition 2.2, p. 739]. □

The p-neighbors can be organized into a weighted **p-neighbor graph** whose vertices are the classes in $\mathrm{Gen}(L)$, where two vertices are connected by an edge iff they are p-neighbors, and where the multiplicity of each edge is the number of distinct p-neighboring lattices of L which are equivalent $\sim_{\mathbb{Z}}$ to L'. From the above theorem, we see that the p-neighbor graph is regular and that if $p \nmid 2\det(L)$ then the it is p^{n-2}-regular (i.e. every class has exactly p^{n-2} neighbors for the prime p).

2 Theta Functions

2.1 Definitions and Convergence

We say that $m \in \mathbb{Z}$ is **represented** by an integer-valued quadratic form Q in n variables if there is a solution $\vec{x} \in \mathbb{Z}^n$ to the equation $Q(\vec{x}) = m$. Similarly we say that m is **locally represented** by Q if there is a solution of $Q(\vec{x}) = m$ with $\vec{x} \in \mathbb{R}^n$ and also a solution $\vec{x} \in (\mathbb{Z}/M\mathbb{Z})^n$ for every $M \in \mathbb{N}$. Our main purpose in this section will be to study the **representation numbers**

$$r_Q(m) := \#\{\vec{x} \in \mathbb{Z}^n \mid Q(\vec{x}) = m\}$$

of a positive definite quadratic form Q over $\mathbb{Z}$, in order to understand something about which numbers $m \in \mathbb{Z} \geq 0$ are represented by Q. Our assumption here that Q is positive definite ensures that $r_Q(m) < \infty$, since there are only finitely many lattice points (in $\mathbb{Z}^n$) in the compact solid ellipsoid $\mathcal{E}_m : Q(\vec{x}) \leq m$ when $\vec{x} \in \mathbb{R}^n$.

It will also be important to consider the **(integral) automorphism group of** Q, which is defined as the set of invertible integral linear transformations preserving Q, i.e.

$$\mathrm{Aut}(Q) := \{M \in M_n(\mathbb{Z}) \mid Q(M\vec{x}) = Q(\vec{x}) \text{ for all } \vec{x} \in \mathbb{Z}^n\}.$$

Our previous compactness observation also tells us that $\#\mathrm{Aut}(Q) < \infty$, since any automorphism of Q is determined by its action on a basis of $\mathbb{Z}^n$ and by taking m large enough we can arrange that the (finitely many) integral vectors in $\mathcal{E}_m$ span $\mathbb{Z}^n$. Because automorphisms preserve the values $Q(\vec{x})$ of all vectors, they preserve the set of integral vectors inside the ellipsoid $\mathcal{E}_m$, and so there are only finitely many possible images of any specified spanning set.

In this setting it makes sense to define the **theta series of** Q as the Fourier series generating function for the representation numbers $r_Q(m)$ given by

$$\Theta_Q(z) := \sum_{m=0}^{\infty} r_Q(m) e^{2\pi i m z}.$$

From this perspective, our main goal will be to understand the symmetries of this generating function very well, and to use them to obtain information about the representation numbers $r_Q(m)$.

In order to make $\Theta_Q(z)$ more than just a formal object, we should try to establish a some convergence properties so it can be regarded as an honest function. For this series to converge absolutely we need the exponentials in the sum to be decaying, which happens for $z \in \mathbb{C}$ when $\mathrm{Im}(z) > 0$. For convenience, we denote by $\mathcal{H}$ the complex upper half-plane

$$\mathcal{H} := \{z \in \mathbb{C} \mid \mathrm{Im}(z) > 0\}.$$

The following theorem shows that any Fourier series with moderately (i.e. polynomially) growing coefficients will converge absolutely on $\mathcal{H}$.

Lemma 2.1.1 (Convergence of Fourier series). *The Fourier series*

$$f(z) := \sum_{m=0}^{\infty} a(m) e^{2\pi i m z}.$$

converges absolutely and uniformly on compact subsets of $\mathcal{H}$ to a holomorphic function $f : \mathcal{H} \to \mathbb{C}$ if all of its coefficients $a(m) \in \mathbb{C}$ satisfy $|a(m)| \leq Cm^r$ for some constant $C > 0$.

Proof. See [Miy06, Lemma 4.3.3, p. 117]. □

Because the number of lattice points in a smooth bounded region $\mathcal{R} \subset \mathbb{R}^n$ is approximately $\mathrm{Vol}(\mathcal{R})$, we see that $\sum_{i=0}^{M} r_Q(m) < CM^n$ for some constant C. Therefore for each m individually we must have that $r_Q(m) < C_1 M^{n-1}$ for some constant C_1, so the previous lemma shows that the theta function $\Theta_Q(z)$ converges (absolutely and uniformly) to a holomorphic function when $z \in \mathbb{C}$ and $\mathrm{Im}(z) > 0$. This gives the following important result:

Theorem 2.1.2. *The theta series $\Theta_Q(z)$ of a positive definite integer-valued quadratic form Q converges absolutely and uniformly to a holomorphic function $\mathcal{H} \to \mathbb{C}$.*

2.2 Symmetries of the Theta Function

While it is not obvious at first glance, $\Theta_Q(z)$ has a surprisingly large number of symmetries. From its definition as a Fourier series, it is clearly invariant under the transformation $z \mapsto z+1$, but this is not particularly special since this holds for any Fourier series. However there is an additional symmetry provided to us by Fourier analysis because we can also view the theta function as a sum of a quadratic exponential function over a lattice $\mathbb{Z}^n$, i.e.

$$\Theta_Q(z) = \sum_{m=0}^{\infty} r_Q(m) e^{2\pi i m z} = \sum_{\vec{x} \in \mathbb{Z}^n} e^{2\pi i Q(\vec{x}) z}.$$

This additional lattice symmetry is realized through the Poisson summation theorem:

Theorem 2.2.1 (Poisson Summation Formula). *Suppose that $f(\vec{x})$ is a function on $\mathbb{R}^n$ which decays faster than any polynomial as $|\vec{x}| \to \infty$ (i.e. for all $r \geq 0$ we know that $|\vec{x}|^r f(\vec{x}) \to 0$ as $|\vec{x}| \to \infty$). Then the equality*

$$\sum_{\vec{x} \in \mathbb{Z}^n} f(\vec{x}) = \sum_{\vec{x} \in \mathbb{Z}^n} \hat{f}(\vec{x})$$

holds and the sums on both sides are absolutely convergent, where

$$\hat{f}(\vec{x}) := \int_{\vec{y} \in \mathbb{R}^n} f(\vec{y}) e^{-2\pi i \vec{x} \cdot \vec{y}} \, dy$$

is the Fourier transform of $f(\vec{x})$.

Proof. See [Lan94, pp. 249–250] for a proof of this. □

The important point here is that the Gaussian function $f(x) = e^{-\pi \alpha x^2}$ transforms into a multiple of itself under the Fourier transform (which follows essentially from checking that $e^{-\pi x^2}$ is its own Fourier transform). Writing $z = x + iy \in \mathcal{H}$ in $\Theta_Q(z)$, we see that the y-dependence of each term will look like a decaying Gaussian (while the x-dependence will just oscillate), so Poisson summation allows us to transform each term into itself after a little rescaling. This allows us to establish an additional symmetry for the theta function under the transformation $z \mapsto \frac{-1}{Nz}$ for some $N \in \mathbb{N}$. In the special case where $Q(\vec{x}) = x^2$ we can take $N = 4$ and have the two identities

$$\Theta_{x^2}(-1/4z) = \sqrt{-2iz}\, \Theta_{x^2}(z) \qquad \text{and} \qquad \Theta_{x^2}(z+1) = \Theta_{x^2}(z). \qquad (2.1)$$

By extending these to the group generated by the transformations $z \mapsto \frac{-1}{4z}$ and $z \mapsto z+1$, we obtain following prototypical theorem.

Theorem 2.2.2. *For all* $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \mathrm{SL}_2(\mathbb{Z})$ *with* $4 \mid c$, *we have that*

$$\Theta_{x^2}\left(\frac{az+b}{cz+d}\right) = \varepsilon_d^{-1}\left(\frac{c}{d}\right)\sqrt{cz+d}\;\Theta_{x^2}(z)$$

where $\frac{-\pi}{2} < \arg(\sqrt{z}) \leq \frac{\pi}{2}$,

$$\varepsilon_d := \begin{cases} 1 & \textit{if } d \equiv 1 \mod 4, \\ i & \textit{if } d \equiv 3 \mod 4, \end{cases} \quad \textit{and} \quad \left(\frac{c}{d}\right) := \begin{cases} \left(\frac{c}{|d|}\right) & \textit{if } c > 0 \textit{ or } d > 0, \\ -\left(\frac{c}{|d|}\right) & \textit{if both } c, d < 0. \end{cases}$$

Here when $d > 0$ *the symbol* $\left(\frac{c}{d}\right)$ *agrees with the usual quadratic character mod d.*

Proof. This is stated in [Shi73, (1.10), p. 440] and proved in Proposition 2, p. 457. See also Iwaniec [Iwa97, Theorem 10.10 with $\chi = \mathbf{1}$, pp. 177–178], Knopp [Kno70, Theorem 13, p. 46 and Theorem 3, p. 51], [Miy06, Corollary 4.9.7, p. 194] and Andrianov-Zhuralev [AZ95, Proposition 4.15, p. 42] for proofs. □

For a diagonal quadratic form $Q(\vec{x}) = \sum_{i=1}^{n} a_i x_i^2$ of level N, this formula is enough to see that $\Theta_Q(z)$ transforms into a multiple of itself under the element $z \mapsto \frac{-1}{Nz}$. However to obtain a transformation formula for a general theta series $\Theta_Q(z)$ similar to Theorem 2.2.2, a more general strategy is needed. One approach is to compute the transformation formulae for more general theta series involving a linear term, and then to specialize this term to zero. Another approach is to obtain identities for how a related generalized higher dimensional theta function (similar to $\Theta_{x^2}(z)$) transforms with respect to a special subgroup of $\mathrm{Sp}_{2n}(\mathbb{Z})$ called the **theta group**. In the case where $n = 1$, the theta group consists the elements $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ of $\mathrm{Sp}_2(\mathbb{Z}) = \mathrm{SL}_2(\mathbb{Z})$ where both products ab and cd are even. Either approach allows one to show the following important transformation formula:

Theorem 2.2.3. *Suppose* Q *is a non-degenerate positive definite quadratic form over* $\mathbb{Z}$ *in* n *variables with level* N. *Then for all* $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \mathrm{SL}_2(\mathbb{Z})$ *with* $N \mid c$, *we have that*

$$\Theta_Q\left(\frac{az+b}{cz+d}\right) = \left(\frac{\det(Q)}{d}\right)\left[\varepsilon_d^{-1}\left(\frac{c}{d}\right)\sqrt{cz+d}\right]^n\;\Theta_Q(z),$$

where $\sqrt{z}$, ε_d, *and* $\left(\frac{c}{d}\right)$ *are defined in Theorem 2.2.2.*

Proof. A nice discussion of theta series and their transformation formulas (by the first approach) can be found in [Iwa97, Chap. 10], and a somewhat simpler discussion of transformation formulas for theta series in an even number of variables along these lines is given in the appendix to Chap. 1 of Eichler's book [Eic66, pp. 44–52]. The second approach described above can be found (in much greater generality) in [AZ95, Chap. 1, §3–4, pp. 11–37], especially Theorem 3.13 on p. 22. □

Here the 2×2 matrices that give symmetries of $\Theta_Q(z)$ form a subgroup of $\mathrm{SL}_2(\mathbb{Z})$ called the **level N congruence group**, which is usually denoted as

$$\Gamma_0(N) := \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \,\middle|\, c \equiv 0 \pmod{N} \right\}.$$

2.3 Modular Forms

It is useful to understand theta series in the context of all functions that have symmetries with respect to the action of congruence subgroups $\Gamma_0(N)$ by linear fractional transformations on $\mathcal{H}$. This idea leads us to define a very important class of functions called **modular forms**, whose symmetry properties essentially depend on three parameters: the **weight** $k \in \frac{1}{2}\mathbb{Z}$, the **level** $N \in \mathbb{N}$, and the **character** $\chi : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$. If the weight $k \notin \mathbb{Z}$, then we must specify an additional function $\varepsilon := \varepsilon(\gamma, k)$ called a **multiplier system**. For theta series this is called the "theta multiplier", but we will not be concerned with its exact form here.

Definition 2.3.1. We define a **modular form** of weight k, level N, Dirichlet character χ and multiplier system ε to be a holomorphic function $f : \mathcal{H} \to \mathbb{C}$ which transforms with respect to $\Gamma_0(N)$ under the rule

$$f\left(\frac{az+b}{cz+d}\right) = \varepsilon(\gamma, k)\chi(d)(cz+d)^k f(z) \tag{2.2}$$

for all $\gamma := \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma_0(N)$, and satisfies the additional technical condition that $f(z)$ is also "holomorphic" at the boundary values $\mathbb{P}^1(\mathbb{Q}) := \mathbb{Q} \cup \{\infty\}$ of the quotient $\Gamma_0(N)\backslash\mathcal{H}$.

It is standard notation to let $M_k(N, \chi)$ denote the $\mathbb{C}$-vector space of all modular forms of weight $k \in \frac{1}{2}\mathbb{Z}$, level N and character χ, where we assume the **trivial multiplier system** $\varepsilon(\gamma, k) := 1$ if $k \in \mathbb{Z}$ and the **theta multiplier system** $\varepsilon(\gamma, k) := \varepsilon_d^{-1}\left(\frac{c}{d}\right)$ if $k \notin \mathbb{Z}$.

We can now rephrase the symmetries of the theta function $\Theta_Q(z)$ using the language of modular forms. Good references for the general theory of modular forms are [Iwa97, DS05, Miy06, Kob93, Shi94], and the theta multiplier is described in detail in [Kno70, Chap. 4] and [Iwa97, Chap. 10]. One important observation to make about modular forms $f(z)$ is that the element $\left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right] \in \Gamma_0(N)$, and so the transformation formula above shows that $f(z+1) = f(z)$.[3] This periodicity together with the holomorphy of $f(z)$ shows that any modular form can be written as a complex **Fourier series**

[3] To justify this, notice that both the trivial and theta multiplier systems have value 1 on this element.

$$f(z) = \sum_{m=0}^{\infty} a(m) e^{2\pi i m z} = \sum_{m=0}^{\infty} a(m) q^m \qquad \text{where } q := e^{2\pi i z} \tag{2.3}$$

and the **Fourier coefficients** $a(m) \in \mathbb{C}$.

Corollary 2.3.2. *Suppose Q is a non-degenerate positive definite quadratic form over $\mathbb{Z}$ in n variables with level N. Then $\Theta_Q(z) \in M_{\frac{n}{2}}(N, \chi)$ is a modular form of weight $\frac{n}{2}$, level N and character $\chi(\cdot) = \left(\frac{(-1)^{\lfloor \frac{n}{2} \rfloor} \det(Q)}{\cdot}\right)$ (and multiplier system $\varepsilon(\gamma, k)$ specified above).*

Proof. This follows because $\varepsilon_d^2 = \left(\frac{-1}{d}\right)$, and so $\left[\varepsilon_d^{-1}\left(\frac{c}{d}\right)\right]^n = \left(\frac{-1}{d}\right)^{\lfloor \frac{n}{2} \rfloor} \cdot \begin{cases} \varepsilon_d^{-1}\left(\frac{c}{d}\right) & \text{if } n \text{ is odd,} \\ 1 & \text{if } n \text{ is even.} \end{cases}$ □

Remark 2.3.3. Note that here the "level" N refers both the level of the quadratic form as well as the level of the modular form (i.e. we use symmetries from $\Gamma_0(N)$).

To understand modular forms structurally, it is important to understand the action of $\Gamma_0(N)$ on $\mathcal{H}$ by linear fractional transformations $z \mapsto \frac{az+b}{cz+d}$. When $N = 1$, then $\Gamma_0(N)$ is all of $\mathrm{SL}_2(\mathbb{Z})$ and there is a well-known fundamental domain $\mathcal{F}$ for this action given by

$$\mathcal{F} := \{z \in \mathcal{H} \mid |z| \geq 1 \text{ and } |\mathrm{Re}(z)| \leq \tfrac{1}{2}\},$$

together with some identifications of its boundary. After these identifications have been made, the resulting fundamental domain $\mathcal{F}$ is not compact. However $\mathcal{F}$ *can be naturally extended to a compact surface* by adding one point (usually called ∞ or $i\infty$) which we imagine to be at the topmost end of the y-axis. This point is called a **cusp of** $\mathrm{SL}_2(\mathbb{Z})$ due to the apparent pointyness of $\mathcal{F}$ as we move along the y-axis towards $i\infty$. In general, $\Gamma_0(N)$ has finite index in $\mathrm{SL}_2(\mathbb{Z})$ and so its fundamental domain will be a union of finitely many translates of $\mathcal{F}$ (with slightly different boundary identifications). This larger fundamental domain is again not compact, but here it can be made compact by the addition of finitely many "boundary" points which we call **cusps of** $\Gamma_0(N)$. These cusps can always be represented by elements of $\mathbb{P}^1(\mathbb{Q})$ since they will be the image of the cusp $i\infty$ under some element of $\Gamma_0(N) \subset \mathrm{SL}_2(\mathbb{Q})$, and we have the identification $i\infty = \infty \in \mathbb{P}^1(\mathbb{Q}) \subset \mathbb{P}^1(\mathbb{C})$.

These cusps play a very important role in the theory of modular forms. For example, they can be used to define a natural subspace of modular forms which vanish at all cusps, called the **cusp forms** $S_k(N, \chi) \subseteq M_k(N, \chi)$. Also, for each cusp $\mathcal{C}$ of $\Gamma_0(N)$ we can usually construct a special modular form $E_{\mathcal{C}}(z)$ associated to $\mathcal{C}$ which has value 1 at $\mathcal{C}$ and vanishes at all other cusps. We call the space spanned by all of these functions associated to cusps the space of **Eisenstein series** $E_k(N, \chi) \subseteq M_k(N, \chi)$.

The Eisenstein series associated to cusps can be understood very explicitly, and is usually considered to be the "easier" part of $M_k(N, \chi)$. For example, for the cusp $\mathcal{C} = i\infty$ of $\mathrm{SL}_2(\mathbb{Z})$, the associated Eisenstein series $E_{\mathcal{C}}(z)$ of weight $k \in 2\mathbb{Z} > 2$ is given by

$$G_k(z) := \tfrac{1}{2} \sum_{\substack{(c,d)\in\mathbb{Z}^2 \\ \gcd(c,d)=1}} \frac{1}{(cz+d)^k} = 1 - \frac{2k}{B_{2k}} \sum_{m\geq 1} \sigma_{k-1}(m) q^m \in M_k(N=1, \chi = \mathbf{1}) \tag{2.4}$$

where B_{2k} is the $(2k)$th Bernoulli number, $\sigma_{k-1}(m) := \sum_{0<d|m} d^{k-1}$ is the usual divisor function and $q := e^{2\pi i z}$. (See [Miy06, Lemma 4.1.6, p. 100 and Theorem 3.2.3, p. 90].) We can also interpret the Fourier expansion in (2.3) as being associated with the cusp $i\infty$, since we can view q as a local parameter in the neighborhood of $i\infty$.

Facts about Modular forms: We now state several fundamental structural results in the theory of modular forms that are useful for understanding theta series:

1. The space of modular forms $M_k(N, \chi)$ with fixed invariants (k, N, χ) is a finite dimensional vector space over $\mathbb{C}$ [Miy06, §2.5, pp. 57–61].
2. The space $M_k(N, \chi)$ can be decomposed uniquely as a direct sum of cusp forms (of functions vanishing at all cusps) and Eisenstein series (spanned by the Eisenstein series associated to the cusps of $\Gamma_0(N)$) [Miy06, Theorem 2.1.7, p. 44 and Theorem 4.7.2, p. 179].
3. Any Eisenstein series has Fourier coefficients $a_E(m)$ which can be as large as $c_\varepsilon m^{k-1+\varepsilon}$ for any $\varepsilon > 0$ and some constant $c_\varepsilon \in \mathbb{R} > 0$ [Miy06, Theorem 4.7.3, p. 181].
4. Any cusp form has Fourier coefficients $a_f(m)$ which are (trivially) no larger than $c_\varepsilon m^{\frac{k}{2}+\varepsilon}$ for any $\varepsilon > 0$ and some constant $c_\varepsilon \in \mathbb{R} > 0$ [Miy06, Corollary 2.1.6, p. 43].

For our purposes, it is important to note that the upper bound on Eisenstein coefficients is not far from the truth, and is best possible when $k > 2$. When $k \in 2\mathbb{Z} > 2$, this bound is attained by the Eisenstein series (2.4).

2.4 Asymptotic Statements About $r_Q(m)$

To apply our knowledge of modular forms to study the numbers m represented by Q, we write the theta series as

$$\Theta_Q(z) = E(z) + C(z)$$

where $E(z)$ is an Eisenstein series and $C(z)$ is a cusp form. Looking at the mth Fourier coefficient of this equation gives a decomposition of the representation numbers as

$$r_Q(m) = a_E(m) + a_C(m).$$

From our informal discussion of modular forms above we know that the Eisenstein Fourier coefficients $a_E(m)$ are about as large as m^{k-1} as $m \to \infty$, and when $n = 5$ one can show using (2.6) that

$$|a_E(m)| \gg m^{\frac{3}{2}}$$

when they are non-zero, and when $n \geq 4$ this occurs $\iff$ m is locally represented by Q. Similarly we know that the cusp form Fourier coefficients satisfy

$$|a_C(m)| \ll m^{\frac{5}{4}+\varepsilon},$$

so if the Eisenstein coefficients are non-zero, then we know that $r_Q(m)$ is non-zero and so m is represented by Q if m is sufficiently large. This asymptotic estimate only improves when Q has more variables, giving the following theorem originally due to Tartakowski:

Theorem 2.4.1 (Tartakowski, [Tar29]). *If Q is a positive definite quadratic form over $\mathbb{Z}$ in $n \geq 5$ variables, then every sufficiently large number $m \in \mathbb{N}$ that is locally represented by Q is represented by Q.*

For $n \leq 4$, the above results are not enough to show that the Eisenstein coefficients are asymptotically larger than the cusp form coefficients, so more care is needed. The case $n = 4$ was first handled by Kloosterman by a clever refinement of the Circle Method (described briefly below), and has since been absorbed into the theory of modular forms as a consequence of the Ramanujan bound $|a_f(p)| \leq 2\sqrt{p}$ for prime coefficients of weight 2 cusp forms. This case also involves additional local considerations at finitely many primes p where Q is anisotropic over $\mathbb{Q}_p$.

The case $n = 3$ is even more delicate, and involves additional arithmetic and analytic tools to understand (e.g. spinor genera, the Shimura lifting of half-integral weight forms, analytic bounds on square-free coefficients of half-integral weight forms). In particular it was handled by Duke and Schulze-Pillot, and then by Schulze-Pillot in the papers [DSP90, SP00]. For more details on asymptotic results, see the survey papers [Han04, Duk97, Iwa87, SP04].

The case $n = 2$ of binary forms is a genuinely arithmetic problem (since for weight $k = 1$ both cusp forms and Eisenstein series coefficients satisfy $a(m) \ll m^{\varepsilon}$ for any $\varepsilon > 0$ [Ser77, §5.2(c), p. 220]) and it exhibits a much closer connection to explicit class field theory for quadratic extensions than the asymptotic results described here.

2.5 The Circle Method and Siegel's Formula

The origins of the many of the modern analytic techniques in the theory of quadratic forms have their origins in the famous "circle method" of Hardy, Littlewood and Ramanujan. The idea of this method is that one can express the number of representations $r_Q(m)$ for $Q = a_1x_1^2 + \cdots + a_nx_n^2$ as an integral over the unit circle which can be well-approximated by taking small intervals about angles which are rational multiples of 2π (where small here means small relative to the overall denominator of the rational multiples one considers). These rational angle contributions can be thought of locally (in terms of Gauss sums), and so we learn that local considerations give a good approximation of the number of representations $r_Q(m)$ for $Q = a_1x_1^2 + \cdots + a_nx_n^2$ when n is large enough. In the language of modular forms this method produces an Eisenstein series $E_Q(z)$ (called a "singular series") with multiplicative Fourier coefficients that agrees with the theta series $\theta_Q(z)$ at all rational points (and at ∞) so the difference $\theta_Q(z) - E_Q(z)$ is a cusp form and so must have asymptotically smaller Fourier coefficients than $E(z)$. This cusp form can be analyzed to various degrees, but the most naive bound for its Fourier coefficients gives non-trivial asymptotic information for the asymptotic behavior of $r_Q(m)$ for $m \geq 5$. (See [MW06, Chap. 6] and [Kno70, Chap. 5, pp. 63–87] for details.) The case $n = 4$ can also be handled, but requires an essential refinement of Kloosterman to obtain additional cancellation. (See [Iwa97, §11.4–5, pp. 190–199] and [IK04, §20.3–5, pp. 467–486] for more details.)

Siegel used these ideas to give quantitative meaning to the Fourier coefficients in the singular series both in terms of the underlying space of modular forms (as an Eisenstein series), but also in terms of the "local densities" associated to the quadratic form Q. In particular he proved the following theorem:

Theorem 2.5.1 (Siegel). *Suppose $Q(\vec{x})$ is a positive definite integer-valued quadratic form in $n \geq 5$ variables, whose theta series $\Theta_Q(z)$ is written (uniquely) as a sum of an Eisenstein series $E(z)$ and a cusp form $C(z)$. Then the Eisenstein series*

$$E(z) = \sum_{m \geq 0} a_E(m) e^{2\pi i m z}$$

can be expressed in two different ways:

Firstly, $E(z)$ can be recovered as a weighted sum of theta series over the genus of Q:

$$E(z) = \frac{\sum_{Q' \in \mathrm{Gen}(Q)} \frac{\Theta_{Q'}(z)}{|\mathrm{Aut}(Q')|}}{\sum_{Q' \in \mathrm{Gen}(Q)} \frac{1}{|\mathrm{Aut}(Q')|}}, \tag{2.5}$$

showing that $E(z)$ is a genus invariant. (That is, the theta series of any $Q' \in \mathrm{Gen}(Q)$ will have the same Eisenstein series $E(z)$.)

Secondly, the Fourier coefficients $a_E(m)$ can be expressed as an infinite local product

$$a_E(m) = \prod_{\textit{places } v} \beta_{Q,v}(m) \tag{2.6}$$

where the numbers $\beta_{Q,v}(m)$ are the ***local representation densities of*** *Q* ***at*** *m, defined by the limit*

$$\beta_{Q,v}(m) := \lim_{U \to \{m\}} \frac{\mathrm{Vol}_{\mathbb{Z}_v^n}(Q^{-1}(U))}{\mathrm{Vol}_{\mathbb{Z}_v}(U)} \tag{2.7}$$

where U runs over a sequence of open subsets of $\mathbb{Z}_v$ containing m with common intersection $\{m\}$, and the volumes appearing are the natural translation-invariant volumes on n-dimensional and 1-dimensional affine space over $\mathbb{Z}_v$ of total volume one.

Proof. See Siegel's Lecture notes [Sie63] or his original series of papers [Sie35, Sie36, Sie37]. □

These formulas are extremely important for the analytical theory of quadratic forms, and can be used to provide precise asymptotics for $r_Q(m)$ as $m \to \infty$. Extensions of this technique led Siegel to prove analogous results for more general kinds of theta functions which count representations of a quadratic form by another quadratic form. These are examples of "Siegel modular forms" which have analogous symmetries for the symplectic group Sp_{2r}. (See [AZ95] for more details.)

The formulas of Siegel were later generalized by Weil to a more representation-theoretic context by means of a certain very simple representation of a symplectic group called the "Weil representation" that we will meet later. This representation can be used to give a proof of Siegel's formulas in the case where Q is a positive definite quadratic form in $n \geq 5$ variables, and has been extended by Kudla and Rallis [KR88b, KR88a] to cover many more cases, including $n \geq 3$. It is interesting to see the progression of ideas from the circle method to modular forms to the Weil representation, and to notice that while the language used to obtain these results changes to suit our deepening perspective and context, the essential features (and technical difficulties) of the result remain very much the same.

These structural results about theta series and modular forms can also be generalized to understand theta series of totally definite $\mathcal{O}_F$-valued quadratic forms over totally real number fields F. These theta series are then Hilbert modular forms for a congruence subgroup of the group $\mathrm{GL}_2(\mathcal{O}_F)$ where $\mathcal{O}_F$ is the ring of integers of F. They can also be generalized to understand the number of representations of a smaller quadratic form Q' by Q, where this can be viewed in the lattice picture as counting the number of isometric embeddings of the quadratic lattice L' into L (which are quadratic lattices associated to Q' and Q respectively). In this context, the resulting theta series is a Siegel modular form for some congruence subgroup of the symplectic group $\mathrm{Sp}_{2n'}(\mathbb{Z})$, where Q' is a quadratic form in n' variables. (Notice that in the special case where $n' = 1$ we have $\mathrm{Sp}_2 = \mathrm{SL}_2$.) In both of these settings, Siegel's formulas remain essentially unchanged.

2.6 Mass Formulas

One useful application of the generalizations of Siegel's formula to representing quadratic forms Q' by a quadratic form Q is when we take $Q' = Q$. In this case, a generalization of Siegel's first formula (2.5) applied to the Qth-Fourier coefficient of the associated Siegel modular form gives

$$a_E(Q) = \frac{\sum_{Q'' \in \mathrm{Gen}(Q)} \frac{r_{Q''}(Q)}{|\mathrm{Aut}(Q'')|}}{\sum_{Q'' \in \mathrm{Gen}(Q)} \frac{1}{|\mathrm{Aut}(Q'')|}} = \frac{1}{\sum_{Q'' \in \mathrm{Gen}(Q)} \frac{1}{|\mathrm{Aut}(Q'')|}} \tag{2.8}$$

because the number of representations $r_Q(Q'')$ of any quadratic form $Q'' \in \mathrm{Gen}(Q)$ by Q is given

$$r_Q(Q'') = \begin{cases} \#\mathrm{Aut}(Q) & \text{if } Q'' \sim_{\mathbb{Z}} Q, \\ 0 & \text{if } Q'' \not\sim_{\mathbb{Z}} Q. \end{cases}$$

From an extension of Siegel's second formula (2.6), we also see that $a_E(Q)$ can be written as a product of local densities (though in this case an extra factor of 2 is needed). This motivates the definition of the **mass of a quadratic form** Q, denoted by $\mathrm{Mass}(Q)$, as

$$\mathrm{Mass}(Q) := \sum_{Q'' \in \mathrm{Gen}(Q)} \frac{1}{|\mathrm{Aut}(Q'')|}.$$

By Siegel's theorems we see that the mass is a local quantity, and can be computed from local knowledge about Q over $\mathbb{Z}_v$ at all places v.

Explicit computations of the mass are simple in principle, but often a bit painful to make explicit. These are known as "exact mass formulas", and they provide very useful information about the class number h_Q of a genus $\mathrm{Gen}(Q)$. As an example of this, using the fact that every quadratic form has at least two automorphisms (e.g. $\vec{x} \mapsto \pm\vec{x}$) we can see that

$$\mathrm{Mass}(Q) = \sum_{Q'' \in \mathrm{Gen}(Q)} \frac{1}{|\mathrm{Aut}(Q'')|} \leq \frac{h(Q)}{2}.$$

Therefore if the $\mathrm{Mass}(Q)$ is large then we know that the genus must contain many distinct classes. However the size of the mass of a positive definite form can be shown by local considerations to grow as we vary Q in an infinite family (e.g. if n grows, or if $\det(Q)$ grows and $n \geq 2$), so the class number can also be shown to get very large in these situations. One interesting application of this is the following result of Pfeuffer and Watson:

Theorem 2.6.1. *There are only finitely many (classes of) primitive positive definite quadratic forms Q over $\mathbb{Z}$ in $n \geq 2$ variables with class number $h_Q = 1$.*

In a long series of papers [Wat63]–[Wat84], Watson enumerated many of these class number one forms. More generally, Pfeuffer showed that there are finitely many totally definite primitive integer-valued quadratic forms Q in $n \geq 2$ variables with $h_Q = 1$ as we vary over all totally real number fields. (See [Pfe71, Pfe78] for details.)

It should also be noted that this is not the end of the story for mass formulas. There are many other connections (e.g. to Tamagawa numbers, Eisenstein series on orthogonal groups, and computational enumeration of classes in a genus) that we do not have space to mention here. As an example of one continuation of the story, in the past few years Shimura has defined a somewhat different notion of "mass" and "mass formula" for quadratic forms which instead of dividing the number of representations by the number of automorphisms, it counts the number of *equivalence classes* of representations in a genus *under the action of the automorphism group*. For a nice discussion of these see [Shi06b,Shi06a] and [Shi10, §37], as well as the more detailed [Shi04, §12–13]. These are very interesting, but do not fit within the framework we are describing here. They are also a good example of how a well-established theory is still evolving in new ways, and that there are many avenues left for future researchers to explore!

2.7 An Example: The Sum of 4 Squares

We conclude with a concrete example of how Siegel's formulas can be used to understand how many ways we can represent certain numbers as a sum of four squares. This question can be treated in many different ways, but the most definitive result is the following exact formula of Jacobi which he derived via the theory of elliptic functions.

Theorem 2.7.1 (Jacobi [Jac]). *For $m \in \mathbb{N}$, we have*

$$r_{x^2+y^2+z^2+w^2}(m) = 8 \cdot \sum_{\substack{0<d \mid m \\ 4 \nmid d}} d.$$

We will now derive some special cases of this result for certain m by using Siegel's formulas in Theorem 2.5.1. To do this, we first note that $Q(\vec{x}) = x^2 + y^2 + z^2 + w^2$ has class number $h_Q = 1$, so Siegel's first formula gives

$$a_E(m) = \frac{\sum_{Q'' \in \mathrm{Gen}(Q)} \frac{r_{Q''}(m)}{|\mathrm{Aut}(Q'')|}}{\sum_{Q'' \in \mathrm{Gen}(Q)} \frac{1}{|\mathrm{Aut}(Q'')|}} = \frac{\frac{r_Q(m)}{\cancel{|\mathrm{Aut}(Q)|}}}{\frac{1}{\cancel{|\mathrm{Aut}(Q)|}}} = r_Q(m).$$

Now we can apply Siegel's second formula to give the purely local formula

$$r_Q(m) = a_E(m) = \prod_v \beta_{Q,v}(m)$$

for $r_Q(m)$ in terms of local densities $\beta_{Q,v}(m)$ defined in (2.7). We now compute this infinite product to evaluate $r_Q(m)$ for some $m \in \mathbb{N}$. For convenience of notation, from now on we use the abbreviation $\beta_v(m) := \beta_{x^2+y^2+z^2+w^2,\, v}(m)$.

2.7.1 Canonical Measures for Local Densities

To compute the local densities $\beta_v(m)$ defined in (2.7) we use the "canonical" Haar measures μ on $\mathbb{Z}_v$ (i.e. additively invariant) uniquely defined by the normalizations

$$\mu_{\mathbb{Z}_p}(\mathbb{Z}_p) = 1, \qquad \mu_{\mathbb{R}}([0,1]) = 1.$$

Even if one is unfamiliar with the measure $\mu_{\mathbb{Z}_p}$, the important thing is that we can easily compute the measure of any set we are interested in. In particular, because we can write $\mathbb{Z}_p$ as the disjoint union

$$\mathbb{Z}_p = \bigsqcup_{a \in \mathbb{Z}/p^i\mathbb{Z}} a + p^i\mathbb{Z}_p$$

and each of these cosets has the same measure (by the additive invariance), we see that $\mu_{\mathbb{Z}_p}(p^i\mathbb{Z}_p) = \frac{1}{p^i}$ and also $\mu_{\mathbb{Z}_p^n}(p^i\mathbb{Z}_p^n) = \frac{1}{p^{n\cdot i}}$.

2.7.2 Computing $\beta_\infty(m)$

When $v = \infty$ we have $\mathbb{Z}_v = \mathbb{R}$, so we see that the local density $\beta_\infty(m)$ is the volume of a thin "shell" around the ellipsoid $x^2 + y^2 + z^2 + w^2 = m$ divided by the "thickness" of the shell (in m-space), which is some measure of the "surface area" of the 4-sphere of radius $r = \sqrt{m}$. To compute $\beta_\infty(m)$ we need to know the "volume" of the 4-ball $B_{4,r} : x^2 + y^2 + z^2 + w^2 \leq r^2$ is given by the well-known formula

$$\mathrm{Vol}(B_{4,r}) = \tfrac{\pi^2}{2} r^4.$$

(There are many ways to see this, for example as a consequence of Pappus's Centroid Theorem [Eve76, §6.18, p. 166] once the volume of the 3-ball $B_{3,r}$ is known to be $\frac{4}{3}\pi r^3$.)

We now compute $\beta_\infty(m)$ using the open sets $U = U_\varepsilon := (m-\varepsilon, m+\varepsilon)$, giving

$$\begin{aligned}
\beta_\infty(m) &:= \lim_{U \supset \{m\}, U \to \{m\}} \frac{\mathrm{Vol}_{\mathbb{R}^n}(Q^{-1}(U))}{\mathrm{Vol}_{\mathbb{R}}(U)} \\
&= \lim_{\varepsilon \to 0} \frac{\mathrm{Vol}_{\mathbb{R}^n}(Q^{-1}(U_\varepsilon))}{\mathrm{Vol}_{\mathbb{R}}(U_\varepsilon)} \\
&= \lim_{\varepsilon \to 0} \frac{\frac{\pi^2}{2}(\sqrt{m+\varepsilon}^4 - \sqrt{m-\varepsilon}^4)}{2\varepsilon} \\
&= \lim_{\varepsilon \to 0} \frac{\frac{\pi^2}{2}((m+\varepsilon)^2 - (m-\varepsilon)^2)}{2\varepsilon} \qquad (2.9) \\
&= \lim_{\varepsilon \to 0} \frac{\frac{\pi^2}{2}((\cancel{m^2} + 2m\varepsilon + \cancel{\varepsilon^2}) - (\cancel{m^2} - 2m\varepsilon + \cancel{\varepsilon^2}))}{2\varepsilon} \\
&= \lim_{\varepsilon \to 0} \frac{\pi^2}{\cancel{2}} \frac{\cancel{4}m\cancel{\varepsilon}}{\cancel{2}\cancel{\varepsilon}} \\
&= \pi^2 m.
\end{aligned}$$

2.7.3 Understanding $\beta_p(m)$ by Counting

When $v = p$, we can think of $\mathbb{Z}_p$ as coming from the quotients $\mathbb{Z}/p^i\mathbb{Z}$ where i is very large (i.e. $\mathbb{Z}_p = \varprojlim \mathbb{Z}/p^i\mathbb{Z}$). Because of this we can interpret the local density $\beta_p(m)$ as a statement about the number of solutions of $Q(\vec{x}) \equiv m \pmod{p^i}$ for sufficiently large powers p^i. More precisely, we have

Lemma 2.7.2. *When $v = p$ is a prime number and $Q(x)$ is a quadratic form in n variables, then we may write $\beta_p(m)$ as*

$$\beta_p(m) = \lim_{i \to \infty} \frac{\#\{\vec{x} \in (\mathbb{Z}/p^i\mathbb{Z})^n \mid Q(\vec{x}) \equiv m \pmod{p^i}\}}{p^{(n-1)i}}.$$

Proof. This follows from the definition by choosing open sets $U_i := p^i\mathbb{Z}_p$. Then

$$\mathrm{Vol}_{\mathbb{Z}_p}(U_i) = \frac{1}{p^i}$$

and each solution $\vec{x}$ of $Q(\vec{x}) \equiv m \pmod{p^i}$ gives a p-adic coset $\vec{x} + p^i\mathbb{Z}_p^n$ of solutions in $Q^{-1}(U_i)$. Therefore since $\mathrm{Vol}(p^i\mathbb{Z}_p^n) = \frac{1}{p^{ni}}$, we have

$$\mathrm{Vol}_{\mathbb{Z}_p^n}(Q^{-1}(U_i)) = \frac{1}{p^{ni}} \cdot \#\{\vec{x} \in (\mathbb{Z}/p^i\mathbb{Z})^n \mid Q(\vec{x}) \equiv m \pmod{p^i}\}$$

and so

$$\begin{aligned}
\beta_p(m) &= \lim_{i\to\infty} \frac{\mathrm{Vol}_{\mathbb{Z}_p^n}(Q^{-1}(U_i))}{\mathrm{Vol}_{\mathbb{Z}_p}(U_i)} \\
&= \lim_{i\to\infty} \frac{\frac{1}{p^{ni}} \cdot \#\{\vec{x} \in (\mathbb{Z}/p^i\mathbb{Z})^n \mid Q(\vec{x}) \equiv m \pmod{p^i}\}}{\frac{1}{p^i}} \\
&= \lim_{i\to\infty} \frac{\#\{\vec{x} \in (\mathbb{Z}/p^i\mathbb{Z})^n \mid Q(\vec{x}) \equiv m \pmod{p^i}\}}{p^{(n-1)i}}.
\end{aligned}$$

□

Philosophically we should think of this formula as the number of solutions $\pmod{p^i}$ divided by the "expected number" of solutions (based solely on knowing the dimension of $Q(\vec{x}) = m$ is $n-1$). To see that this limit actually exists, we need to invoke Hensel's lemma which (as a consequence) says that if i is sufficiently large then the sequence defining $\beta_p(m)$ is constant. In particular, for $Q(\vec{x}) = x_1^2 + \cdots + x_n^2$ it is enough to compute the (non-zero) solutions $\pmod{p}$ if $p > 2$ and $\pmod{8}$ if $p = 2$. In the next few sections we compute the local densities $\beta_p(m)$ by counting these numbers of solutions.

2.7.4 Computing $\beta_p(m)$ for All Primes p

Counting solutions to a polynomial equation over finite fields $\mathbb{Z}/p\mathbb{Z}$ can be done explicitly by the method of "exponential sums" (sometimes called Gauss sums or Jacobi sums), and this gives particularly simple formulas for degree 2 equations. One such formula is

Lemma 2.7.3. *Suppose $p \in \mathbb{N}$ is a prime > 2, then*

$$r_{x^2+y^2+z^2+w^2,\,p}(m) = \begin{cases} p^3 - p & \text{if } p \nmid m, \\ p^3 + p(p-1) & \text{if } p \mid m. \end{cases}$$

which gives the following explicit local density formulas:

Lemma 2.7.4. *Suppose $p \in \mathbb{N}$ is a prime > 2, then*

$$\beta_{x^2+y^2+z^2+w^2,\,p}(m) = \begin{cases} 1 - \frac{1}{p^2} & \text{if } p \nmid m, \\ \left(1 - \frac{1}{p^2}\right)\left(1 + \frac{1}{p}\right) & \text{if } p \mid m \text{ but } p^2 \nmid m. \end{cases}$$

Remark 2.7.5. The formula when $p \mid m$ follows by counting all solutions except $\vec{x} = \vec{0}$, since that solution will not lift by Hensel's lemma to a solution of $Q(\vec{x}) = m \pmod{p^2}$.

When $p = 2$ we need to understand the local densities $\pmod 8$, which we do by explicitly enumerating the values $Q(\vec{x})$ of all vectors $\vec{x} \in (\mathbb{Z}/8\mathbb{Z})^4$, giving

Lemma 2.7.6. *Suppose $p = 2$, then*

$$\beta_{x^2+y^2+z^2+w^2,\,2}(m) = \begin{cases} 1 & \textit{if}\; p \nmid m, \\ \frac{3}{2} & \textit{if}\; p \mid m \;\textit{but}\; p^2 \nmid m. \end{cases}$$

2.7.5 Computing $r_Q(m)$ for Certain m

We are now in a position to compute the number of representations $r_Q(m)$ for some simple numbers m. To warm up, we see that when $m = 1$ we have

$$\begin{aligned} r_Q(1) &= \prod_v \beta_v(1) = \beta_\infty(1)\,\beta_2(1) \prod_{p>2} \beta_p(1) && (2.10)\\ &= \left(\pi^2 \cdot 1\right)\,(1) \prod_{p>2} \left(1 - \frac{1}{p^2}\right) && (2.11)\\ &= \pi^2 \left(\frac{1}{1-\frac{1}{2^2}}\right) \prod_p \left(1 - \frac{1}{p^2}\right) && (2.12)\\ &= \frac{4\pi^2}{3} \prod_p \left(1 - \frac{1}{p^2}\right) && (2.13)\\ &= \frac{4\pi^2}{3} \frac{1}{\zeta(2)} && (2.14)\\ &= \frac{4\cancel{\pi^2}}{3} \frac{6}{\cancel{\pi^2}} && (2.15)\\ &= 8 && (2.16) \end{aligned}$$

which we could also have worked out (perhaps more quickly) by observing that if $Q(\vec{x}) = \sum_{i=1}^4 x_i^2 = 1$, then we must have $|x_i| \leq 1$ and at all but one x_i is zero.

Now suppose that $m = p > 2$ is prime. Then our computation at of $r_Q(p)$ looks almost the same as when $m = 1$ with the exception that the factors at $v = \infty$ and $v = p$ have changed. This gives

$$\begin{aligned} r_Q(p) &= r_Q(1) \cdot \frac{\beta_\infty(p)}{\beta_\infty(1)} \cdot \frac{\beta_p(p)}{\beta_p(1)} && (2.17)\\ &= r_Q(1) \cdot \frac{\cancel{\pi^2} p}{\cancel{\pi^2}} \cdot \frac{\cancel{\left(1 - \frac{1}{p^2}\right)} \left(1 + \frac{1}{p}\right)}{\cancel{\left(1 - \frac{1}{p^2}\right)}} && (2.18) \end{aligned}$$

$$= r_Q(1) \cdot p \cdot \left(1 + \tfrac{1}{p}\right) \tag{2.19}$$

$$= 8(p+1) \tag{2.20}$$

Finally, we suppose that m is an odd squarefree number t. Then the computation changes at $v = \infty$ and at all primes $p \mid t$, giving

$$r_Q(t) = r_Q(1) \cdot \frac{\beta_\infty(t)}{\beta_\infty(1)} \cdot \prod_{p \mid t} \frac{\beta_p(t)}{\beta_p(1)} \tag{2.21}$$

$$= r_Q(1) \cdot t \cdot \prod_{p \mid t} \tfrac{p+1}{p} \tag{2.22}$$

$$= 8 \prod_{p \mid t} (p+1). \tag{2.23}$$

We see that this agrees with Jacobi's divisor sum formula for $r_Q(m)$ in Theorem 2.7.1 since the positive divisors of t are exactly the terms appearing when the product $\prod_{p \mid t}(p+1)$ is fully expanded. One could continue to prove Jacobi's formula for $r_Q(m)$ for any $m \in \mathbb{N}$ by extending this computation, though the computations of the local densities $\beta_p(m)$ when $p = 2$ and at primes where $p^2 \mid m$ become somewhat more involved.

3 Quaternions and Clifford Algebras

In this section, we describe some important algebraic structures naturally associated with quadratic forms. One of them is the Clifford algebra, which one can think of an algebra that enhances a quadratic space with a multiplication law. The other is the Spin group, which is an algebraic group that is the "double cover" of the special orthogonal group and can be constructed naturally in terms of the Clifford algebra.

3.1 Definitions

Quadratic forms are closely connected with quadratic extensions, both those which are commutative (quadratic fields and their rings of integers) and also non-commutative (quaternion algebras and their maximal orders). We now explore some connections with non-commutative algebras of a particularly nice kind (known as "central simple algebras"), and describe their basic structure. Good references for central simple algebras are [Jac89, §4.6], [GS06, §1–2], [Lam05, Chaps. III–IV] and [Shi10, Chap. IV].

We begin by defining a **central simple algebra** A as a finite-dimensional (possibly non-commutative) algebra over a field k whose center is k and which contains no proper non-zero two-sided ideals. To make the dependence on k explicit, we sometimes write A as A/k. We say that the **dimension** of A/k is the dimension of A as a vector space over k.

Theorem 3.1.1. *Suppose that A_1 and A_2 are central simple algebras over k. Then the tensor product $A_1 \otimes_k A_2$ is also a central simple algebra over k.*

Proof. See [Jac89, Corollary 3, p. 219]. □

Another nice property of central simple algebras is that we can freely extend the base field k and preserve the property of being central simple (though now with a larger center!):

Theorem 3.1.2. *Suppose that A/k is a central simple algebra and K is a field containing k, then $A/K := A \otimes_k K$ is a central simple algebra over K of the same dimension as A/k.*

Proof. See [Jac89, Corollary 2, p. 219] and the discussion on the top of p. 220. □

The simplest examples of central simple algebras are the matrix algebras $M_n(k)$ (which have dimension n^2). Notice that any central simple algebra over k which is commutative must be just k itself, so in general central simple algebras are non-commutative. The next simplest example of a central simple algebra which is not a field (i.e. non-commutative) is called a **quaternion algebra**, and can be defined in terms of a basis $\mathcal{B} = \{1, i, j, \kappa\}$ satisfying the relations $i^2 = a$, $j^2 = b$, $\kappa := ij = -ji$ for some fixed $a, b \in k^\times$ (where we always assume that $\mathrm{char}(k) \neq 2$). This quaternion algebra is often referred to by the symbol $\left(\frac{a,b}{k}\right)$, though various different choices of a and b may give rise to isomorphic quaternion algebras (e.g. $\left(\frac{1,-1}{k}\right) \cong \left(\frac{4,-1}{k}\right)$).

If $A/k \cong M_n(k)$ for some n, we say that A is **split**. If it happens that $A \otimes_k K$ is split for some extension K of k, we say that A/k is **split by** K, or that K is a **splitting field** for A/k. The following theorem (and proof) shows that it is not too difficult to find a splitting field for any A/k:

Theorem 3.1.3. *If A/k is a central simple algebra over k, then A/k is split by some finite separable extension K/k.*

Proof. The existence of a finite extension splitting A follows from [Jac89, Theorem 4.8, p. 221] and the discussion on the top of p. 220. To see that they are not hard to construct explicitly, see [Jac89, Theorem 4.12, p. 224]. Finally separability of the extension follows from (the proof of) [GS06, Proposition 2.2.5, p. 22]. □

Since base change doesn't change the dimension of a central simple algebra, and we can always enlarge our base field so that A splits, we have the following useful corollary and definition:

Corollary 3.1.4. *The dimension of a central simple algebra is always a square.*

Definition 3.1.5. If A/k has dimension n^2, then we say that A has **degree** n.

We can use this idea to define a norm map $N_{A/k} : A \to k$ by extending scalars to the separable closure k^{sep}, which splits A by Theorem 3.1.3, giving an isomorphism $A/k^{\text{sep}} \cong M_n(k^{\text{sep}})$. We then define the **norm** $N_{A/k}(x)$ as the determinant of x under this isomorphism. Since $\det(x)$ is constant on conjugacy classes, the norm is independent of the choice of isomorphism, and is invariant under the Galois action as well, hence is in k. Since the determinant is multiplicative, we see that

$$N_{A/k}(\alpha\beta) = N_{A/k}(\alpha)N_{A/k}(\beta) \qquad \text{for all } \alpha, \beta \in A.$$

If it happens that every non-zero element of A is invertible (i.e. $\alpha \in A-\{0\} \implies$ there is some $\alpha' \in A$ so that $\alpha\alpha' = 1$ and $\alpha'\alpha = 1$) then we say that A is a **division algebra**. One can think of division algebras as natural non-commutative generalizations of (finite degree) field extensions K/k. In fact any non-zero element α of a central simple algebra A of degree n generates a commutative subalgebra $k[\alpha] \subseteq A$ of degree $[k[\alpha] : k]$ dividing n. In the case of a quaternion algebra one can realize the norm map in terms of a conjugation operation explicitly on the basis (by taking $\alpha = a + bi + cj + d\kappa \mapsto \bar{\alpha} := a - bi - cj - d\kappa$), giving the norm as $N_{A/k}(\alpha) = \alpha\bar{\alpha}$. The property of being a division algebra can be easily characterized in terms of the norm map.

Theorem 3.1.6. *A central simple algebra A over k is a division algebra iff the condition $N_{A/k}(\alpha) = 0 \iff \alpha = 0$ holds.*

Proof. Notice that α is invertible in $A \iff$ the left multiplication map $L_\alpha : A/k \to A/k$ is an invertible linear map (by taking $\alpha^{-1} := L_\alpha^{-1}(1)$). However L_α is invertible $\iff$ its linear extension $L_\alpha^{\text{sep}} : A/k^{\text{sep}} \to A/k^{\text{sep}}$ over the separable closure k^{sep} of k is invertible, which happens iff $\det(L_\alpha^{\text{sep}}) = N_{A/k}(\alpha)^n \neq 0$, where n is the degree of A over k. (See also [Pie82, §16.3, Corollary a, p. 300].)

In the special case where A is quaternion algebra this follows more directly by noticing that if α is invertible then its unique two-sided inverse has the form $\alpha^{-1} = \bar{\alpha} \cdot (N_{A/k}(\alpha))^{-1}$, which exists iff $N_{A/k}(\alpha) \neq 0$. □

The following important structural result of Wedderburn shows that division algebras play a crucial role in the study of central simple algebras. It is also the starting point for defining the Brauer group, which we will not discuss here, but is discussed in Parimala's lecture notes [Par] in this volume.

Theorem 3.1.7 (Wedderburn). *Every central simple algebra A over k is isomorphic to a matrix ring over a division algebra, i.e.*

$$A \cong M_n(D)$$

where D is a (unique) division algebra over k, and $n \in \mathbb{N}$ is the degree of A/k.

Proof. See [GS06, Theorem 2.1.3, p. 18]. □

We now specialize to consider quaternion algebras, which are very closely related to quadratic spaces and questions about quadratic forms. One important connection is given by considering the **associated quadratic space** $(V, Q) := (A/k, N_{A/k})$ of the quaternion algebra A/k.

Lemma 3.1.8. *A quaternion algebra A/k is uniquely determined (up to isomorphism) by its associated quadratic space.*

Proof. When $\mathrm{char}(k) \neq 2$ this is [Lam05, Theorem 2.5(a)–(b), pp. 57–58], and more generally this follows from [Knu91, Chap. V, Proposition 2.4.1, p. 256]. □

In this language we have the following useful corollary of Theorem 3.1.6.

Corollary 3.1.9. *A quaternion algebra A/k is a division algebra iff its associated (4-dimensional) quadratic space is anisotropic.*

By combining this with the theory of local invariants of quadratic spaces in Sect. 1.6, we have the following uniqueness result:

Theorem 3.1.10. *There is a unique quaternion division algebra over each of the local fields $\mathbb{Q}_p$ and $\mathbb{R}$.*

Proof. Over $\mathbb{R}$ we see that $\left(\frac{a,b}{\mathbb{R}}\right)$ is determined by the signs of a and b, and that this is split iff at least one of them is > 0. The remaining case gives $a = b = -1$, which gives the Hamiltonian quaternions $\mathcal{H}$ and is the unique division algebra over $\mathbb{R}$.

Over $\mathbb{Q}_p$ this follows from Lemma 3.1.8 and the fact that there is a unique 4-dimensional anisotropic quadratic space over $\mathbb{Q}_p$ (characterized by the Hilbert symbol relation $c_p = (-1, -d_p)_p$) [Cas78, Lemma 2.6, p. 59]. □

Therefore, since every non-split quaternion algebra is a division algebra we see that

Theorem 3.1.11. *There are exactly two quaternion algebras (up to isomorphism) over each of the local fields $k = \mathbb{Q}_p$ or $\mathbb{R}$: the split algebra $M_2(k)$, and a division algebra D.*

When $k = \mathbb{Q}_p$ or $\mathbb{R}$, the dichotomy of Theorem 3.1.11 is often referred to as saying that a quaternion algebra A/k is either **split** or **ramified** (when it is a division algebra). The term "ramified" is used here because in the associated valuation theory of local division algebras (which is discussed in [Shi10, §21, particularly Theorem 21.17, p. 108]), the division quaternion algebra D has a valuation ring with maximal ideal $\mathfrak{p}$ satisfying $\mathfrak{p}^2 = (p) := p\mathbb{Z}$, which agrees with the usual notion of ramification in algebraic number theory.

To decide whether the local quaternion algebra A/k above is split or ramified, one can use the easily computable **(local) Hilbert symbol**

$$(\cdot,\cdot)_v : \mathbb{Q}_v^\times/(\mathbb{Q}_v^\times)^2 \times \mathbb{Q}_v^\times/(\mathbb{Q}_v^\times)^2 \longrightarrow \{\pm 1\} \tag{3.1}$$

which is a non-degenerate multiplicative symmetric bilinear form on the (non-zero) squareclasses of $\mathbb{Q}_v$. The Hilbert symbol arises naturally in the study of Class Field Theory [Neu99, Chap. V, §3, with $n = 2$], and is defined by the (not obviously symmetric) relation $(a, b)_v = 1 \iff a \in N_{K_v/\mathbb{Q}_v}(K_v^\times)$ where $K_v := \mathbb{Q}_v(\sqrt{b})$. The Hilbert symbol has many interesting properties:

Theorem 3.1.12. *The local Hilbert symbol defined in (3.1) satisfies the following properties:*

1. $(a, b)_v = (b, a)_v$, *(symmetry)*
2. $(a_1 a_2, b)_v = (a_1, b)_v (a_2, b)_v$, *(bilinearity)*
3. $(a, b)_v = 1$ *for all* $b \in \mathbb{Q}_v^\times/(\mathbb{Q}_v^\times)^2 \implies a \in (\mathbb{Q}_v^\times)^2$, *(non-degeneracy)*
4. $(a, -a)_v = (a, 1 - a)_v = 1$, *(symbol)*
5. $(a, b)_p = 1$ *if* $p \neq 2$ *and* $\mathrm{ord}_p(a), \mathrm{ord}_p(b) \in 2\mathbb{Z}$.

Proof. This follows from [Cas78, Lemma 2.1, 42] except for $(a, 1 - a)_v = 1$, which follows since $1 - a = N_{\mathbb{Q}_v(\sqrt{a})/\mathbb{Q}_v}(1 + \sqrt{a})$. See also [Neu99, Chap. V, Proposition 3.2, p. 334] for the anaogous proofs over number fields. □

Hilbert symbols are also an example of a "symbol" in the sense of K-theory (see [Lam05, Chap. V, §6 and Chap. X, §6, p. 362] and [NSW08, Chap. VI, §4, p. 356]), but for our purposes it is enough to be able to explicitly compute them, which can be done with the tables on [Cas78, pp. 43–44]. The question of computing Hilbert symbols (and splitting of quaternion algebras) over number fields is discussed in Voight's paper [Voi] in this volume.

Finally, the Hilbert symbol also satisfies the global "reciprocity" relation, from which quadratic reciprocity can be easily proved.

Theorem 3.1.13. *For all* $a, b \in \mathbb{Q}^\times$, *we have the product formula*

$$\prod_v (a, b)_v = 1,$$

and all but finitely many factors are one.

Proof. See [Cas78, Lemma 3.4, 46] or [Neu99, Chap. VI, Theorem 8.1, p. 414] for the analogous result over number fields. □

This theorem has the following important parity consequence for quaternion algebras $A/\mathbb{Q}$.

Corollary 3.1.14. *Given any quaternion algebra* $A/\mathbb{Q}$, *the set of places* v *where* $A/\mathbb{Q}_v$ *is ramified has even cardinality.*

Proof. By writing $A/\mathbb{Q}$ as $\left(\frac{a,b}{\mathbb{Q}}\right)$ for some $a, b \in \mathbb{Q}^\times$, we see that $A/\mathbb{Q}_v$ is ramified $\iff (a, b)_v = -1$, and the product formula guarantees this happens an even number of times. □

3.2 The Clifford Algebra

Good basic references for Clifford algebras over fields of characteristic $\neq 2$ are [Jac89, §4.8], [Lam05, Chap. V], and [Cas78, Chap. 10]. The valuation theory of central simple algebras over number fields can be found in [Shi10], and Clifford algebras over general rings are discussed thoroughly in [Knu91, Chaps. IV–V]. A very in-depth treatment of Clifford algebras as well as automorphic forms on their associated Spin groups can be found in the recent book of Shimura [Shi04].

Given a quadratic space (V, Q) over a field K (of characteristic $\neq 2$) of dimension n, we define its **Clifford algebra** $C(V)$ as the K-algebra generated by all formal multiplications of scalars $k \in K$ and vectors $\vec{v} \in V$ subject to the family of "squaring relations" that $\vec{v}^{\,2} = \vec{v} \cdot \vec{v} = Q(\vec{v})$ for all $\vec{v} \in V$. More formally, we can construct the Clifford algebra as a quotient $C(V) := T(V)/I(V)$ of the tensor algebra $T(V) = \oplus_{i=0}^{\infty}(\otimes^i V)$ by the ideal of relations

$$I(V) := \begin{array}{c}\text{the ideal of } T(V) \text{ generated by the set} \\ \{\vec{v}^{\,2} - Q(\vec{v}),\ \ k \cdot \vec{v} - k\vec{v} \ \text{ for all } \vec{v} \in V, k \in K\}.\end{array}$$

This shows that $C(V)$ is well-defined (and we will soon see that it is non-zero!).

One useful observation is that there is also a nice relationship between multiplication in $C(V)$ and the inner product $B(\vec{v}, \vec{w})$. We see this by expanding out

$$Q(\vec{v} + \vec{w}) = (\vec{v} + \vec{w})^2 \tag{3.2}$$

$$= \vec{v}^{\,2} + \vec{v} \cdot \vec{w} + \vec{w} \cdot \vec{v} + \vec{w}^{\,2} \tag{3.3}$$

$$= Q(\vec{v}) + (\vec{v} \cdot \vec{w} + \vec{w} \cdot \vec{v}) + Q(\vec{w}) \tag{3.4}$$

and comparing this with the polarization identity (1.2), which shows that

$$\vec{v} \cdot \vec{w} + \vec{w} \cdot \vec{v} = 2B(\vec{v}, \vec{w}). \tag{3.5}$$

This relation can be used to give a unique presentation of any element $\alpha \in C(V)$ in terms of a given choice of basis $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ of V, since we can reverse the order of adjacent elements to present them in terms of the basis of all possible products of distinct vectors $\vec{v}_i \in \mathcal{B}$ with indices i in increasing order. Because these products are indexed by the 2^n subsets I of $\{1, \ldots, n\}$, we see that

Theorem 3.2.1. *The dimension of the Clifford algebra is* $\dim_K(C_0(V)) = 2^n$.

An interesting special case of the Clifford algebra is when the quadratic form Q is identically zero. In this case, by taking a basis $\mathcal{B}$ for V and applying the relations above we see that $\vec{v}_i \cdot \vec{v}_j = -\vec{v}_j \cdot \vec{v}_i$ when $i \neq j$ and $\vec{v}_i^{\,2} = 0$. This shows that $C(V)$ is just the exterior algebra $\oplus_i(\bigwedge^i V)$. When Q is not identically zero we can think of $C(V)$ as a deformation of the exterior algebra that encodes the arithmetic of Q.

Another interesting fact about the Clifford algebra is that it has a natural $(\mathbb{Z}/2\mathbb{Z})$-grading (called the **parity**) coming from the $\mathbb{Z}$-grading on the tensor algebra $T(V)$ and the fact that the relations in $I(V)$ only involve relations among elements of the same parity. We say that an element of $C(V)$ is said to be **even** or **odd** if it can be written as a sum of elements of $T(V)$ of even or odd degree respectively. Given this, we can consider the subalgebra $C_0(V)$ of even elements in $C(V)$, called the **even Clifford algebra** of V. It follows from our basis description of $C(V)$ above and simple facts about binomial coefficients that

Theorem 3.2.2. *The dimension of the even Clifford algebra is* $\dim_K(C_0(V)) = 2^{n-1}$.

Both $C(V)$ and $C_0(V)$ have a **canonical involution** $\alpha \mapsto \tilde{\alpha}$ defined (on the pure tensor elements) by reversing the order of every product of vectors, i.e.

$$\alpha := \vec{v}_1 \cdots \vec{v}_k \longmapsto \vec{v}_k \cdots \vec{v}_1 =: \tilde{\alpha},$$

and then extending this map linearly to the entire algebra. We can use this involution to define a multiplicative **norm function** $N : C(V) \to K$ by the product

$$N(\alpha) := \alpha \cdot \tilde{\alpha}.$$

To see that $N(\alpha) \in K$ notice that if $\alpha := \vec{v}_1 \cdots \vec{v}_k$ then

$$N(\alpha) = \alpha \cdot \tilde{\alpha} = (\vec{v}_1 \cdots \vec{v}_k) \cdot (\vec{v}_k \cdots \vec{v}_1) = Q(\vec{v}_k) \cdots Q(\vec{v}_1) \in K, \tag{3.6}$$

by repeatedly using the relation $\vec{v}_i{}^2 = Q(\vec{v}_i)$. This also shows that $N(\alpha) = \tilde{\alpha} \cdot \alpha$.

We will be interested in multiplicative subgroups of $C(V)$, so we say that an element $\alpha \in C(V)$ is **invertible** if there is some $\alpha^{-1} \in C(V)$ so that $\alpha \cdot \alpha^{-1} = 1$. Notice that if α is invertible, then

$$\alpha^{-1} = \frac{\tilde{\alpha}}{N(\alpha)}$$

and we also have $\alpha^{-1} \cdot \alpha = 1$, so our inverses are "two-sided". It is also useful to notice that we already have a good understanding of what vectors $\vec{v} \in V$ are invertible.

Lemma 3.2.3. *Suppose that* $\vec{v} \in V \subset C(V)$. *Then* $\vec{v}$ *is invertible* $\iff$ $\vec{v}$ *is anisotropic.*

Proof. This follows since $\vec{v}$ is invertible $\iff N(\vec{v}) = \vec{v} \cdot \vec{v} = Q(\vec{v}) \neq 0$. □

Finally, we mention the following theorem that explains the structure of the Clifford algebra as a central simple algebra.

Theorem 3.2.4. *Suppose that V is a non-degenerate quadratic space of dimension n. Then $C(V)$ is a central simple algebra if n is even and $C_0(V)$ is a central simple algebra when n is odd.*

Proof. See [Jac89, Theorem 4.14, p. 237], [Lam05, Theorems 2.4 and 2.5, p. 110], [Shi10, Theorem 23.8, p. 125]. A more general version of this holds over rings, where the word "Azumaya" replaces "central simple". (See [Knu91, Chap. IV, Theorem 2.2.3, p. 203 and Theorem 3.2.4(1), p. 210] for proofs, and [Sal99, Chap. 2] for a discussion of Azumaya.) □

3.3 Connecting Algebra and Geometry in the Orthogonal Group

One important feature of the orthogonal group $O(V)$ is that can be used to describe equivalences of quadratic spaces and quadratic lattices, however this is not very useful unless one can somehow describe the elements of $O(V)$. One approach for doing this is to try to use the geometry of V to construct explicit transformations in $O(V)$. We now describe how this is done (over fields K of characteristic $\mathrm{char}(K) \neq 2$).

Given some $\vec{v} \in V$ with $Q(\vec{v}) \neq 0$, we can define the **reflection symmetry** $\tau_{\vec{v}} \in O(V)$ defined by sending $\vec{v} \mapsto -\vec{v}$ and pointwise fixing all vectors in the orthogonal complement $(K\vec{v})^{\perp}$ of the line spanned by $\vec{v}$. Using the standard projection formulas of linear algebra (e.g. [Str09, §4.2]) we see that $\tau_{\vec{v}}$ can be written explicitly as

$$\tau_{\vec{v}}(\vec{w}) = \vec{w} - 2\frac{B(\vec{v},\vec{w})}{B(\vec{v},\vec{v})}\vec{v}, \tag{3.7}$$

which is only defined if $B(\vec{v},\vec{v}) = Q(\vec{v}) \neq 0$. Notice that since we are reversing the direction of a line, and stabilizing its complement, we know $\det(\tau_{\vec{v}}) = -1$. One useful property of these reflection symmetries is that they can be explicitly seen to act transitively on vectors of a given non-zero length. More precisely,

Lemma 3.3.1. *Suppose that $\vec{v}, \vec{w} \in V$ satisfy $Q(\vec{v}) = Q(\vec{w})$.*

(a) *If $Q(\vec{v}) = Q(\vec{w}) \neq 0$. Then $\alpha\vec{v} = \vec{w}$ where α is a product of at most two reflection symmetries.*
(b) *If $Q(\vec{v} - \vec{w}) \neq 0$, then*

$$\tau_{\vec{v}-\vec{w}}(\vec{v}) = \vec{w}.$$

Proof. This can be found in [Cas78, pp. 19–20] among other places, though we give the argument here. Part (b) follows from a direct computation with (3.7). Part (a) follows from (b) if $Q(\vec{v} - \vec{w}) \neq 0$, otherwise the polarization identity (1.2) ensures that $Q(\vec{v}+\vec{w}) \neq 0$, and so part (b) allows us to find a symmetry interchanging $\vec{v}$ and $-\vec{w}$. From here the symmetry $\tau_{\vec{w}}$ interchanges $\vec{w}$ and $-\vec{w}$, giving $\alpha = \tau_{\vec{w}} \cdot \tau_{\vec{v}+\vec{w}}$. □

This transitive action of reflection symmetries shows that they generate the full orthogonal group $O(V)$ for any non-degenerate quadratic space.

Theorem 3.3.2. *If (V, Q) is a non-degenerate quadratic space, then every element $\beta \in O(V)$ can be written as a product of reflection symmetries, i.e.*

$$\beta = \tau_{\vec{v}_1} \cdots \tau_{\vec{v}_k}$$

for some vectors $\vec{v}_i \in V$ with $Q(\vec{v}_i) \neq 0$.

Proof. This is proved [Cas78, Lemma 4.3, pp. 20–21]. This follows by induction on the dimension on V, since for any $\vec{v}$ with $Q(\vec{v}) \neq 0$ we can find some product of symmetries α so that $\alpha\vec{v} = \beta\vec{v}$. Therefore $\alpha^{-1}\beta$ fixes $K\vec{v}$ and also $W := (K\vec{v})^{\perp}$, and we are reduced to showing that $\alpha^{-1}\beta$ is a product of symmetries on W. When $\dim(W) = 1$, this holds because $\beta : \vec{v} \mapsto \pm\vec{v}$, completing the proof. □

There is a particularly interesting map called the **spinor norm map**, denoted $\mathrm{sn}(\alpha)$, from $O^+(V)$ to the squareclasses $K^\times/(K^\times)^2$ that can be defined easily by using the reflection symmetry description of $O(V)$. To do this we write $\alpha \in O^+(V)$ as a product of symmetries $\tau_{\vec{v}}$ and define

$$\alpha = \tau_{\vec{v}_1} \cdots \tau_{\vec{v}_k} \longmapsto Q(\vec{v}_1) \cdots Q(\vec{v}_k) =: \mathrm{sn}(\alpha). \tag{3.8}$$

This gives a squareclass because we could have rescaled any of the $\vec{v}_i$ without changing its associated symmetry $\tau_{\vec{v}_i}$, but we would change $Q(\vec{v}_i)$ by a non-zero square. However it is more work to show that $\mathrm{sn}(\alpha)$ is independent of our particular presentation of α as a product of symmetries.

Lemma 3.3.3. *Suppose that (V, Q) is a non-degenerate quadratic space. Then any even element in the center of $C(V)$ is a scalar.*

Proof. This can be shown by taking an orthogonal basis $\{\vec{e}_i\}$ for V, which necessarily satisfies $\vec{e}_i\vec{e}_j = -\vec{e}_j\vec{e}_i$, and imposing the commutation relation. This is done explicitly in [Cas78, Lemma 2.3, p. 174], [O'M71, §54:4, pp. 135–136], [Shi10, Corollary 23.9, p. 126] and somewhat indirectly in [Lam05, Theorem 3.4, p. 92 and Theorem 2.2, p. 109]. □

Theorem 3.3.4. *Suppose that (V, Q) is a non-degenerate quadratic space. Then the spinor norm map sn $: O(V) \to F^\times/(F^\times)^2$ is a well-defined group homomorphism.*

Proof. To see that $\mathrm{sn}(\alpha)$ is well-defined, notice that any two expressions for α as a product of transpositions gives rise to an expression for the identity map as a product of an even number of reflection symmetries

$$\prod_i \tau_{\vec{v}_i} = \mathrm{id} \in SO(V),$$

and $\mathrm{sn}(\alpha)$ is well-defined iff $\prod_i Q(\vec{v}_i) \in (F^\times)^2$. Lemma 3.4.1 allows us to interpret $\tau_{\vec{v}}$ as conjugation by $\vec{v}$ in $C(V)$ and letting $u := \prod_i \vec{v}_i \in C_0(V)$ we see that $u\vec{w}u^{-1} = \vec{w}$ for all $\vec{w} \in V$. Therefore by Lemma 3.3.3 we know that $u \in K^\times$, and so $\prod_i Q(\vec{v}_i) = u\tilde{u} = u^2 \in (K^\times)^2$. (This argument also appears in [Cas78, Corollary 3, p. 178], [Shi10, §24.8, p. 131], [O'M71, §55, p. 137] and [Lam05, Theorem 1.13, p. 108].) □

3.4 The Spin Group

Now that we understand some basic properties of the Clifford algebra and the orthogonal group, we are ready to construct a very useful "two-fold cover" of the special orthogonal group $SO(V)$ called the spin group. Aside from being interesting in its own right, the spin group plays a very important role in the theory of indefinite quadratic forms.

As a first step, we notice that conjugation in the Clifford algebra is a very interesting operation because it naturally produces elements of the orthogonal group. For example, if it happens that $u \in C(V)^\times$ satisfies $u^{-1}Vu \subseteq V$ then we claim that this conjugation gives an isometry of V, and so it is an element of the orthogonal group $O(V)$. To see this, for any $\vec{x} \in V$ we compute

$$Q(u^{-1}\vec{x}u) = (u^{-1}\vec{x}u)(u^{-1}\vec{x}u) = u^{-1} \cdot \vec{x} \cdot \vec{x} \cdot u = Q(\vec{x}).$$

Amazingly, we can even identify exactly which element of $O(V)$ this conjugation gives us.

Lemma 3.4.1. *Suppose that $u \in V$ satisfies $Q(u) \neq 0$ and that for all $\vec{x} \in V$ the conjugation $\varphi_u : \vec{x} \mapsto u^{-1}\vec{x}u \in V$. Then $\varphi_u \in O(V)$ and φ_u gives the negative reflection symmetry $-\tau_u$.*

Proof. We have already seen that $\varphi_u \in O(V)$, so we only need to identify φ_u explicitly as

$$\begin{aligned}
\varphi_u = u^{-1}\vec{x}u &= \frac{u}{Q(u)}\vec{x}u \\
&= \frac{1}{Q(u)}\left[(u\vec{x} + \vec{x}u)u - xu^2\right] \\
&= \frac{1}{Q(u)}\left[2B(x,u)u - xQ(u)\right] \\
&= -x + \frac{2B(x,u)}{Q(u)}u \\
&= -\tau_u(\vec{x}).
\end{aligned}$$

□

This leads us to define the multiplicative subgroup

$$U_0 := \{u \in (C_0(V))^\times \mid u^{-1}Vu \subseteq V\} \subseteq C_0(V)$$

on which we have a natural **conjugation map** $\varphi : U_0 \to O(V)$ defined by sending $u \mapsto (\varphi_u : \vec{x} \mapsto u^{-1}\vec{x}u)$. It takes a little work to see that the image of this map is in $SO(V)$.

Lemma 3.4.2. *The conjugation map above gives a homomorphism $\varphi : U_0 \to SO(V)$.*

Proof. If $u \in U_0$ then by Theorem 3.3.2 we can find r anisotropic vectors $\vec{v}_i$ so that $\alpha := \prod_i \vec{v}_i$ gives $\varphi(\alpha) = \varphi(u)$, and so $\beta := \alpha \cdot u^{-1}$ has $\varphi(\beta) = \mathrm{id} \in O(V)$. This is equivalent to the commutation relation $\beta\vec{v} = \vec{v}\beta$ for all $\vec{v} \in V$. However by expressing β as a unique linear combination of ordered monomials with respect to some fixed orthogonal basis $\{\vec{w}_i\}$for V, this commutation relation for $\vec{w}_i$ says that each monomial containing $\vec{w}_i$ must have even degree, and so $\beta \in C_0(V)$. Therefore $\alpha \in C_0(V)$, r is even and $\varphi(u) = \varphi(\alpha) \in SO(V)$.

This argument can be found in [Cas78, Theorem 3.1, pp. 176–177], [Shi10, Theorem 24.6, pp. 129–130], and there Shimura points out that this result is originally due to Lipschitz [:1959], though a special case was shown by Clifford. □

This map is very useful for connecting the Clifford algebra and the special orthogonal group, as it provides a natural and explicit covering.

Lemma 3.4.3. *Suppose that (V, Q) is non-degenerate quadratic space. Then the conjugation map $\varphi : U_0 \to SO(V)$ is surjective with kernel $K^\times$, and so $U_0/(K^\times) \xrightarrow{\sim} SO(V)$.*

Proof. Surjectivity follows from Theorem 3.3.2 and Lemma 3.4.1. To see that $\mathrm{Ker}(\varphi) = K^\times$, use Theorem 3.3.2 and Lemma 3.3.3.

This can also be found in [Cas78, Theorem 3.1, p. 176] and [Shi10, Theorem 24.6(iii), p. 129]. □

Using this Lemma, we define the **spin group** $\mathrm{Spin}(V)$ as the elements of $\alpha \in U_0$ with norm $N(\alpha) = 1$. The spin group an algebraic section of the covering map $\varphi : U_0 \to SO(V)$, which we soon show is a "double covering" of its image. A helpful observation for doing this is that the spinor norm of an element of $\mathrm{Spin}(V)$ under this composition can be computed fairly easily.

Lemma 3.4.4. *Suppose that (V, Q) is non-degenerate quadratic space. Then for any $u \in U_0$ we have $sn(u) = N(u)(K^\times)^2$.*

Proof. This is shown in [Cas78, Corollary 1, p. 177] and [Shi10, §24.8 above (24.7a), p. 131], but we give the argument below.

We first show that any $\alpha \in U_0$ can be written as a product of an even number of anisotropic vectors $\vec{v}_i \in V$. To see this, we use the proof of Lemma 3.4.2 to see that $\alpha := \prod_i \vec{v}_i \in (C_0(V))^\times$ and that $\varphi(\alpha \cdot u^{-1}) = \mathrm{id} \in O(V)$. Therefore $\alpha \cdot u^{-1}$ commutes with V, hence is in the center of $C(V)$, and applying Lemma 3.3.3 shows that $u = c\prod_i \vec{v}_i$ for some $c \in K^\times$.

Given that $u \in U_0$ can be written as a product of anisotropic vectors $u = \vec{v}_1 \cdots \vec{v}_r$, the lemma follows from computing $N(u) = \prod_{i=1}^r Q(\vec{v}_i) = sn(u)$ using (3.6). □

From this it follows that the spinor norm of the image of any element of $\mathrm{Spin}(V)$ under φ must be trivial (i.e. $sn(\varphi(\mathrm{Spin}(V))) = (K^\times)^2$), and so the image of $\mathrm{Spin}(V)$ is contained in the **spinor kernel** $\kappa(V) := \ker(sn)$. In fact

Lemma 3.4.5. *The map* $\varphi : \mathrm{Spin}(V) \to SO(V)$ *has image* $\kappa(V)$ *and kernel* $\{\pm 1\}$. *The image* $\mathrm{Im}(\varphi(\mathrm{Spin}(V))) = \kappa(V)$ *since any element* $\alpha \in U_0$ *with* $sn(\alpha) = (K^\times)^2$ *must have* $N(\alpha) = 1$, *and also* $\mathrm{Ker}(\varphi(\mathrm{Spin}(V))) = \{\pm 1\}$.

Proof. The image is $\kappa(V)$ because any $\alpha \in U_0$ with $sn(\alpha) = (K^\times)^2$ must have $N(\alpha) = 1$, and the kernel consists of the elements $c \in F^\times$ with $N(c) = c^2 = 1$. □

We can conveniently summarize our results in the following exact commutative diagram:

$$\begin{tikzcd}
 & \{\pm 1\} \arrow[d, hook] & K^\times \arrow[d, hook] & & \\
1 \arrow[r] & \mathrm{Spin}(V) \arrow[r] \arrow[d, dashed, two heads, "\varphi"] & U_0 \arrow[r, "\mathrm{Norm}"] \arrow[d, two heads, "\varphi"] & K^\times \arrow[r] \arrow[d, "\mathrm{id}"] & 1 \\
1 \arrow[r] & \kappa(V) \arrow[r] & SO(V) \arrow[r, "sn"] & K^\times/(K^\times)^2 &
\end{tikzcd}$$

This shows our main result

Theorem 3.4.6. $\mathrm{Spin}(V)$ *is a double covering of the spinor kernel* $\kappa(V) \subseteq SO(V)$, *and the obstruction to this being a covering map are presence of non-trivial squareclasses of* $K^\times$ *in the image of the spinor norm map* sn.

In the special case where $K^\times$ has only one squareclass (e.g. when K is algebraically closed), we have that

Corollary 3.4.7. *If* $K^\times = (K^\times)^2$, *then* $\mathrm{Spin}(V)$ *is a double cover of* $SO(V)$.

Another interesting special case arises when $K = \mathbb{R}$ and (V, Q) is positive definite. Here all spinor norms are positive, hence they are in the identity squareclass $(\mathbb{R}^\times)^2$, so $\kappa(V) = SO(V)$ and again $\mathrm{Spin}(V)$ is a double cover of $SO(V)$.

3.5 Spinor Equivalence

In Sect. 1.9 we have seen that equivalence of quadratic forms can be viewed as the equivalence of quadratic lattices in a quadratic space (V, Q) by the action of $O(V)$. There are other more refined notions of equivalence that are useful as well. For example, equivalence of quadratic lattices under $SO(V)$ is called **proper equivalence**, and plays an essential role in the theory of binary quadratic forms. In this section we are interested in defining a notion of equivalence called "spinor equivalence" that comes from the action of the spin group $\mathrm{Spin}(V)$ and plays an important role for understanding indefinite quadratic forms in $n \geq 3$ variables.

We say that two quadratic forms over $\mathbb{Z}$ are **locally spinor equivalent** if for every place v their associated quadratic lattices are in the same $\kappa(V_v)$-orbit, where $\kappa(V_v)$ is the local spinor kernel group at v.

In our definition of the genus $\mathrm{Gen}(Q)$ earlier, we saw that it could also be locally realized by the action of a product of local groups $O(V_v)$ giving local isometries, and in Sect. 4.5 we will give a precise adelic version of this statement. One important thing to check is that the local equivalence defining the genus is weaker than the corresponding global equivalence defining classes. This is obvious for the definition of $\mathrm{Gen}(Q)$, but must be forcibly imposed in the case of the spinor genus (because $O(V) \not\subseteq \prod_v \kappa(V_v)$).

Suppose that Q has a corresponding quadratic lattice $L \subset (V, Q)$. Then we define the **spinor genus** of Q, denoted $\mathrm{Spn}(Q)$, to be the set of all quadratic forms Q' whose corresponding lattice $L' \subset (V, Q)$ is locally spinor equivalent to L after performing a global isometry (in $O(V)$). The importance of the spinor genus comes from the following beautiful observation of Eichler that the associated spin groups have a "strong approximation" property, which essentially says that the $\mathbb{Q}$-rational points of the "adelic spin group" are dense in the adelic group $\mathrm{Spin}(V)_{\mathbb{A}}$. While we avoid a more precise statement here, this adelic formulation of algebraic groups will play a central role in Sect. 4.

Theorem 3.5.1 (Eichler [Eic52]). *Suppose that Q is a non-degenerate indefinite quadratic form over $\mathbb{Z}$ in $n \geq 3$ variables. Then there is exactly one class of quadratic forms in its spinor genus.*

Proof. This is proved in [Cas78, Theorem 7.1, p. 186] and [Shi10, Theorem 32.15, p. 192]. □

This theorem allows is to understand statements about indefinite forms in $n \geq 3$ variables by performing various local computations. There is also a (somewhat modified) version of Siegel's Theorem 2.5.1 due to Schulze-Pillot [SP84] that holds for quadratic forms when one averages over Spinor genus $\mathrm{Spn}(Q)$ instead of a genus $\mathrm{Gen}(Q)$. From our computations in Sect. 2.7, we see that this kind of formula gives direct access to the arithmetic of Q when there is only one class in its spinor genus.

4 The Theta Lifting

4.1 Classical to Adelic Modular Forms for GL_2

It is convenient to understand the transformation property (2.2) of modular forms by viewing them as functions on the algebraic group GL_2 with certain invariance properties. We do this in two steps, first by lifting the function $f(z)$ on $\mathcal{H}$ to a function $\tilde{f}$ on $\mathrm{GL}_2(\mathbb{R})$, and then by further lifting this to a function $\mathcal{F}$ on the adelic group $\mathrm{GL}_2(\mathbb{A})$ whose transformation properties can be seen most simply. This adelic perspective will also give us a very flexible language to use to describe the lifting of modular forms via the Weil representation. This passage from classical to adelic modular forms is described in [Gel75b, §3], [Shi97, §10] and [Hid00, §3.1.5].

Given a modular form $f(z) : \mathcal{H} \to \mathbb{C}$ of integral weight $k \in \mathbb{Z}$, level N, Dirichlet character χ and trivial multiplier system as in Definition 2.3.1 (so $f \in M_k(N, \chi)$), we can express its defining transformation property (2.2) as the *invariance property* $(f|_{k,\chi}\gamma)(z) = f(z)$ for all $\gamma \in \Gamma_0(N)$ with respect to the **weight-character slash operator**

$$(f|_{k,\chi}\,\gamma)(z) := f(\gamma\cdot z)(cz+d)^{-k}\chi(d)^{-1}, \qquad \text{where } \gamma = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \Gamma_0(N). \quad (4.1)$$

To create an invariant function on $\mathrm{GL}_2^+(\mathbb{R})$, we first notice that the weight-character slash operator cannot be extended to allow $\gamma \in \mathrm{GL}_2^+(\mathbb{R})$ because the character factor $\chi(d)$ does not make sense in this generality, though the weight factor $(cz+d)^k$ does make sense. However if we ignore the character χ in (4.1) (i.e. take $\chi = 1$ there), then we do get a well-defined **weight slash operator**

$$(f|_k\, g)(z) := f(g\cdot z)(cz+d)^{-k}, \qquad \text{where } g = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \mathrm{GL}_2^+(\mathbb{R}).$$

For these operators the transformation property of $f \in M_k(N, \chi)$ becomes the twisted invariance $(f|_k\gamma)(z) = \chi(d)f(z)$ for all $\gamma \in \Gamma_0(N)$. We can also use this to make a twisted invariant function $\tilde{f}$ on $\mathrm{GL}_2^+(\mathbb{R})$ by noticing that the linear fractional transformation action of $\mathrm{GL}_2^+(\mathbb{R})$ on $\mathcal{H}$ is transitive. By choosing a distinguished point $i \in \mathcal{H}$, we can define $\tilde{f} : \mathrm{GL}_2^+(\mathbb{R}) \to \mathbb{C}$ by

$$\tilde{f}(g) := (f|_k\, g)(i),$$

which satisfies $\tilde{f}(\gamma g) = \chi(d)\tilde{f}(g)$ for all $\gamma \in \Gamma_0(N)$.

At this point we do not yet have a truly invariant function on $\mathrm{GL}_2^+(\mathbb{R})$ unless the character χ is trivial. To incorporate the character χ into our formalism, we need to find a natural place for this "mod N character" to live. The necessary congruence structure is provided by the groups $\mathrm{GL}_2(\mathbb{Q}_p)$ at the non-archimedean (i.e. p-adic) places, and the natural structure combining all of these completions

$\mathrm{GL}_2(\mathbb{Q}_v)$ is the "adelization" $\mathrm{GL}_2(\mathbb{A})$ of the algebraic group GL_2 over $\mathbb{Q}$, defined in the next section. For our purposes in this section we will not be too interested in the role of the character when passing from a classical modular form to an adelic one, however we will be very interested in adelic modular forms in general as the natural "invariant" setting for discussing modular forms.

4.2 Adelizations and Adelic Modular Forms

In this section we give a general definition for adelic modular forms for a general algebraic group G. This agrees with the definitions above when $G = \mathrm{GL}_2$, and in future sections we will want to consider G to be either the symplectic group Sp_{2n} or the special orthogonal group $SO(Q)$ of a definite rational quadratic form Q.

We first define the **adelization of an affine/linear algebraic group** G over the ring of integers $\mathcal{O}$ of a number field F (defined as the zero set of an ideal of relations in a polynomial ring $\mathcal{O}[\vec{x}]$) to be the **restricted direct product** $\prod_v' G(F_v)$ of the local algebraic groups $G(F_v)$ over all places v of F, which is the subset of the usual direct product $\prod_v G(F_v)$ satisfying the restriction that $g_\mathbb{A} = (g_v)_v$ is subject to the restriction that $g_v \in G(\mathcal{O}_v)$ for all but finitely many v. The restricted direct product has several advantages over the usual direct product – it is small enough to be locally compact (since every element has all but finitely many components in the compact group $G(\mathcal{O}_v)$), but it is large enough to contain all rational points $G(F)$.

For the convenience of the reader, we will consistently use subscripts (e.g. $\mathbb{A}, v, \infty,$ $\mathbf{f}$) to denote the kind of element (resp. adelic, local, archimedean, non-archimedean/finite) that the element $g_\bullet \in G(F_\bullet)$ represents. We also denote the center of G by Z, to which the same conventions apply for $z \in Z(F_\bullet)$. Elements without subscripts will represent rational elements (i.e. we take $g \in G(F)$). In most cases $Z(F) = F^\times$ and $Z_\mathbb{A} := Z(F_\mathbb{A})$ are the ideles of F. It is also common to denote the compact groups $G(\mathcal{O}_v)$ as K_v, with K_v denoting a fixed choice of the maximal compact subgroup in $G(F_v)$ when v is archimedean.

In the case where $G = \mathrm{SL}_2$ and $F = \mathbb{Q}$, we can use this notion of an adelic group to further lift a classical modular form $f(z)$ to an "invariant" function on $\mathrm{SL}_2(\mathbb{A})$. To do this we extend the original weight-character slash operator to an operator on $\mathrm{SL}_2(\mathbb{A})$ by writing the Dirichlet character $\chi(d)$ as a product of prime-power characters $\chi_p : \mathbb{Z}_p/p^{\nu_p}\mathbb{Z}_p \cong \mathbb{Z}/p^{\nu_p}\mathbb{Z} \to \mathbb{C}$ where $\nu_p := \mathrm{ord}_p(N)$. These χ_p can be thought of as characters on the p-adic congruence subgroups

$$K_p(N) := \{g = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \mathrm{SL}_2(\mathbb{Z}_p) \mid c \in p^{\nu_p}\mathbb{Z}_p\},$$

and so χ can be thought of as a character on the compact product group $K_\mathbf{f}(N) := \prod_p K_p(N)$ by the formula

$$\chi : (x_p)_{\{p\in\mathbf{f}\}} \longmapsto \prod_p \chi_p(x_p)$$

in which all but finitely many factors $\chi_p(x_p) = 1$. With this reformulation of the Dirichlet character χ, we define the **adelic slash operator** by the formula

$$(f|_{k,\chi,\mathbb{A}}\, g_{\mathbb{A}})(z) := f(g_\infty \cdot z)(c_\infty z + d_\infty)^{-k} \prod_p \chi_p(d_p)^{-1}, \qquad \text{where } g_{\mathbb{A}} = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \mathrm{SL}_2(\mathbb{A}),$$

and notice that for all $\gamma \in \Gamma_0(N)$ (considered as an elements of $\mathrm{SL}_2(\mathbb{A})$ by the canonical diagonal embedding $\gamma \mapsto (\gamma, \gamma, \dots)$) we have the invariance property that

$$(f|_{k,\chi,\mathbb{A}}\, \gamma)(z) = (f|_{k,\chi}\, \gamma)(z) = f(z).$$

We can now lift the classical modular form $f(z) \in M_k(N, \chi)$ to an adelic function $\mathcal{F} : \mathrm{SL}_2(\mathbb{A}) \to \mathbb{C}$ by defining its dependence on $g_{\mathbb{A}}$ through its action on the distinguished point $i \in \mathcal{H}$ as

$$\mathcal{F}(g_{\mathbb{A}}) := (f|_{k,\chi,\mathbb{A}}\, g_{\mathbb{A}})(i). \tag{4.2}$$

One can easily verify that this adelic lift $\mathcal{F}$ satisfies the following important invariance properties:

- $\mathcal{F}(g' g_{\mathbb{A}}) = \mathcal{F}(g_{\mathbb{A}})$ for all $g' \in \mathrm{SL}_2(\mathbb{Q})$,
- $\mathcal{F}(g_{\mathbb{A}} k_{\mathbf{f}}) = \mathcal{F}(g_{\mathbb{A}}) \cdot \chi(k_{\mathbf{f}})$ for all $k_{\mathbf{f}} \in K_{\mathbf{f}}(N)$.

One could work a little harder to show that $\mathcal{F}(g_{\mathbb{A}})$ is an "adelic automorphic form" for the group $G = \mathrm{SL}_2$ in the sense defined below, but for our purposes in these notes the most important properties are the "rational left-invariance" and "right K-finiteness" properties just mentioned. These are the features of adelic automorphic forms that will be most prominent as we perform our explicit theta-lift.

Since our main goal is to move automorphic forms form one group to another, it will be important to have a definition of automorphic forms that is general enough to cover all cases of interest. In general, one defines an **adelic automorphic form** on a linear algebraic group G to be a function $\mathcal{F} : G_{\mathbb{A}} \to \mathbb{C}$ satisfying:

1. $\mathcal{F}$ is left-invariant for the rational group: $\mathcal{F}(g \cdot g_{\mathbb{A}}) = \mathcal{F}(g_{\mathbb{A}})$ for all $g \in G(F)$ and for all $g_{\mathbb{A}} \in G_{\mathbb{A}}$.
2. $\mathcal{F}$ has a central adelic (Hecke) character $\psi : Z_{\mathbb{A}} : Z(\mathbb{A}) \to \mathbb{C}^\times$ so that $\mathcal{F}(z_{\mathbb{A}} \cdot g_{\mathbb{A}}) = \psi(z_{\mathbb{A}}) \cdot \mathcal{F}(g_{\mathbb{A}})$ for all $z_{\mathbb{A}} \in Z_{\mathbb{A}}$ and for all $g_{\mathbb{A}} \in G_{\mathbb{A}}$.
3. $\mathcal{F}$ is right-"$K_{\mathbb{A}}$-finite", meaning that the span of $\mathcal{F}$ as a function under the action of $K_{\mathbb{A}}$ by the right regular representation $k_{\mathbb{A}} : \mathcal{F} \mapsto \tilde{\mathcal{F}}(g_{\mathbb{A}}) := \mathcal{F}(g_{\mathbb{A}} \cdot k_{\mathbf{f}})$ is a finite-dimensional vector space over F.
4. $\mathcal{F}_\infty$ is smooth and "$\mathfrak{z}_\infty$-finite", where $\mathfrak{z}_\infty$ is the center of the universal enveloping algebra for G_∞: meaning that the image of $\mathcal{F}$ under $\mathfrak{z}$ spans a finite-dimensional vector space over F_∞. We note that $\mathfrak{z}_\infty$ can also be interpreted as the ring of bi-invariant differential operators on G_∞, and in the case of $\mathrm{GL}_2(\mathbb{R})$ that $\mathfrak{z}_\infty$ is the ring $\mathbb{C}[\Delta]$ where Δ is the hyperbolic Laplacian operator $-y^2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)$.

5. $\mathcal{F}$ has moderate growth: meaning that there are constants C and $M \in \mathbb{R} > 0$ so that $|\mathcal{F}([\begin{smallmatrix} a & 0 \\ 0 & 1 \end{smallmatrix}]\, g_{\mathbb{A}})| \leq C|a|_{\mathbb{A}}^{M}$ for all $a \in \mathbb{A}_F^{\times}$ with $|a|_{\mathbb{A}} > c$ for some c, and all $g_{\mathbb{A}}$ in any fixed compact subset of $G_{\mathbb{A}}$.

The most important conditions for us will be conditions (1)–(3). Condition (4) is a generalization of the usual holomorphy condition for classical modular forms (since any real-analytic function is an eigenfunction of the Laplacian operator with eigenvalue zero), and condition (5) is a technical growth condition used to exclude poorly behaved functions. In the case where G is an orthogonal group of a definite quadratic form conditions (4) and (5) can be safely omitted because the archimedean component is already compact.

References for this section: For adelizations of algebraic groups and modular forms for an adelic group see [Shi97, §8, 10, 11], for definitions of adelic modular forms on GL_2 [Gel75b, §1.3, pp. 40–53] and [Hid00, §3.1]. Some motivation for this reformulation of the classical language can be found in the brief Corvallis article of Piatetski-Shapiro [PS79]. See also Bump's book [Bum97, §3.1–2] for a discussion of the adelic approach to automorphic forms for the important groups GL_1 and GL_2.

4.3 The Weil Representation

To see how theta functions arise in terms of representation theory, we now define the **Weil representation** whose symmetries will be closely related to the Fourier transform. We will not go through the explicit construction of the Weil representation, but instead content ourselves to list its defining properties below and go on to use the Weil representation to produce classical theta functions. Some references that explicitly construct the Weil representation are [LV80], or Gelbart's book [Gel76]. Other places where the Weil representation is used in a similar way are [Pra93, Pra98, Gel79], Kudla's lecture notes [Kud08], and Gelbart's book, [Gel75a, §7A, pp. 134–150].

These considerations give rise to the adelic Weil Representation $\mathcal{W}$: $\mathrm{Sp}_2(F)\backslash\mathrm{Sp}_2(\mathbb{A}) \to \mathrm{GL}(S(V_{\mathbb{A}}))$ on the space of Schwartz functions on $V_{\mathbb{A}}$, defined by the following transformation formulas:

1. $\left(\mathcal{W}\left(\begin{bmatrix} a & 0 \\ 0 & a^{-1} \end{bmatrix}\right)\Phi\right)(\vec{v}) = \chi_V(a) \cdot |a|_{\mathbb{A}}^{\frac{n}{2}} \cdot \Phi(a\vec{v})$
2. $\left(\mathcal{W}\left(\begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}\right)\Phi\right)(\vec{v}) = e_{\mathbb{A}}(xQ(\vec{v})) \cdot \Phi(\vec{v})$
3. $\left(\mathcal{W}\left(\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\right)\Phi\right)(\vec{v}) = \hat{\Phi}(-\vec{v})$

Here the character $\chi_V(\cdot) := (\cdot, (-1)^{n/2}\det(Q))_F$ and $e_{\mathbb{A}}$ denotes the adelic exponential, defined by

$$e_{\mathbb{A}}((x_v)_v) := e^{2\pi i\, x_\infty} \cdot \prod_p e^{-2\pi i\, \mathrm{Frac}_p(x_p)}$$

where $\mathrm{Frac}_p(x_p) \in \mathbb{Q}/\mathbb{Z}$ is defined as any rational number with p-power denominator for which $x_p \in \mathrm{Frac}_p(x_p) + \mathbb{Z}_p$. (Notice that the **adelic exponential** is always given by a finite product since any adele $x_{\mathbb{A}}$ will have all but finitely many $x_p \in \mathbb{Z}_p$.) The adelic exponential $e_{\mathbb{A}}$ is also used to define the **adelic Fourier transform**

$$\hat{\Phi}(\vec{w}) := \int_{\vec{v} \in V_{\mathbb{A}}} e_{\mathbb{A}}(H(\vec{v}, \vec{w}))\Phi(\vec{v})\, d_{\mathbb{A}}\vec{v}$$

of any Schwartz function $\Phi \in S(V_{\mathbb{A}})$, where the additive Haar measure $d_{\mathbb{A}}\vec{v}$ is normalized so that $\mathrm{Vol}(V_{\mathbb{A}}) = 1$. An explicit reference for these formulas are [Gel76, Theorem 2.22, p. 37]; also [Kud08, I.1.6, p. 3] (though there is a minor typo writing x for a in the second formula there).

These formulas uniquely define the Weil representation, since any element of $\mathrm{Sp}_2(F_{\mathbb{A}})$ can be expressed as a product of these elements (using the Bruhat decomposition for $\mathrm{Sp}_2 = \mathrm{SL}_2$). They are also visibly trivial on elements of $\mathrm{Sp}_2(F)$, because the adelic absolute value $|\cdot|_{\mathbb{A}}$, the rational Hilbert symbol $(\cdot,\cdot)_F := \prod_v (\cdot,\cdot)_{F_v}$, and the adelic exponential $e_{\mathbb{A}}(\cdot)$ are all trivial on rational elements.

Remark 4.3.1. We have not explicitly defined **(adelic) Schwartz functions on** $V_{\mathbb{A}}$, which are just finite linear combinations of an infinite product $\prod_v \Phi_v(x_v)$ of Schwartz functions Φ_v on F_v where Φ_v is the characteristic function of $\mathbb{Z}_p$ at all but finitely many places. For more details, see [Bum97, §3.1, pp. 256–257].

4.4 Theta Kernels and Theta Liftings

For convenience, we now let W denote the non-degenerate 2-dimensional symplectic vector space over F, and identify $\mathrm{SL}_2(F)$ with $\mathrm{Sp}(W)$. The Weil representation restricted to our pair $G \times H := \mathrm{Sp}(W) \times O(V)$ incorporates both an invariance under the orthogonal group, and a Fourier transform from the Weyl element.

To produce theta functions from this, we will need to introduce the familiar classical features of a self-dual function on a lattice. In the adelic context, our lattice is provided by the rational points $V(F)$ and the adelic self-dual function is taken to be the local product $\phi_{\mathbb{A}}(\vec{v}_{\mathbb{A}}) := \prod_v \phi_v(\vec{v}_v)$ of the familiar Gaussian exponential $\phi_\infty(\vec{v}) := e^{-\pi Q(\vec{v})}$ at ∞ and the characteristic function of the completion of some fixed lattice L on V at all non-archimedean places p. (To fix ideas we can take $F = \mathbb{Q}$ and take the standard lattice $L = \mathbb{Z}^n$, giving $\phi_p(\vec{v}) :=$ characteristic function of $\mathbb{Z}_p^n$.)

Finally we must sum the values of our function $\phi_{\mathbb{A}}$ over the rational lattice $V(F)$, which gives rise to a "theta distribution"

$$\theta : \phi \mapsto \sum_{\vec{v} \in V(F)} \phi(\vec{v})$$

on functions $\phi \in S(V)$. We will be interested in the behavior of this distribution under the action of the Weil representation, so we define the **theta kernel**

$$\theta_\phi(g,h) := \sum_{\vec{v}\in V(F)} (\mathcal{W}(g,h)\phi)(\vec{v}) = \sum_{\vec{v}\in V(F)} (\mathcal{W}(g)\phi)(h^{-1}\vec{v})$$

as the value of the theta distribution under this action. This theta kernel already feels very similar to a theta series (since we are summing a quadratic Gaussian over a rational lattice, and our choice $\phi_{\mathbb{A}}$ is supported only on the integral lattice), though it depends on two variables $g \in \mathrm{Sp}_2(\mathbb{A})$ and $h \in O_Q(\mathbb{A})$. However it is more appropriate to think of this as a part of an adelic "theta machine" that will allow us to produce many theta series on $\mathrm{Sp}_2 = \mathrm{SL}_2$ after eliminating the orthogonal variable h in some way.

We now study the rational invariance properties of the theta kernel, which allow us to think of our a priori "adelic" construction as something "automorphic". The main observation is

Lemma 4.4.1. *The theta kernel* $\theta_\phi(g,h)$ *is a function on* $\mathrm{Sp}(W)\backslash\mathrm{Sp}(W_{\mathbb{A}}) \times O(V)\backslash O(V_{\mathbb{A}})$.

Proof. At the end of the previous section we noted that the Weil representation transformation 2 is invariant when $x \in F$. Transformation 1 with $a \in F^\times$ performs a rational scaling of the values, which leaves the rational lattice $V(F)$ invariant, and transformation 3 preserves the theta kernel because the (adelic) Poisson summation formula tells us that the sum of a function on the standard lattice is the same as the sum using its Fourier transform. Therefore the theta kernel has the rational lattice symmetries of being left-invariant under $\mathrm{Sp}(W)$. It is also left-invariant under $O(V(F))$ because that action just permutes $V(F)$. Therefore we have shown that the theta kernel is rationally left-invariant, and so it descends to a function on the rational left cosets as desired. □

This rational bi-invariance is exactly what allows us to use the theta kernel to move automorphic forms between the orthogonal and symplectic groups. Given an automorphic form $F(h_{\mathbb{A}})$ on the orthogonal group $O(V_{\mathbb{A}})$, we can define its **theta lift** by the integral

$$(\Theta(F))(g_{\mathbb{A}}) := \int_{h_{\mathbb{A}}\in O(V)\backslash O(V_{\mathbb{A}})} F(g_{\mathbb{A}})\, \theta_\phi(g_{\mathbb{A}}, h_{\mathbb{A}})\, dh_{\mathbb{A}} \qquad (4.3)$$

of F against the theta kernel with respect to the choice of adelic Haar measure $dh_{\mathbb{A}}$ on $O(V_{\mathbb{A}})$ giving the adelic stabilizer $\mathrm{Stab}_{\mathbb{A}}(L) \subset O(V_{\mathbb{A}})$ volume 1. The theta lift $\Theta(F)(g_{\mathbb{A}})$ formally inherits the symplectic invariance of the theta kernel, and gives an automorphic form on $\mathrm{Sp}(W_{\mathbb{A}})$ when the integral converges. In the next few sections we will give an explicit example of how this process can be used to produce the classical theta series of a positive definite integer-valued quadratic form.

4.5 Some Simple Automorphic Forms on the Orthogonal Group

To actually use the theta lift, we must have at our disposal a supply of automorphic forms on the orthogonal quotient $O(V)\backslash O(V_{\mathbb{A}})$. In this section we describe the simplest of these, characteristic functions of a point, which are surprisingly useful for our purposes.

We begin by giving a classical interpretation of the orthogonal quotient $O(V)\backslash O(V_{\mathbb{A}})$. Given the rational quadratic space (V, Q), we define an action of the adelic orthogonal group $O(V_{\mathbb{A}})$ on the set of all (quadratic) $\mathcal{O}_F$-lattices in (V, Q) by using the following local-global statement for lattices in a rational vector space.

Lemma 4.5.1. *There is a natural bijection between lattices L in an n-dimensional F-rational vector space V and the tuples $(L_p)_p$ of local lattices $L_p \subset V_p := V \otimes_F F_p$ satisfying the property that all but finitely many L_p are equal to $\mathcal{O}_p^n$. (Here p runs over the set of (non-zero) primes of F.)*

Proof. This is proved in [Shi10, Lemma 21.6, pp. 102–103] and [Wei67, Theorem 2, p. 84]. The relevant maps in each direction are

$$L \mapsto (L_p := L \otimes_{\mathcal{O}_F} \mathcal{O}_p)_p \qquad \text{and} \qquad (L_p)_p \mapsto L := \bigcap_p (V \cap L_p). \qquad \square$$

With this lemma, we define an action of $h_{\mathbb{A}} \in O(V_{\mathbb{A}})$ on the lattices in (V, Q) by acting locally on the associated tuple of local lattices:

$$h_{\mathbb{A}} : L \longmapsto (h_p L_p)_p \overset{Lemma}{\longleftrightarrow} h_{\mathbb{A}} L.$$

This action produces a new tuple of local lattices, which differs from the first tuple at only finitely many places (by the restricted direct product condition on $O(V_{\mathbb{A}})$), and so corresponds to a unique lattice in (V, Q). Notice that this action makes no use of the non-archimedian component h_∞ of $h_{\mathbb{A}}$.

We now fix a lattice L in (V, Q), and interpret the action of $O(V_{\mathbb{A}})$ on L classically.

Lemma 4.5.2. *The orbit of L under $O(V_{\mathbb{A}})$ is the genus of L.*

Proof. From the definition of the action, we see that the new lattice $L' := h_A L$ is locally isometric to the lattice L at all primes p, so it is in the genus of L. Since $L' \subset (V, Q)$ we see they are also isometric at the archimedian place ∞, hence $L' \in \mathrm{Gen}(L)$.

Similarly, any lattice in $\mathrm{Gen}(L)$ can be realized as $h_{\mathbb{A}} L$ by taking h_p to be the element of the orthogonal group carrying L_p to L'_p at the finitely many primes where $L_p \neq L'_p$, and taking all other components h_v as the identity. $\square$

This interpretation can be extended a little further, by trying to describe the classes in the genus $\mathrm{Gen}(L)$ adelically. If we define the adelic stabilizer

$$K_{\mathbb{A}} := \mathrm{Stab}_{\mathbb{A}}(L) := \{h_{\mathbb{A}} \in O(V_{\mathbb{A}}) \mid h_A L = L\}$$

then we have a bijection

$$\begin{array}{ccc} O(V_{\mathbb{A}})/\mathrm{Stab}_{\mathbb{A}}(L) & \overset{1-1}{\longleftrightarrow} & \mathrm{Gen}(L) \\ h_A & \longmapsto & h_A L \end{array}$$

Taking this one step further, we have the important bijection

$$O(V)\backslash O(V_{\mathbb{A}})/\mathrm{Stab}_{\mathbb{A}}(L) \quad \overset{1-1}{\longleftrightarrow} \quad \text{classes in } \mathrm{Gen}(L)$$

because two lattices in (V, Q) are in the same class if they are isometric, hence they differ by the action of an element of $O(V)$.

This finite quotient $O(V)\backslash O(V_{\mathbb{A}})/\mathrm{Stab}_{\mathbb{A}}(L)$ corresponding to the classes in the genus can also be thought of as the analogue of the usual upper half-plane $\mathcal{H}$ (for SL_2) for the orthogonal group $O(V)$. Under this analogy, we see that the analogue of modular functions for $O(V)$ are just functions on this finite set of points (labelled by the classes L_i in $\mathrm{Gen}(L)$). The simplest of these are the characteristic functions Φ_{L_i} of each point, and the constant function 1, both of which we will see play a special role the theory of theta series. Since the non-archimedian part of $\mathrm{Stab}_{\mathbb{A}}(L)$ is a compact group, we can easily verify that both of these functions are adelic automorphic forms in the sense of Sect. 4.2.

4.6 Realizing Classical Theta Functions as Theta Lifts

We now compute the theta lifting of the characteristic function Φ_{L_j} of the double coset of $O(V_{\mathbb{A}})$ corresponding to a chosen quadratic lattice $L_j \in \mathrm{Gen}(L)$, with respect to the fixed choice of function $\phi_{\mathbb{A}}(\vec{v}_{\mathbb{A}})$ described above. We will see that this an adelic automorphic form that classically corresponds to a certain multiple of the familiar theta series

$$\Theta_{L_j}(z) := \sum_{\vec{v}\in L_j} e^{2\pi i Q(\vec{v})z} = \sum_{m\in\mathbb{Z}\geq 0} r_{L_j}(m)\, e^{2\pi i m z}.$$

By our definition of the theta lift in (4.3), we are interested in computing

$$\Theta(\Phi_{L_j})(g_{\mathbb{A}}) = \int_{O(V)\backslash O(V_{\mathbb{A}})} \Phi_{L_j}(h_{\mathbb{A}})\, \theta_{\phi}(g_{\mathbb{A}}, h_{\mathbb{A}})\, dh_{\mathbb{A}} \tag{4.4}$$

as an automorphic form on $\mathrm{SL}_2 = \mathrm{Sp}_2$. To do this we first decompose $O(V_{\mathbb{A}})$ as a union of double cosets corresponding to the classes in the genus of L (i.e. with respect to the adelic stabilizer $K_{\mathbb{A}}$), giving

$$O(V_{\mathbb{A}}) = \bigsqcup_{i \in I} O(V_F)\, \alpha_{i,\mathbb{A}}\, K_{\mathbb{A}}$$

for some fixed choice of representatives $\alpha_i := \alpha_{i,\mathbb{A}} \in O(V_{\mathbb{A}})$ where $\mathrm{Gen}(L) = \bigsqcup_{i \in I} \mathrm{Cls}(L_i)$ and $L_i = \alpha_i L$. Since the action of $O(V_{\mathbb{A}})$ on lattices only depends on the non-archimedean components of $\alpha_{i,\mathbb{A}}$, to simplify our lives we choose the $\alpha_{i,\mathbb{A}}$ to have trivial archimedean components $\alpha_{i,\infty} = 1 \in O(V_\infty)$. It will also be convenient to define the adelic stabilizers $K_{i,\mathbb{A}} := \mathrm{Stab}_{\mathbb{A}}(L_i)$ of the other lattices $L_i \in \mathrm{Gen}(L)$. With this we compute

$$\begin{aligned}
\Theta(\Phi_{L_j})(g_{\mathbb{A}}) &= \int_{O(V_F)\backslash O(V_{\mathbb{A}})} \Phi_{L_j}(h_{\mathbb{A}})\, \theta_\phi(g_{\mathbb{A}}, h_{\mathbb{A}})\, dh_{\mathbb{A}} && (4.5)\\
&= \int_{O(V_F)\backslash \bigsqcup_{i \in I} O(V_F)\alpha_i K_{\mathbb{A}}} \Phi_{L_j}(h_{\mathbb{A}})\, \theta_\phi(g_{\mathbb{A}}, h_{\mathbb{A}})\, dh_{\mathbb{A}} && (4.6)\\
&= \sum_{i \in I} \int_{O(V_F)\backslash O(V_F)\alpha_i K_{\mathbb{A}}} \Phi_{L_j}(h_{\mathbb{A}})\, \theta_\phi(g_{\mathbb{A}}, h_{\mathbb{A}})\, dh_{\mathbb{A}} && (4.7)\\
&= \sum_{i \in I} \int_{O(V_F)\backslash O(V_F)\alpha_i K_{\mathbb{A}}\alpha_i^{-1}} \Phi_{L_j}(h_{\mathbb{A}}\alpha_i)\, \theta_\phi(g_{\mathbb{A}}, h_{\mathbb{A}}\alpha_i)\, dh_{\mathbb{A}} && (4.8)\\
&= \sum_{i \in I} \int_{O(V_F)\backslash O(V_F)K_{i,\mathbb{A}}} \Phi_{L_j}(h_{\mathbb{A}}\alpha_i)\, \theta_\phi(g_{\mathbb{A}}, h_{\mathbb{A}}\alpha_i)\, dh_{\mathbb{A}} && (4.9)\\
&= \sum_{i \in I} \int_{(O(V_F)\cap K_{i,\mathbb{A}})\backslash K_{i,\mathbb{A}}} \Phi_{L_j}(h_{\mathbb{A}}\alpha_i)\, \theta_\phi(g_{\mathbb{A}}, h_{\mathbb{A}}\alpha_i)\, dh_{\mathbb{A}} && (4.10)\\
&= \sum_{i \in I} \frac{1}{|\mathrm{Aut}(L_i)|} \int_{K_{i,\mathbb{A}}} \Phi_{L_j}(h_{\mathbb{A}}\alpha_i)\, \theta_\phi(g_{\mathbb{A}}, h_{\mathbb{A}}\alpha_i)\, dh_{\mathbb{A}}, && (4.11)
\end{aligned}$$

where the last step follows because $O(V_F) \cap K_{i,\mathbb{A}}$ is the finite group of rational automorphisms $\mathrm{Aut}(L_i)$, since $K_{i,\mathbb{A}}$ is the adelic stabilizer of L_i. (Note: We are also implicitly using the invariance of the left Haar measure $dh_{\mathbb{A}}$ under right multiplication, because the orthogonal group is unimodular. See [Wei82, p. 23], [Vos98, §14.4, p. 137] and [Ono66, §2, p. 123] for a justification of this bi-invariance.)

At this point we have "unfolded" our integral to the point where it factors as a product of local integrals (since $K_{i,\mathbb{A}}$ is the product of the local stabilizers $K_{i,v}$ over all places v), each of which we can try to evaluate separately. We first notice that for each summand we have an integral over $K_{i,\mathbb{A}} = \alpha_i K_{\mathbb{A}} \alpha_i^{-1}$, giving

$$\Theta(\Phi_{L_j})(g_{\mathbb{A}}) = \sum_{i\in I} \frac{1}{|\mathrm{Aut}(L_i)|} \int_{K_{i,\mathbb{A}}} \Phi_{L_j}(h_{\mathbb{A}}\alpha_i)\,\theta_\phi(g_{\mathbb{A}}, h_{\mathbb{A}}\alpha_i)\,dh_{\mathbb{A}} \tag{4.12}$$

$$= \sum_{i\in I} \frac{1}{|\mathrm{Aut}(L_i)|} \int_{K_{\mathbb{A}}} \Phi_{L_j}(\alpha_i h_{\mathbb{A}})\,\theta_\phi(g_{\mathbb{A}}, \alpha_i h_{\mathbb{A}})\,dh_{\mathbb{A}} \tag{4.13}$$

whose integrals do not depend on $i \in I$. To analyze the internal integral, notice that $h_{\mathbb{A}} \in K_{\mathbb{A}}$, giving that

$$\Phi_{L_j}(\alpha_i h_{\mathbb{A}}) \neq 0 \iff \alpha_i h_{\mathbb{A}} \in O(V_F)\cdot \alpha_j \cdot K_{\mathbb{A}} \tag{4.14}$$

$$\iff \alpha_i \in O(V_F)\cdot \alpha_j \cdot K_{\mathbb{A}} \tag{4.15}$$

$$\iff \alpha_i = \alpha_j \tag{4.16}$$

and so all terms with $\alpha_i \neq \alpha_j$ vanish. Thus

$$\Theta(\Phi_{L_j})(g_{\mathbb{A}}) = \frac{1}{|\mathrm{Aut}(L_j)|} \int_{h_{\mathbb{A}}\in K_{\mathbb{A}}} \theta_\phi(g_{\mathbb{A}}, \alpha_j h_{\mathbb{A}})\,dh_{\mathbb{A}} \tag{4.17}$$

At this point we have to unwind the theta kernel to evaluate the integral, now heavily using the fact that this is a product of local integrals. We can simplify the non-archimedean orthogonal action with the observation that

$$h_p^{-1}\alpha_j^{-1}\vec{v} \in L_p \text{ for all primes } p \iff \vec{v} \in \alpha_j h_p L_p = L_{j,p} \text{ for all } p \tag{4.18}$$

$$\iff \vec{v} \in L_j. \tag{4.19}$$

This together with our choice that $\alpha_{j,\infty} = 1 \in O(V_\infty)$ gives

$$\Theta(\Phi_{L_j})(g_{\mathbb{A}}) = \frac{1}{|\mathrm{Aut}(L_j)|} \int_{h_{\mathbb{A}}\in K_{\mathbb{A}}} \theta_\phi(g_{\mathbb{A}}, \alpha_j h_{\mathbb{A}})\,dh_{\mathbb{A}} \tag{4.20}$$

$$= \frac{1}{|\mathrm{Aut}(L_j)|} \int_{h_{\mathbb{A}}\in K_{\mathbb{A}}} \sum_{\vec{v}\in V_F} (\mathcal{W}(g_{\mathbb{A}})\phi_{\mathbb{A}})(h_{\mathbb{A}}^{-1}\alpha_j^{-1}\vec{v})\,dh_{\mathbb{A}} \tag{4.21}$$

$$= \frac{1}{|\mathrm{Aut}(L_j)|} \int_{h_{\mathbb{A}}\in K_{\mathbb{A}}} \sum_{\vec{v}\in L_j} (\mathcal{W}(g_{\mathbb{A}})\phi_{\mathbb{A}})(h_\infty^{-1}\vec{v})\,dh_{\mathbb{A}}. \tag{4.22}$$

To understand the non-archimedean symplectic action we take advantage of the invariance of the theta lift under $\mathrm{Sp}(W_F)$ by invoking the "strong approximation" property of symplectic groups, (a special case of) which states that

$$\mathrm{Sp}(W_{\mathbb{A}}) = \mathrm{Sp}(W)\cdot \mathrm{Sp}(F_\infty) \prod_p \mathrm{Sp}(\mathcal{O}_p).$$

This means that we can adjust the element $g_{\mathbb{A}}$ (by left-multiplying with some element $g_F \in \mathrm{Sp}(W)$) so that its new local components g_p live in $\mathrm{Sp}(\mathcal{O}_p)$ for all primes p. By using the transformation formulas of the Weil representation we see that each component g_p acts trivially on the characteristic function $\phi_p(\vec{v})$ of L_p. Thus we can express our theta lift as depending only on the archimedean component g_∞ of $g_{\mathbb{A}}$, giving

$$\Theta(\Phi_{L_j})(g_{\mathbb{A}}) = \frac{1}{|\mathrm{Aut}(L_j)|} \int_{h_{\mathbb{A}} \in K_{\mathbb{A}}} \sum_{\vec{v} \in L_j} (\mathcal{W}(g_\infty)\phi_\infty)(h_\infty^{-1}\vec{v})\, dh_{\mathbb{A}}. \tag{4.23}$$

Since we are interested in the classical modular form $f(z)$ on $\mathcal{H}$ associated the adelic modular form $\Theta(\Phi_L)$, we need only evaluate this on elements $g_\infty \in \mathrm{SL}_2(\mathbb{R})$ for which $g_\infty \cdot i = z \in \mathcal{H}$. We notice that when $x, y \in \mathbb{R}$ with $y > 0$, the elements

$$g_{\infty,z} := \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{y} & 0 \\ 0 & \sqrt{y}^{-1} \end{bmatrix}$$

satisfy $g_{\infty,z} \cdot i = x + iy \in \mathcal{H}$. For these elements $g_{\infty,z}$, the action of the Weil representation in (4.23) can be written more explicitly as

$$\begin{aligned}
(\mathcal{W}(g_{\infty,z})\phi_\infty)(h_\infty^{-1}\vec{v}) &= \left(\mathcal{W}\left(\begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}\right) \mathcal{W}\left(\begin{bmatrix} \sqrt{y} & 0 \\ 0 & \sqrt{y}^{-1} \end{bmatrix}\right) \phi_\infty\right)(h_\infty^{-1}\vec{v}) && (4.24)\\
&= y^{\frac{n}{4}} \left(\mathcal{W}\left(\begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}\right) \phi_\infty\right)(\sqrt{y}\, h_\infty^{-1}\vec{v}) && (4.25)\\
&= y^{\frac{n}{4}} e^{2\pi i x Q(\vec{v})} \phi_\infty(\sqrt{y}\, h_\infty^{-1}\vec{v}) && (4.26)\\
&= y^{\frac{n}{4}} e^{2\pi i x Q(\vec{v})} e^{-2\pi Q(\sqrt{y}\, h_\infty^{-1}\vec{v})} && (4.27)\\
&= y^{\frac{n}{4}} e^{2\pi i x Q(\vec{v})} e^{2\pi i \cdot i y Q(h_\infty^{-1}\vec{v})} && (4.28)\\
&= y^{\frac{n}{4}} e^{2\pi i x Q(\vec{v})} e^{2\pi i \cdot i y Q(\vec{v})} && (4.29)\\
&= y^{\frac{n}{4}} e^{2\pi i z Q(\vec{v})}. && (4.30)
\end{aligned}$$

Substituting this back into (4.23) gives

$$\begin{aligned}
\Theta(\Phi_{L_j})(g_{\infty,z}) &= \frac{1}{|\mathrm{Aut}(L_j)|} \int_{h_{\mathbb{A}} \in K_{\mathbb{A}}} \sum_{\vec{v} \in L_j} (\mathcal{W}(g_{\infty,z})\phi_\infty)(h_\infty^{-1}\vec{v})\, dh_{\mathbb{A}} && (4.31)\\
&= \frac{1}{|\mathrm{Aut}(L_j)|} \int_{h_{\mathbb{A}} \in K_{\mathbb{A}}} \sum_{\vec{v} \in L_j} y^{\frac{n}{4}} e^{2\pi i Q(\vec{v}) z}\, dh_{\mathbb{A}} && (4.32)
\end{aligned}$$

$$= \frac{\mathrm{Vol}_{\mathbb{A}}(K_{\mathbb{A}})}{|\mathrm{Aut}(L_j)|} \sum_{\vec{v} \in L_j} y^{\frac{n}{4}} e^{2\pi i Q(\vec{v}) z} \tag{4.33}$$

$$= \frac{1}{|\mathrm{Aut}(L_j)|} \sum_{\vec{v} \in L_j} y^{\frac{n}{4}} e^{2\pi i Q(\vec{v}) z}. \tag{4.34}$$

Now using the relation (4.2) with $k = n/2$ and trivial Dirichlet character χ we have $g_{\infty,z}$ has $(cz+d)^k = y^{-k/2}$ and can see that $\Theta(\Phi_{L_j})$ corresponds to the classical weight k modular form

$$f(z) := \chi(d)(cz+d)^k \cdot \Theta(\Phi_{L_j})(g_{\infty,z}) \tag{4.35}$$

$$= y^{-n/4} \cdot \Theta(\Phi_{L_j})(g_{\infty,z}) \tag{4.36}$$

$$= \frac{1}{|\mathrm{Aut}(L_j)|} \sum_{\vec{v} \in L_j} e^{2\pi i Q(\vec{v}) z}. \tag{4.37}$$

But this is just the usual theta series $\Theta_{L_j}(z)$ weighed by the rational factor $\frac{1}{|\mathrm{Aut}(L_j)|}$, so we have indeed recovered the classical theta function as the theta lift of the characteristic function of the double coset $O(V)\,\alpha_{j,\mathbb{A}}\,K_{\mathbb{A}}$ of the adelic orthogonal group corresponding to the lattice $L_j \in \mathrm{Gen}(L)$.

References

[:1959] Correspondence [signed “R. Lipschitz”]. *Ann. of Math. (2)*, 69:247–251, 1959. Attributed to A. Weil.

[AZ95] Anatolii Nikolaevich Andrianov and Vladimir Georgievich Zhuravlëv. *Modular forms and Hecke operators*, volume 145 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1995. Translated from the 1990 Russian original by Neal Koblitz.

[Bak81] Anthony Bak. *K-theory of forms*, volume 98 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, N.J., 1981.

[BH83] James William Benham and John Sollion Hsia Spinor equivalence of quadratic forms. *J. Number Theory*, 17(3):337–342, 1983.

[Bha] Manjul Bhargava. 2009 Arizona Winter School lecture notes on “The parametrization of rings of small rank”.

[Bum97] Daniel Bump. *Automorphic forms and representations*, volume 55 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1997.

[Cas78] John William Scott Cassels. *Rational quadratic forms*, volume 13 of *London Mathematical Society Monographs*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1978.

[DS05] Fred Diamond and Jerry Shurman. *A first course in modular forms*, volume 228 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2005.

[DSP90] William Duke and Rainer Schulze-Pillot. Representation of integers by positive ternary quadratic forms and equidistribution of lattice points on ellipsoids. *Invent. Math.*, 99(1):49–57, 1990.

[Duk97] William Duke. Some old problems and new results about quadratic forms. *Notices Amer. Math. Soc.*, 44(2):190–196, 1997.

[Eic52] Martin Eichler. Die Ähnlichkeitsklassen indefiniter Gitter. *Math. Z.*, 55:216–252, 1952.

[Eic66] Martin Eichler. *Introduction to the theory of algebraic numbers and functions*. Translated from the German by George Striker. Pure and Applied Mathematics, Vol. 23. Academic Press, New York, 1966.

[EKM08] Richard Elman, Nikita Karpenko, and Alexander Merkurjev. *The algebraic and geometric theory of quadratic forms*, volume 56 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2008.

[Eve76] Howard Eves. *An introduction to the history of mathematics*. Holt, Rinehart and Winston, fourth edition, 1976.

[Gel75a] Stephen Gelbart. *Automorphic forms and representations of adele groups*. Department of Mathematics, University of Chicago, Chicago, Ill., 1975. Lecture Notes in Representation Theory.

[Gel75b] Stephen S. Gelbart. *Automorphic forms on adèle groups*. Princeton University Press, Princeton, N.J., 1975. Annals of Mathematics Studies, No. 83.

[Gel76] Stephen S. Gelbart. *Weil's representation and the spectrum of the metaplectic group*. Lecture Notes in Mathematics, Vol. 530. Springer-Verlag, Berlin, 1976.

[Gel79] Stephen Gelbart. Examples of dual reductive pairs. In *Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 1*, Proc. Sympos. Pure Math., XXXIII, pages 287–296. Amer. Math. Soc., Providence, R.I., 1979.

[Ger08] Larry J. Gerstein. *Basic quadratic forms*, volume 90 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2008.

[GS06] Philippe Gille and Tamás Szamuely. *Central simple algebras and Galois cohomology*, volume 101 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006.

[Han04] Jonathan Hanke. Some recent results about (ternary) quadratic forms. In *Number theory*, volume 36 of *CRM Proc. Lecture Notes*, pages 147–164. Amer. Math. Soc., Providence, RI, 2004.

[Hid00] Haruzo Hida. *Modular forms and Galois cohomology*, volume 69 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2000.

[IK04] Henryk Iwaniec and Emmanuel Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[Iwa87] Henryk Iwaniec. Spectral theory of automorphic functions and recent developments in analytic number theory. In *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Berkeley, Calif., 1986)*, pages 444–456, Providence, RI, 1987. Amer. Math. Soc.

[Iwa97] Henryk Iwaniec. *Topics in classical automorphic forms*, volume 17 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1997.

[Jac] Carl Gustav Jacob Jacobi. *Fundamenta nova theoriae functionum ellipticarum.* Königsberg, 1829. In Latin. Reprinted with corrections in: Carl Gustav Jacob Jacobi. Gesammelte Werke. 8 volumes. Berlin, 1881–1891. 1. 49–239. reprinted new york (chelsea, 1969) and available from the american mathematical society. edition.

[Jac89] Nathan Jacobson. *Basic algebra. II.* W. H. Freeman and Company, New York, second edition, 1989.

[Kap03] Irving Kaplansky. *Linear algebra and geometry*. Dover Publications Inc., Mineola, NY, revised edition, 2003. A second course.

[Kne66] Martin Kneser. Strong approximation. In *Algebraic Groups and Discontinuous Subgroups (Proc. Sympos. Pure Math., Boulder, Colo., 1965)*, pages 187–196. Amer. Math. Soc., Providence, R.I., 1966.

[Kno70] Marvin I. Knopp. *Modular functions in analytic number theory*. Markham Publishing Co., Chicago, Ill., 1970.

[Knu91] Max-Albert Knus. *Quadratic and Hermitian forms over rings*, volume 294 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1991. With a foreword by I. Bertuccioni.

[Kob93] Neal Koblitz. *Introduction to elliptic curves and modular forms*, volume 97 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1993.

[KR88a] Stephen S. Kudla and Stephen Rallis. On the Weil-Siegel formula. *J. Reine Angew. Math.*, 387:1–68, 1988.

[KR88b] Stephen S. Kudla and Stephen Rallis. On the Weil-Siegel formula. II. The isotropic convergent case. *J. Reine Angew. Math.*, 391:65–84, 1988.

[Kud08] Stephen S. Kudla. Some extensions of the Siegel-Weil formula. In *Eisenstein series and applications*, volume 258 of *Progr. Math.*, pages 205–237. Birkhäuser Boston, Boston, MA, 2008.

[Lam05] Tsit Yuen Lam. *Introduction to quadratic forms over fields*, volume 67 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2005.

[Lan94] Serge Lang. *Algebraic number theory*, volume 110 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1994.

[Lan95] Serge Lang. *Algebra*. Addison-Wesley Publishing Company, Inc., Reading, MA, third edition, 1995.

[LV80] Gérard Lion and Michèle Vergne. *The Weil representation, Maslov index and theta series*, volume 6 of *Progress in Mathematics*. Birkhäuser Boston, Mass., 1980.

[Miy06] Toshitsune Miyake. *Modular forms*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, english edition, 2006. Translated from the 1976 Japanese original by Yoshitaka Maeda.

[MW06] Carlos J. Moreno and Samuel S. Wagstaff, Jr. *Sums of squares of integers*. Discrete Mathematics and its Applications (Boca Raton). Chapman & Hall/CRC, Boca Raton, FL, 2006.

[Neu99] Jürgen Neukirch. *Algebraic number theory*, volume 322 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999. Translated from the 1992 German original and with a note by Norbert Schappacher, With a foreword by G. Harder.

[NSW08] Jürgen Neukirch, Alexander Schmidt, and Kay Wingberg. *Cohomology of number fields*, volume 323 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 2008.

[O'M71] Onorato Timothy O'Meara. *Introduction to quadratic forms*. Springer-Verlag, New York, 1971. Second printing, corrected, Die Grundlehren der mathematischen Wissenschaften, Band 117.

[Ono66] Takashi Ono. On Tamagawa numbers. In *Algebraic Groups and Discontinuous Subgroups (Proc. Sympos. Pure Math., Boulder, Colo., 1965)*, pages 122–132. Amer. Math. Soc., Providence, R.I., 1966.

[Par] Raman Parimala. 2009 Arizona Winter School lecture notes on "Some aspects of the algebraic theory of quadratic forms".

[Pfe71] Horst Pfeuffer. Einklassige Geschlechter totalpositiver quadratischer Formen in totalreellen algebraischen Zahlkörpern. *J. Number Theory*, 3:371–411, 1971.

[Pfe78] Horst Pfeuffer. Darstellungsmasse binärer quadratischer Formen über totalreellen algebraischen Zahlkörpern. *Acta Arith.*, 34(2):103–111, 1977/78.

[Pie82] Richard S. Pierce. *Associative algebras*, volume 88 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1982. Studies in the History of Modern Science, 9.

[PR94] Vladimir Platonov and Andrei Rapinchuk. *Algebraic groups and number theory*, volume 139 of *Pure and Applied Mathematics*. Academic Press Inc., Boston, MA, 1994. Translated from the 1991 Russian original by Rachel Rowen.

[Pra93] Dipendra Prasad. Weil representation, Howe duality, and the theta correspondence. In *Theta functions: from the classical to the modern*, volume 1 of *CRM Proc. Lecture Notes*, pages 105–127. Amer. Math. Soc., Providence, RI, 1993.

[Pra98] Dipendra Prasad. A brief survey on the theta correspondence. In *Number theory (Tiruchirapalli, 1996)*, volume 210 of *Contemp. Math.*, pages 171–193. Amer. Math. Soc., Providence, RI, 1998.

[PS79] Ilya I. Piatetski-Shapiro. Classical and adelic automorphic forms. An introduction. In *Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 1*, Proc. Sympos. Pure Math., XXXIII, pages 185–188. Amer. Math. Soc., Providence, R.I., 1979.

[Sah60] Chih-han Sah. Quadratic forms over fields of characteristic 2. *Amer. J. Math.*, 82:812–830, 1960.

[Sal99] David J. Saltman. *Lectures on division algebras*, volume 94 of *CBMS Regional Conference Series in Mathematics*. Published by American Mathematical Society, Providence, RI, 1999.

[Ser77] Jean-Pierre Serre. Modular forms of weight one and Galois representations. In *Algebraic number fields: L-functions and Galois properties (Proc. Sympos., Univ. Durham, Durham, 1975)*, pages 193–268. Academic Press, London, 1977.

[SH98] Rudolf Scharlau and Boris Hemkemeier. Classification of integral lattices with large class number. *Math. Comp.*, 67(222):737–749, 1998.

[Shi73] Goro Shimura. On modular forms of half integral weight. *Ann. of Math. (2)*, 97:440–481, 1973.

[Shi94] Goro Shimura. *Introduction to the arithmetic theory of automorphic functions*, volume 11 of *Publications of the Mathematical Society of Japan*. Princeton University Press, Princeton, NJ, 1994. Reprint of the 1971 original, Kano Memorial Lectures, 1.

[Shi97] Goro Shimura. *Euler products and Eisenstein series*, volume 93 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1997.

[Shi04] Goro Shimura. *Arithmetic and analytic theories of quadratic forms and Clifford groups*, volume 109 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2004.

[Shi06a] Goro Shimura. Integer-valued quadratic forms and quadratic Diophantine equations. *Doc. Math.*, 11:333–367 (electronic), 2006.

[Shi06b] Goro Shimura. Quadratic Diophantine equations, the class number, and the mass formula. *Bull. Amer. Math. Soc. (N.S.)*, 43(3):285–304 (electronic), 2006.

[Shi10] Goro Shimura. *Arithmetic of quadratic forms*. Springer Monographs in Mathematics. Springer, New York, 2010.

[Sie35] Carl Ludwig Siegel. Über die analytische Theorie der quadratischen Formen. *Ann. of Math. (2)*, 36(3):527–606, 1935.

[Sie36] Carl Ludwig Siegel. Über die analytische Theorie der quadratischen Formen. II. *Ann. of Math. (2)*, 37(1):230–263, 1936.

[Sie37] Carl Ludwig Siegel. Über die analytische Theorie der quadratischen Formen. III. *Ann. of Math. (2)*, 38(1):212–291, 1937.

[Sie63] Carl Ludwig Siegel. *Lectures on the analytical theory of quadratic forms*. Notes by Morgan Ward. Third revised edition. Buchhandlung Robert Peppmüller, Göttingen, 1963.

[SP84] Rainer Schulze-Pillot. Darstellungsmaße von Spinorgeschlechtern ternärer quadratischer Formen. *J. Reine Angew. Math.*, 352:114–132, 1984.

[SP00] Rainer Schulze-Pillot. Exceptional integers for genera of integral ternary positive definite quadratic forms. *Duke Math. J.*, 102(2):351–357, 2000.

[SP04] Rainer Schulze-Pillot. Representation by integral quadratic forms—a survey. In *Algebraic and arithmetic theory of quadratic forms*, volume 344 of *Contemp. Math.*, pages 303–321. Amer. Math. Soc., Providence, RI, 2004.

[Str09] Gilbert Strang. *Introduction to Linear Algebra*. Wellesley Cambridge Press, New York, fourth edition, 2009.

[Tar29] W. Tartakowsky. Die gesamtheit der zahlen, die durch eine positive quadratische form $f(x_1, x_2, \ldots, x_s)$ $(s \geq 4)$ darstellbar sind. i, ii. *Bull. Ac. Sc. Leningrad*, 2(7):111–122; 165–196, 1929.

[Tor05] Gonzalo Tornaria. *The Brandt module of ternary quadratic lattices*. PhD thesis, University of Texas, Austin, 2005.

[Voi] John Voight. Computing with quaternion algebras: Identifying the matrix ring.

[Vos98] Valentin Evgenévich Voskresenskiĭ. *Algebraic groups and their birational invariants*, volume 179 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1998. Translated from the Russian manuscript by Boris Kunyavski [Boris È. Kunyavskiĭ].

[Wat63] G. L. Watson. One-class genera of positive quadratic forms. *J. London Math. Soc.*, 38:387–392, 1963.

[Wat84] G. L. Watson. One-class genera of positive quadratic forms in seven variables. *Proc. London Math. Soc. (3)*, 48(1):175–192, 1984.

[Wei67] André Weil. *Basic number theory*. Die Grundlehren der mathematischen Wissenschaften, Band 144. Springer-Verlag New York, Inc., New York, 1967.

[Wei82] André Weil. *Adeles and algebraic groups*, volume 23 of *Progress in Mathematics*. Birkhäuser Boston, Mass., 1982. With appendices by M. Demazure and Takashi Ono.

Integral Positive Ternary Quadratic Forms

William C. Jagy

Abstract We discuss some families of integral positive ternary quadratic forms. Our main example is $f(x,y,z) = x^2 + y^2 + 16nz^2$, where n is positive, squarefree, and $n = u^2 + v^2$ with $u, v \in \mathbf{Z}$.

Key words Ternary quadratic forms • Spinor genus

Subject Classification: Primary 11E20; Secondary 11D85, 11E12, 11E25

1 Notation

As in [4, 13], and Sect. 7 of [15], we let the integer sextuple

$$\langle a, b, c, r, s, t \rangle$$

refer to the quadratic form

$$f(x,y,z) = ax^2 + by^2 + cz^2 + ryz + szx + txy.$$

The Gram matrix for the form is the matrix of second partial derivatives:

$$\begin{pmatrix} 2a & t & s \\ t & 2b & r \\ s & r & 2c \end{pmatrix}.$$

W.C. Jagy (✉)
Mathematical Sciences Research Institute, 17 Gauss Way, Berkeley, CA 94720-5070, USA
e-mail: jagy@msri.org

K. Alladi et al. (eds.), *Quadratic and Higher Degree Forms*, Developments in Mathematics 31, DOI 10.1007/978-1-4614-7488-3_6,

So our Gram matrix is symmetric, positive definite, and has integer entries.

We define our discriminant Δ as half the determinant of the matrix above, so

$$\Delta = 4abc + rst - ar^2 - bs^2 - ct^2.$$

All our forms are positive and primitive ($\gcd(a,b,c,r,s,t) = 1$). Note that we do allow some of r,s,t to be odd at times. When r,s,t are all even, we refer to the form as classically integral.

2 Introduction

In a 1995 letter to J.S. Hsia and R. Schulze-Pillot, Irving Kaplansky pointed out some simple properties of

$$\langle 2,2,4k^2+1,2,2,0\rangle$$

or

$$f(x,y,z) = 2x^2 + 2y^2 + (4k^2+1)z^2 + 2yz + 2zx.$$

When k is odd, then $f \neq m^2$, in notation going back to Jones and Pall [11], where this means that all prime factors of m are congruent to $1 \pmod 4$.

We give the simple proof, while changing the focus to

$$\langle 2,2,4n+1,2,2,0\rangle$$

where n is odd, squarefree, and $n = u^2 + v^2$ in integers. Furthermore the numbers not represented will be all nm^2.

Lemma 2.1. *Let n be positive, odd, squarefree, and $n = u^2 + v^2$ in integers. Then*

$$\langle 2,2,4n+1,2,2,0\rangle \neq nm^2.$$

Proof: We have the identity

$$2x^2 + 2y^2 + (4n+1)z^2 + 2yz + 2zx = (x+y+z)^2 + (x-y)^2 + 4nz^2.$$

That is to say, $\langle 2,2,4n+1,2,2,0\rangle$ represents all numbers that can be expressed as $U^2 + V^2 + 4nz^2$ with $U+V+z$ even. So, assume we have

$$U^2 + V^2 + 4nz^2 = nm^2, \quad U + V + z \equiv 0 \pmod 2.$$

As n,m are odd, it follows that $U+V$ is odd, so z is also odd and nonzero. Then

$$U^2 + V^2 = n(m^2 - 4z^2),$$

and

$$U_1^2 + V_1^2 = m^2 - 4z^2 = (m + 2z)(m - 2z).$$

Now, $m + 2z \equiv 3 \pmod 4$, $m - 2z \equiv 3 \pmod 4$. There is some prime $q \equiv 3 \pmod 4$ such that $q^{2i+1} \parallel m + 2z$. However, $(m + 2z)(m - 2z)$ is the sum of two squares, so we also have $q^{2j+1} \parallel m - 2z$, from which it follows that $q|m$, a contradiction. ○

Our discussion of the genus containing $\langle 2, 2, 4n + 1, 2, 2, 0\rangle$ is simplified by

Lemma 2.2. *Let k be any positive integer. Then $\langle 1, 1, 16k, 0, 0, 0\rangle$ and $\langle 2, 2, 4k + 1, 2, 2, 0\rangle$ are in the same genus.*

Proof: We use Proposition 4 on page 410 of Lehman [13], using his terminology and notation, once for each form. Divisor, reciprocal, and level are defined on page 402, while conditions we need on the relationship of the form and its reciprocal are given in Proposition 2 on page 403.

First, we take $f = \langle a, b, c, r, s, t\rangle = \langle 1, 1, 16k, 0, 0, 0\rangle$, which has discriminant $64k$, level $64k$, and divisor $m = 4$. Next, we find its reciprocal $\phi = \langle \alpha, \beta, \gamma, \rho, \sigma, \tau\rangle = \langle 16k, 16k, 1, 0, 0, 0\rangle$, which has discriminant $1024k^2$, level $64k$, and divisor $\mu = 64k$. So we have $a = \gamma = 1$.

Lehman defines the collection of genus symbols on page 410. As $m = 4$ is not divisible by any odd prime or by 16 or 32, none of the genus symbols $(f|\cdot)$ are defined. As $\mu = 64k$ and $\gamma = 1$, for any odd prime dividing k we have $(\phi|p) = (\gamma|p) = (1|p) = 1$. Then, as $16, 32|\mu$, we have $(\phi|4) = (-1)^{(\gamma-1)/2} = (-1)^0 = 1$, then $(\phi|8) = (-1)^{(\gamma^2-1)/8} = (-1)^0 = 1$.

We need to take a cyclic permutation of variables in our second form to use these results, so, reusing most of the letters, take $h = \langle a, b, c, r, s, t\rangle = \langle 4k + 1, 2, 2, 0, 2, 2\rangle$, which has discriminant $64k$, level $64k$, and divisor $m = 4$. The reciprocal is $\eta = \langle \alpha, \beta, \gamma, \rho, \sigma, \tau\rangle = \langle 4, 8k + 1, 8k + 1, 2, -4, -4\rangle$, which has discriminant $1024k^2$, level $64k$, and divisor $\mu = 64k$. This time $a = 4k + 1$ and $\gamma = 8k + 1$. This works out, insofar as the conditions in Proposition 2 are that $\gcd(a, \gamma) = \gcd(a, m\mu) = \gcd(\gamma, m\mu) = 1$.

Once again, with $m = 4$, Lehman gives no value for any of the genus symbols $(h|\cdot)$. For any odd prime $p|k$, we get $(\eta|p) = (\gamma|p) = (8k + 1|p) = (1|p) = 1$. Then, as $16, 32|\mu$, we have $(\eta|4) = (-1)^{(\gamma-1)/2} = (-1)^{4k} = 1$, then $(\eta|8) = (-1)^{(\gamma^2-1)/8} = (-1)^{8k^2+2k} = 1$.

We have calculated discriminant, level, and collection of genus symbols for f, h and found agreement, so our two forms are in the same genus by Proposition 4 of [13]. ○

We introduce a celebrated result of Duke and Schulze-Pillot, which is the Corollary to Theorem 3 in [6]:

Theorem 2.3. *Let $q(x_1, x_2, x_3)$ be a positive integral ternary quadratic form. Then every large integer n represented primitively by a form in the spinor genus of q is represented by q itself and the representing vectors are asymptotically uniformly distributed on the ellipsoid $q(\mathbf{x}) = n$.*

We will also need a short lemma on binary forms:

Lemma 2.4. *If all prime factors of a positive integer are* $1 \pmod 4$, *then it can be represented primitively as* $x^2 + y^2$, *that is with* $\gcd(x, y) = 1$.

From Lemma 2.4, when n is odd, squarefree, and $n = u^2 + v^2$ in integers, and all prime factors of m are $1 \pmod 4$ as well (although m need not be squarefree), we see that nm^2 is primitively represented by $\langle 1, 1, 16n, 0, 0, 0\rangle$. But Kaplansky's argument has shown that $\langle 2, 2, 4n + 1, 2, 2, 0\rangle \neq nm^2$. It now follows from Theorem 2.3 that $\langle 2, 2, 4n + 1, 2, 2, 0\rangle$ and $\langle 1, 1, 16n, 0, 0, 0\rangle$, while in the same genus, are in fact in different spinor genera, so there are at least two spinor genera in this genus.

J. S. Hsia [9] confirmed for the author that, for both odd and even squarefree $n = u^2 + v^2$, the genus of $\langle 1, 1, 16n, 0, 0, 0\rangle$ has exactly two spinor genera, and that n itself is a **spinor exceptional integer** (a number not represented by one of the spinor genera). He mentioned that the methods were in [7]. He also pointed out his proof that, if there are any spinor exceptions for a genus, there is one that divides 2Δ, this being Theorem 2 in [8]. Our family shows that the smallest spinor exception can be as large as $\Delta/64$.

We return briefly to the base genus, with our $n = 1$. For all numbers except odd squares, the number of representations by $\langle 1, 1, 16, 0, 0, 0\rangle$ is the same as the number of representations by $\langle 2, 2, 5, 2, 2, 0\rangle$. Then, for k **odd**, $r_{\langle 1,1,16,0,0,0\rangle}(k^2) - r_{\langle 2,2,5,2,2,0\rangle}(k^2) = 4\ (-1|k)\ k$. Complete proofs of these facts have been supplied by Alexander Berkovich [3] and Wadim Zudilin [17], in the language of modular forms. In this situation, the odd squares are called the **splitting integers** for the genus, as the Siegel weighted average representation of the odd squares for one spinor genus disagrees with that of the other spinor genus. As it is also possible to calculate the Siegel weighted average of representations for any genus, this allows one to separately calculate $r_{\langle 1,1,16,0,0,0\rangle}(j)$ and $r_{\langle 2,2,5,2,2,0\rangle}(j)$ for any integer j. Splitting integers are used in Sect. 2 of [1] to correctly partition a genus of ten classes into its spinor genera, five classes each. The characterization of splitting integers as disagreement of representation measures is Corollary 1 on page 3 of [1]. An anonymous referee has pointed out that explicit calculation of the difference of representation measures is dealt with in Satz 2 and Korollar 1 of [14].

3 A Rare Phenomenon

We have mentioned that, with n squarefree and $n = u^2 + v^2$, the genus of $\langle 1, 1, 16n, 0, 0, 0\rangle$ has two spinor genera, and n itself is a spinor exception. In this section we prove

Theorem 3.1. *Let* n *be positive, odd, squarefree, and* $n = u^2 + v^2$ *in integers. Then every form in the same spinor genus as* $\langle 1, 1, 16n, 0, 0, 0\rangle$ *also integrally represents* n.

In another section we will prove the same result for even n.

The main tool is a **genus-correspondence**, with the first simple properties conjectured by the author, and proved by Wai Kiu Chan [5]. First we need to describe what we mean by a ternary form representing a multiple of another ternary form.

Suppose we have two positive ternary forms f, g, with Gram matrices F, G, and suppose we have some positive integer k. We will say that f represents kg when there is an integral matrix P such that

$$P^t \, F P = kG.$$

The easiest consequence of such a relationship is that, whenever g integrally represents an integer w, it follows that f integrally represents kw.

Our concern is for the situation when two forms represent prescribed multiples of each other:

Theorem 3.2 (Chan). *Suppose f_0, g_0 are positive ternary forms with* **integral** *discriminant ratio k. Suppose that f_0 represents kg_0 and g_0 represents kf_0. Then, for any $f_1 \in$ gen f_0, there is at least one $g_1 \in$ gen g_0 such that f_1 represents kg_1 and g_1 represents kf_1. Also, for any $g_2 \in$ gen g_0, there is at least one $f_2 \in$ gen f_0 such that g_2 represents kf_2 and f_2 represents kg_2.*

We call this a **genus-correspondence** because it is generally many-to-many, that is, there is generally no well-defined mapping on equivalence classes of forms in either direction.

We are now able to prove Theorem 3.1. Take $n = u^2 + v^2$ to be squarefree and **odd**. Let G_0 be the Gram matrix for $g_0 = \langle 1, 1, 16n, 0, 0, 0 \rangle$, so that

$$G_0 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 32n \end{pmatrix}.$$

Let F_0 be the Gram matrix for $f_0 = \langle 1, 1, 16, 0, 0, 0 \rangle$, so that

$$F_0 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 32 \end{pmatrix}.$$

We have $P^t \, G_0 \, P = nF_0$, with

$$P = \begin{pmatrix} u & v & 0 \\ -v & u & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note that $\det P = n$. We take the adjoint Q so that $PQ = QP = nI$ and for that matter $\det Q = n^2$. We find that $Q^t \, F_0 \, Q = nG_0$, with

$$Q = \begin{pmatrix} u & -v & 0 \\ v & u & 0 \\ 0 & 0 & n \end{pmatrix}.$$

Furthermore, the ratio of the discriminants of f_0, g_0 is $64n/64 = n$. So f_0 represents ng_0 and g_0 represents nf_0, and Theorem 3.2 applies.

Let g_1 be any form in the spinor genus of g_0, written $g_1 \in \text{spn } g_0$. According to Lemma 2.4, for any prime $p \equiv 1 \pmod 4$, we know that $x^2 + y^2$ and therefore g_0 represent np^2 primitively. According to Theorem 2.3, when p is sufficiently large, np^2 is also represented by g_1. From Theorem 3.2, we know that g_1 corresponds with either $f_0 = \langle 1, 1, 16, 0, 0, 0 \rangle$ or $f_1 = \langle 2, 2, 5, 2, 2, 0 \rangle$. However, if f_1 represented ng_1, then f_1 would integrally represent n^2p^2, which is a spinor exception for this genus and is not, in fact, represented by f_1. It follows that g_1 represents nf_0 and f_0 represents ng_1. In particular, g_1 integrally represents n. This completes the proof of Theorem 3.1. ○

Next, consider any $g_2 \in \text{gen } g_0$ but $g_2 \notin \text{spn } g_0$. Then g_2 does not represent n, as n is a spinor exception for gen g_0. So it is not possible for g_2 to represent nf_0. From Theorem 3.2, we find that g_2 represents nf_1, where $f_1 = \langle 2, 2, 5, 2, 2, 0 \rangle$. We have chosen to say that this genus-correspondence respects spinor genus. Formally, we could say this: given a pair of genera with discriminant ratio k and a genus-correspondence. Suppose that f_3 represents kg_3 and g_3 represents kf_3, while f_4 represents kg_4 and g_4 represents kf_4. We say that the genus-correspondence **respects spinor genus** when f_3, f_4 are in the same spinor genus if and only if g_3, g_4 are in the same spinor genus.

We have extensive numerical support for the following:

Conjecture 3.3. *Given two genera G_1, G_2 of positive ternary forms, with integral* **squarefree** *discriminant ratio and with a genus-correspondence. Suppose that G_1, G_2 both have exactly two spinor genera. Then G_1 has spinor exceptional integers if and only if G_2 has spinor exceptional integers, G_1 has splitting integers if and only if G_2 has splitting integers, and the genus-correspondence respects spinor genus. When there are spinor exceptions, the regular spinor genera correspond. When there are splitting integers, the spinor genera that have larger (weighted) representation measures for the smallest splitting integers correspond.*

We should emphasize that a genus need not have splitting integers. The best known example is that of gen $\langle 1, 17, 289, 0, 0, 0 \rangle$, from page 257 of [2]. The example with the smallest discriminant (1375) is gen $\langle 1, 5, 70, 5, 0, 0 \rangle$, just beyond the range of the Brandt and Intrau tables [4]. It was rather surprising that splitting integers were not evidently required for a genus-correspondence to respect spinor genus, as there is then no apparent way to label one spinor genus as "more regular" than the other.

With less detail and far less evidence, we also offer, for four or more spinor genera,

Conjecture 3.4. *Given two genera G_1, G_2 of positive ternary forms, with integral* **squarefree** *discriminant ratio and with a genus-correspondence. Suppose that G_1, G_2 have exactly the same number (some 2^j) of spinor genera. Then the genus-correspondence respects spinor genus.*

Note that, with squarefree discriminant ratio and a genus-correspondence, it is still common for either the genus with larger discriminant or the genus with the smaller discriminant to have fewer spinor genera than the other. Such examples can be quite instructive.

4 Tornaria's Constructions

Gonzalo Tornaria was kind enough to describe the genus-correspondence, in two situations, as a mapping between forms in some canonical shapes. These mappings do not extend to mappings of equivalence classes. The virtue of this approach is the placing of the genus-correspondence as merely one variant of Kaplansky's "descent" steps, used in preparing [10], and described throughout [12]. The similarity to Watson transformations [16] also becomes apparent, although a Watson transformation is a well-defined mapping on equivalence classes of forms, and a Watson transformation does not send a form with some odd prime $p \parallel \Delta$ to a form with $\Delta \not\equiv 0 \pmod p$. The closest parallel we know involving a Watson transformation is the descent of a form (probably regular) with $\Delta = 2592 = 32 \cdot 81$ to one with $\Delta = 32$ that is regular, in that

$$\lambda_9(\langle 5, 9, 17, 6, 5, 3\rangle) = \langle 1, 3, 3, 1, 0, 1\rangle.$$

We have taken some extra care to show how Tornaria's ascent and descent steps may be viewed as inverses, at least to the extent that they interchange forms in one canonical shape with forms in another canonical shape.

Take an odd prime p and a discriminant such that $\Delta \not\equiv 0 \pmod p$. Take any form $f_0 = \langle a, b, c, r, s, t\rangle$ with discriminant Δ. As f_0 is isotropic in $\mathbf{Q}_p$, we may demand that $c \equiv 0 \pmod p$, in that such a value is indeed primitively represented by our form. From $\Delta \equiv rst - ar^2 - bs^2 \not\equiv 0 \pmod p$ we know that r, s are not both divisible by p. If necessary, interchange variables so that $s \not\equiv 0 \pmod p$. Formally, we have taken the Gram matrix A_1 and replaced it by the equivalent $A_2 = P^t\, A_1 P$, where

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The coefficients become $\langle b, a, c, s, r, t\rangle$, and we simply rename these with the original letters. So we now have $\langle a, b, c, r, s, t\rangle$ with $c \equiv 0 \pmod p), s \not\equiv 0 \pmod p$. Next, solve for k in $a + sk \equiv 0 \pmod p$, then find $A_3 = Q^t\, A_2 Q$, with

$$Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ k & 0 & 1 \end{pmatrix},$$

The new coefficients are $\langle a + sk + ck^2, b, c, r, s + 2ck, t + rk\rangle$. Renaming again, we have $\langle a, b, c, r, s, t\rangle$ with $a, c \equiv 0 \pmod{p}, s \not\equiv 0 \pmod{p}$, this being the first of the two canonical shapes. Then we may construct the form

$$g_0(x, y, z) = \frac{1}{p}\, f_0(px, py, z),$$

with coefficients

$$g_0 = \left\langle pa, pb, \frac{c}{p}, r, s, pt \right\rangle .$$

In the descent direction, let $\Delta \equiv 0 \pmod{p}$ and $\Delta \not\equiv 0 \pmod{p^2}$, or $p \parallel \Delta$. Let $g_1 = \langle a, b, c, r, s, t\rangle$ have discriminant Δ. This time we need to explicitly require that the form be isotropic in $\mathbf{Q}_p$. We then demand that $p^2|a$. It follows that $\Delta \equiv rst - bs^2 - ct^2 \not\equiv 0 \pmod{p^2}$. Thus we know that s, t are not both divisible by p. If necessary, transpose s, t so that $s \not\equiv 0 \pmod{p}$. We are taking the Gram matrix B_1 and replacing it by $B_2 = P^t\ B_1 P$, where

$$P \;=\; \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} .$$

Next, solve for an integer k in $t + sk \equiv 0 \pmod{p}$. Construct the matrix

$$Q \;=\; \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & k & 1 \end{pmatrix} ,$$

and take the form with Gram matrix $B_3 = Q^t\ B_2 Q$. The new coefficients are $\langle a, b + rk + ck^2, c, r + 2ck, s, t + sk\rangle$. The value t has thus been replaced by $t + sk$, divisible by p, but without altering the value of a or s. At this point, $\Delta \equiv -bs^2 \pmod{p}$, so that $p|b$. We now have our form in the second canonical shape, $g_1 = \langle a, b, c, r, s, t\rangle$, with a, b, t all divisible by p, indeed $p^2|a$, but $s \not\equiv 0 \pmod{p}$. The new form, with discriminant $\frac{\Delta}{p}$, is given by

$$f_1(x, y, z) = \frac{1}{p}\, g_1(x, y, pz),$$

with coefficients

$$f_1 = \left\langle \frac{a}{p}, \frac{b}{p}, pc, r, s, \frac{t}{p} \right\rangle .$$

5 Even n

We prove the other case of Theorem 3.1, namely

Theorem 5.1. *Let n be positive, even, squarefree, and $n = u^2 + v^2$ in integers. Then every form in the same spinor genus as $\langle 1, 1, 16n, 0, 0, 0\rangle$ also integrally represents n.*

Proof: The genus containing $f_0 = \langle 1, 1, 32, 0, 0, 0\rangle$ consists of three classes, in two spinor genera. The first spinor genus contains the classes $\langle 1, 1, 32, 0, 0, 0\rangle$ and $\langle 2, 2, 9, 2, 2, 0\rangle$, both of which represent 2. The other spinor genus consists of the single class $\langle 1, 4, 9, 4, 0, 0\rangle$, which does not represent 2 or any $2m^2$.

With n even, $g_0 = \langle 1, 1, 16n, 0, 0, 0\rangle$ represents $\frac{n}{2} \cdot \langle 1, 1, 32, 0, 0, 0\rangle$, so that $\langle 1, 1, 32, 0, 0, 0\rangle$ also represents $\frac{n}{2} \cdot \langle 1, 1, 16n, 0, 0, 0\rangle$, and there is thus a genus-correspondence. Consider some $g_1 \in$ spn g_0. From Lemma 2.4, for any prime $p \equiv 1 \pmod 4$, we know that $x^2 + y^2$ represents $(n/2)p^2$ primitively, denote this $(n/2)p^2 = a^2 + b^2, \quad \gcd(a, b) = 1$. As $a^2 + b^2$ is odd, it follows that $\gcd(a - b, a + b) = 1$ as well. So we have the primitive representation $(a-b)^2 + (a+b)^2 = np^2$, which tells us that g_0 primitively represents np^2. When p is sufficiently large, Theorem 2.3 tells us that g_1 represents np^2. By Theorem 3.2, we know that g_1 corresponds with at least one of the three forms in the genus of f_0. However, if $\langle 1, 4, 9, 4, 0, 0\rangle$ should represent $\frac{n}{2} g_1$, it would follow that $\langle 1, 4, 9, 4, 0, 0\rangle$ represented the integer $\frac{n^2p^2}{2}$, which is of the form $2m^2$. It follows that g_0 represents either $\frac{n}{2} \cdot \langle 1, 1, 32, 0, 0, 0\rangle$ or $\frac{n}{2} \cdot \langle 2, 2, 9, 2, 2, 0\rangle$. In either case g_0 represents the integer n. ○

We pause to discuss the influence of Conjecture 3.3. It was necessary to have a separate proof for even n because $4m^2$ is not a spinor exception for the genus containing $\langle 1, 1, 16, 0, 0, 0\rangle$. If we had known a proof of Conjecture 3.3, we could simply have said that any form in the same spinor genus as $\langle 1, 1, 16n, 0, 0, 0\rangle$ represents $n \cdot \langle 1, 1, 16n, 0, 0, 0\rangle$. Similarly, we would not have needed any invocation of Theorem 2.3, which can become unusable if primitive representations of desirable numbers are not available.

Conjecture 3.3 would be an even bigger help in the following related pair of examples, where the conjectured behavior has simply not been proved, although checked as correct for $n \leq 200$. One situation is $n = u^2 + uv + 4v^2$ squarefree, with the genus of $\langle 1, 4, 225n, 0, 0, 1\rangle$. Second, $n = 2u^2 + uv + 2v^2$ squarefree, and the genus of $\langle 2, 2, 225n, 0, 0, 1\rangle$. In these cases n is allowed odd or even. The “base” genus has four forms in two spinor genera of two classes each: $\langle 1, 4, 225, 0, 0, 1\rangle$ and $\langle 1, 15, 60, 15, 0, 0\rangle$ are in one spinor genus, $\langle 6, 6, 25, 0, 0, 3\rangle$ and $\langle 9, 10, 10, 5, 0, 0\rangle$ are in the other. The spinor exceptions are of the form μ^2, where all prime factors of μ are $1, 2, 4, 8 \pmod{15}$, and 2 itself is included. As $9\mu^2$ and $25\mu^2$ are not spinor exceptions, to deal with n divisible by $3, 5, 15$, we would first need to calculate the genera of $\langle 2, 2, 675, 0, 0, 1\rangle$, $\langle 2, 2, 1125, 0, 0, 1\rangle$, and $\langle 1, 4, 3375, 0, 0, 1\rangle$.

6 Involutions

We return to odd squarefree $n = u^2+v^2$ and the genus of $\langle 1, 1, 16n, 0, 0, 0\rangle$. As long as $n \leq 505$, a few interesting things happen. First, the two spinor genera in the genus have the same number of equivalence classes of forms. Second, for each class f, there is a single class g with $f \neq g$, such that f represents $4g$ and g represents $4f$, while g never lies in the same spinor genus as f. So "involution" seems a good term for this, as we have a bijection that interchanges the two spinor genera.

A similar thing happens in these two situations, from the last paragraph of Sect. 5: first, $n = u^2 + uv + 4v^2$, with $\langle 1, 4, 225n, 0, 0, 1\rangle$, or second, $n = 2u^2 + uv + 2v^2$, with $\langle 2, 2, 225n, 0, 0, 1\rangle$, while we keep n squarefree, but add the restriction that n not be divisible by 3 or 5. There are indeed two spinor genera, and they are the same size, checked for $n \leq 200$. The worthwhile detail is that we get one involution where each f has a single $g \neq f$ such that f represents $9g$ and g represents $9f$, so that is one involution. But there is a different involution where f represents $25g$ and g represents $25f$. Both 9 and 25 interchange spinor genera. Nothing special occurs with 4.

This last conjecture has not been checked as thoroughly, but is worthwhile for suggesting possibilities with four spinor genera. In [2], there is a genus with four spinor genera described, containing the form called $B^1 = \langle 1, 20, 400, 0, 0, 0\rangle$. The spinor genera all have three classes. There are two families of spinor exceptions, $5m^2$, all prime factors of m being $1 \pmod 4$, and ϕ^2, where all prime factors of ϕ are $1, 3, 7, 9 \pmod{20}$.

This first step has been checked for $n \leq 1189 = 29 \cdot 41$. Let n be squarefree, and all prime factors of n be either $1 \pmod{20}$ or $9 \pmod{20}$. Then the genus of $\langle 1, 20, 400n, 0, 0, 0\rangle$ has four spinor genera of equal size. Either $n = u^2 + 20v^2$ or $n = 4u^2 + 5v^2$, and it is easy to check that $\langle 1, 20, 400n, 0, 0, 0\rangle$ represents either $n \cdot \langle 1, 20, 400, 0, 0, 0\rangle$ or $n \cdot \langle 4, 5, 400, 0, 0, 0\rangle$. In turn, the relevant form in the "base" genus represents $n \cdot \langle 1, 20, 400n, 0, 0, 0\rangle$. This extends to a genus-correspondence. With all as described, this genus-correspondence respects spinor genus.

Let us label the four spinor genera. Let A_n be regular, let $B_n \neq 5nm^2$, let $C_n \neq n\phi^2$, finally $D_n \neq 5nm^2, n\phi^2$. These next items have been checked only as far as $n \leq 61$. There are involutions with multiplier 25, these interchange A_n with C_n, and then interchange B_n with D_n.

In comparison, with multiplier 4, any form in A_n corresponds with a single one in D_n, but with two forms each in B_n, C_n. Similar comments apply beginning with any of the four spinor genera. So multiplier 4 does give an identifiable involution, (B_n matches with C_n,) but the behavior is not as clean as that with multiplier 25.

Finally, we explain the restriction on n itself. If n is a number that is represented by both the binary forms $x^2 + 20y^2$ and $4x^2 + 5y^2$, such as $n = 21$, then $\langle 1, 20, 400n, 0, 0, 0\rangle$ represents **both** $n \cdot \langle 1, 20, 400, 0, 0, 0\rangle$ **and** $n \cdot \langle 4, 5, 400, 0, 0, 0\rangle$, so that it is not possible to have a genus-correspondence that respects spinor genus, even if the resulting genus does actually possess four spinor genera.

Acknowledgements This article is a selection from the author's talk during a Focused Week at the University of Florida, Gainesville, in March 2010. The author would like to thank the organizers for a delightful event.

We would also like to thank A. Berkovich, W. K. Chan, J. S. Hsia, G. Tornaria, and W. Zudilin for fundamental contributions to this article and permission to include these. We especially wish to thank W. Zudilin for his careful reading and extensive comments.

References

1. J.W. Benham, A. G. Earnest, J. S. Hsia, and D. C. Hung. Spinor regular positive ternary quadratic forms. *Journal of the London Mathematical Society*, 42:1–10, 1990.
2. J.W. Benham and J. S. Hsia. On spinor exceptional representations. *Nagoya Mathematical Journal*, 87:247–260, 1982.
3. A. Berkovich. Personal communication, 2010.
4. H. Brandt and O. Intrau. Tabelle reduzierten positiver ternärer quadratischer Formen. *Abh. der Sächsischen Akad. der Wissenschaften zu Leipzig, Math.-Naturw.*, 45(4), 1958.
5. W. K. Chan. Personal communication, 2008. One page pdf.
6. W. Duke and R. Schulze-Pillot. Representations of integers by positive ternary quadratic forms and equidistribution of lattice points on ellipsoids. *Inventiones Mathematicae*, 99:49–57, 1990.
7. A. G. Earnest, J. S. Hsia, and D. C. Hung. Primitive representations by spinor genera of ternary quadratic forms. *Journal of the London Mathematical Society*, 50:222–230, 1994.
8. J. S. Hsia. Regular positive ternary quadratic forms. *Mathematika*, 28:231–238, 1981.
9. J. S. Hsia. Personal communication, 2004.
10. W. C. Jagy, I. Kaplansky, and A. Schiemann. There are 913 regular ternary quadratic forms. *Mathematika*, 44:332–341, 1997.
11. B. W. Jones and G. Pall. Regular and semi-regular positive ternary quadratic forms. *Acta Mathematica*, 70:165–191, 1939.
12. I. Kaplansky. Notes on the classification of regular ternary forms. Unpublished, 1996.
13. J. Larry Lehman. Levels of positive definite ternary quadratic forms. *Mathematics of Computation*, 58:399–417, 1992.
14. R. Schulze-Pillot. Darstellungsmaße von Spinorgeschlechtern ternärer quadratischer Formen. *J. Riene Angew. Math.*, 352:114–132, 1984.
15. G. L. Watson. *Some problems in the theory of numbers*. PhD thesis, University of London, 1953.
16. G. L. Watson. Transformations of a quadratic form which do not increase the class-number. *Proceedings of the London Mathematical Society*, 12:577–587, 1962.
17. W. Zudilin. Personal communication, 2010.

Some Aspects of the Algebraic Theory of Quadratic Forms

R. Parimala

Abstract This article, based on the lectures at the Arizona Winter School on "Quadratic forms", gives a quick introduction to the algebraic theory of quadratic forms. It discusses some invariants associated to quadratic forms like the Pythagoras number and the u-invariant and touches on some recent progress on these topics.

Key words and Phrases Quadratic forms • Galois cohomology • Invariants • Number fields • Function fields

Mathematics Subject Classification (2010): 11E81, 14G27

This text is based on the lectures given at the Arizona Winter School on "Quadratic forms". The aim of the text is to give a brief introduction to the algebraic theory of quadratic forms. We explain invariants associated to quadratic forms—invariants with values in Galois cohomology as well as numerical invariants. We explain some open questions concerning these invariants and recent progress related to these questions.

There are many good references for this material on the algebraic theory of quadratic forms including [EKM, K, L, Pf] and [S].

The author was partially supported by NSF grant DMS-1001872.

R. Parimala (✉)
Department of Mathematics and Computer Science, Emory University, 400 Dowman Drive,
Atlanta, GA 30322, USA
e-mail: parimala@mathcs.emory.edu

K. Alladi et al. (eds.), *Quadratic and Higher Degree Forms*, Developments
in Mathematics 31, DOI 10.1007/978-1-4614-7488-3_7,

1 Quadratic Forms

Let k be a field with $\operatorname{char} k \neq 2$.

Definition 1.1. A **quadratic form** $q\colon V \to k$ on a vector space V over k is a map satisfying:

(1) $q(\lambda v) = \lambda^2 q(v)$ for $v \in V$, $\lambda \in k$.
(2) The map $b_q\colon V \times V \to k$, defined by

$$b_q(v,w) = \frac{1}{2}[q(v+w) - q(v) - q(w)]$$

is bilinear.

We denote a quadratic form by (V,q), or simply by q. Throughout, we restrict ourselves to the study of quadratic forms on finite-dimensional vector spaces.

The bilinear form b_q is symmetric; q determines b_q and for all $v \in V$, $q(v) = b_q(v,v)$.

For a choice of basis $\{e_1, \ldots, e_n\}$ of V, b_q is represented by a symmetric matrix $A(q) = (a_{ij})$ with $a_{ij} = b_q(e_i, e_j)$. If $v = \sum_{1\leq i\leq n} X_i e_i \in V$, $X_i \in k$, then

$$q(v) = \sum_{1\leq i,j\leq n} a_{ij} X_i X_j = \sum_{1\leq i\leq n} a_{ii} X_i^2 + 2\sum_{i<j} a_{ij} X_i X_j.$$

Thus q is represented by a homogeneous polynomial of degree 2. Clearly, every homogeneous polynomial of degree 2 corresponds to a quadratic form on V with respect to the chosen basis.

Definition 1.2. Two quadratic forms (V_1, q_1), (V_2, q_2) are **isometric** if there is an isomorphism $\phi\colon V_1 \xrightarrow{\sim} V_2$ such that $q_2(\phi(v)) = q_1(v)$ for all $v \in V_1$.

If $A(q_1)$, $A(q_2)$ are the matrices representing q_1 and q_2 with respect to bases B_1 and B_2 of V_1 and V_2 respectively, ϕ yields a matrix $T \in GL_n(k)$, $n = \dim V$, such that

$$T A(q_2)\, T^t = A(q_1).$$

In other words, the symmetric matrices $A(q_1)$ and $A(q_2)$ are congruent. Thus isometry classes of quadratic forms yield congruence classes of symmetric matrices.

Definition 1.3. The form $q\colon V \to k$ is said to be **regular** if $b_q\colon V \times V \to k$ is nondegenerate.

Thus q is regular if and only if the map $V \to V^* = \operatorname{Hom}(V,k)$, defined by $v \mapsto (w \mapsto b_q(v,w))$, is an isomorphism. This is the case if $A(q)$ is invertible.

Let (V,q) be a quadratic form. Then

$$V_0 = \{v \in V \,:\, b_q(v,w) = 0 \text{ for all } w \in V\}$$

is called the **radical** of V. If V_1 is any complementary subspace of V_0 in V, then $q|_{V_1}$ is regular and $(V, q) = (V_0, 0) \perp (V_1, q|_{V_1})$. Note that V is regular if and only if the radical of V is zero.

Henceforth, we shall only be concerned with regular quadratic forms.

Definition 1.4. Let W be a subspace of V and $q\colon V \to k$ be a quadratic form. The **orthogonal complement** of W denoted $W^{\perp}$ is the subspace

$$W^{\perp} = \{v \in V : b_q(v, w) = 0 \text{ for all } w \in W\}.$$

Exercise 1.5. Let (V, q) be a regular quadratic form and W a subspace of V. Then

(1) $\dim(W) + \dim(W^{\perp}) = \dim(V)$.
(2) $(W^{\perp})^{\perp} = W$.

1.1 Orthogonal Sums

Let (V_1, q_1), (V_2, q_2) be quadratic forms. The form

$$(V_1, q_1) \perp (V_2, q_2) = (V_1 \oplus V_2, q_1 \perp q_2),$$

with $q_1 \perp q_2$ defined by

$$(q_1 \perp q_2)(v_1, v_2) = q_1(v_1) + q_2(v_2),\ v_1 \in V_1,\ v_2 \in V_2$$

is called the *orthogonal sum* of (V_1, q_1) and (V_2, q_2).

1.2 Diagonalization

Let (V, q) be a quadratic form. There exists a basis $\{e_1, \ldots, e_n\}$ of V such that $b_q(e_i, e_j) = 0$ for $i \neq j$. Such a basis is called an *orthogonal basis* for q. With respect to an orthogonal basis, b_q is represented by a diagonal matrix.

If $\{e_1, \ldots, e_n\}$ is an orthogonal basis of q and $q(e_i) = d_i$, we write $q = \langle d_1, \ldots, d_n \rangle$. In this case, $V = ke_1 \oplus \cdots \oplus ke_n$ is an orthogonal sum and $q|_{ke_i}$ is represented by $\langle d_i \rangle$. Thus every quadratic form is diagonalizable.

1.3 Hyperbolic Forms

Definition 1.6. A quadratic form (V, q) is said to be **isotropic** if there is a nonzero $v \in V$ such that $q(v) = 0$. It is **anisotropic** if q is not isotropic. A quadratic form (V, q) is said to be **universal** if it represents every element of k; i.e., given $\lambda \in k$, there is a vector $v \in V$ such that $q(v) = \lambda$.

Example 1.7. The quadratic form $X^2 - Y^2$ is isotropic over k. Suppose (V, q) is a regular form which is isotropic. Let $v \in V$ be such that $q(v) = 0$, $v \neq 0$. Since q is regular, there exists $w \in V$ such that $b_q(v, w) \neq 0$. After scaling we may assume $b_q(v, w) = 1$. If $q(w) \neq 0$, we may replace w by $w + \lambda v$, $\lambda = -\frac{1}{2}q(w)$, and assume that $q(w) = 0$. Thus $W = kv \oplus kw$ is a 2-dimensional subspace of V and $q|_W$ is represented by $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ with respect to $\{v, w\}$.

Definition 1.8. A binary quadratic form isometric to $(k^2, \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right))$ is called a **hyperbolic plane**. A quadratic form (V, q) is **hyperbolic** if it is isometric to an orthogonal sum of hyperbolic planes. A subspace W of V such that q restricts to zero on W and $\dim W = \frac{1}{2} \dim V$ is called a **Lagrangian**.

Every regular quadratic form which admits a Lagrangian can easily be seen to be hyperbolic.

Exercise 1.9. Let (V, q) be a regular quadratic form and $(W, q|_W)$ a regular form on the subspace W. Then $(V, q) = (W, q|_W) \perp (W^\perp, q|_{W^\perp})$.

Theorem 1.10 (Witt's Cancellation Theorem). *Let (V_1, q_1), (V_2, q_2), (V, q) be quadratic forms over k. Suppose*

$$(V_1, q_1) \perp (V, q) \cong (V_2, q_2) \perp (V, q).$$

Then $(V_1, q_1) \cong (V_2, q_2)$.

The key ingredient of Witt's cancellation theorem is the following.

Proposition 1.11. *Let (V, q) be a quadratic form and $v, w \in V$ with $q(v) = q(w) \neq 0$. Then there is an isometry $\tau\colon (V, q) \xrightarrow{\sim} (V, q)$ such that $\tau(v) = w$.*

Proof. Let $q(v) = q(w) = d \neq 0$. Then

$$q(v + w) + q(v - w) = 2q(v) + 2q(w) = 4d \neq 0.$$

Thus $q(v + w) \neq 0$ or $q(v - w) \neq 0$. For any vector $u \in V$ with $q(u) \neq 0$, define $\tau_u\colon V \to V$ by

$$\tau_u(z) = z - \frac{2b_q(z, u)u}{q(u)}.$$

τ_u is an isometry called the *reflection with respect to u*.

Suppose $q(v - w) \neq 0$. Then $\tau_{v-w}\colon V \to V$ is an isometry of V which sends v to w. Suppose $q(v + w) \neq 0$. Then $\tau_w \circ \tau_{v+w}$ sends v to w. □

Remark 1.12. The orthogonal group of (V, q) denoted by $O(q)$ is the set of isometries of V onto itself. This group is generated by reflections. This is seen by an inductive argument on $\dim(q)$, using the above proposition.

Theorem 1.13 (Witt's decomposition). *Let (V, q) be a quadratic form (not necessarily regular). Then there is a decomposition*

$$(V,q) = (V_0, 0) \perp (V_1, q_1) \perp (V_2, q_2)$$

where V_0 is the radical of q, $q_1 = q|_{V_1}$ is anisotropic and $q_2 = q|_{V_2}$ is hyperbolic. If $(V,q) = (V_0,0) \perp (W_1, f_1) \perp (W_2, f_2)$ with f_1 anisotropic and f_2 hyperbolic, then

$$(V_1, q_1) \cong (W_1, f_1),\ (V_2, q_2) \cong (W_2, f_2).$$

Remark 1.14. A hyperbolic form (W,f) is determined by $\dim(W)$; for if $\dim(W) = 2n$, $(W,f) \cong nH$, where $H = (k^2, (\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}))$ is the hyperbolic plane.

From now on, we shall assume (V,q) is a regular quadratic form. We denote by q_{an} the quadratic form (V_1, q_1) in Witt's decomposition which is determined by q up to isometry. We call $\frac{1}{2}\dim(V_2)$ the *Witt index* of q. Thus any regular quadratic form q admits a decomposition $q \cong q_{an} \perp (nH)$, with q_{an} anisotropic and H denoting the hyperbolic plane. We also sometimes denote by H^n the sum of n hyperbolic planes.

2 Witt Group of Forms

2.1 Witt Groups

We set

$$W(k) = \{\text{isomorphism classes of regular quadratic forms over } k\}/\sim$$

where the Witt equivalence $\sim$ is given by:

$$(V_1, q_1) \sim (V_2, q_2) \qquad \iff \qquad \begin{array}{l} \text{there exist } r,\ s \in \mathbb{Z} \text{ such that} \\ (V_1, q_1) \perp H^r \cong (V_2, q_2) \perp H^s \end{array}.$$

$W(k)$ is a group under orthogonal sum:

$$[(V_1, q_1)] \perp [(V_2, q_2)] = [(V_1, q_1) \perp (V_2, q_2)].$$

The zero element in $W(k)$ is represented by the class of hyperbolic forms. For a regular quadratic form (V,q), $(V,q) \perp (V,-q)$ has Lagrangian

$$W = \{(v,v)\ :\ v \in V\}$$

so that $(V,q) \perp (V,-q) \cong H^n$, $n = \dim(V)$. Thus, $[(V,-q)] = -[(V,q)]$ in $W(k)$.

It follows from Witt's decomposition theorem that every element in $W(k)$ is represented by a unique anisotropic quadratic form up to isometry. Thus $W(k)$ may be thought of as a group made out of isometry classes of anisotropic quadratic forms over k.

The abelian group $W(k)$ admits a ring structure induced by tensor product on the associated bilinear forms. For example, if $q_1 \cong \langle a_1, \ldots, a_n\rangle$ and q_2 is a quadratic form, then $q_1 \otimes q_2 \cong a_1 q_2 \perp a_2 q_2 \perp \cdots \perp a_n q_2$.

Definition 2.1. Let $I(k)$ denote the ideal of classes of even-dimensional quadratic forms in $W(k)$. The ideal $I(k)$ is called the **fundamental ideal**. $I^n(k)$ stands for the nth power of the ideal $I(k)$.

Definition 2.2. Let $P_n(k)$ denote the set of isomorphism classes of forms of the type

$$\langle\langle a_1, \ldots, a_n\rangle\rangle := \langle 1, a_1\rangle \otimes \cdots \otimes \langle 1, a_n\rangle.$$

Elements in $P_n(k)$ are called **n-fold Pfister forms**.

The ideal $I(k)$ is generated additively by the forms $\langle 1, a\rangle$, $a \in k^*$. Moreover, the ideal $I^n(k)$ is generated additively by n-fold Pfister forms. For instance, for $n = 2$, the generators of $I^2(k)$ are of the form

$$\langle a, b\rangle \otimes \langle c, d\rangle \cong \langle 1, ac, ad, cd\rangle - \langle 1, cd, -bc, -bd\rangle = \langle\langle ac, ad\rangle\rangle - \langle\langle cd, -bc\rangle\rangle$$

Example 2.3. If $k = \mathbb{C}$, every 2-dimensional quadratic form over k is isotropic.

$$W(k) \cong \mathbb{Z}/2\mathbb{Z}$$

$$[(V, q)] \mapsto \dim(V) \pmod 2$$

is an isomorphism.

Example 2.4. Let $k = \mathbb{F}_{p^n}$, $p \neq 2$, be a finite field. Then $k^* = k \setminus \{0\}$ has two square classes, $\{1, u\}$. Every 3-dimensional quadratic form over k is isotropic. Further, $W(k) \cong \mathbb{Z}/4\mathbb{Z}$ if -1 is not a square in $\mathbb{F}_{p^n}$ and $W(k) \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ if -1 is a square in $\mathbb{F}_{p^n}$ (cf. [L], Corollary 3.6).

Example 2.5. If $k = \mathbb{R}$, every quadratic form q is represented by

$$\langle 1, \ldots, 1, -1, \ldots, -1\rangle$$

with respect to an orthogonal basis. The number r of $+1$'s and the number s of -1's in the diagonalization above are uniquely determined by the isomorphism class of q. The *signature* of q is defined as $r - s$. The signature yields a homomorphism $\operatorname{sgn}\colon W(\mathbb{R}) \to \mathbb{Z}$ which is an isomorphism.

2.2 Quadratic Forms Over p-Adic Fields

Let k be a finite extension of the field $\mathbb{Q}_p$ of p-adic numbers. We call k a *non-dyadic* p-adic field if $p \neq 2$. The field k has a discrete valuation v *extending* the p-adic valuation on $\mathbb{Q}_p$. Let π be a uniformizing parameter for v and κ the residue field for v. The field κ is a finite field of characteristic $p \neq 2$. Let u be a unit in k^* such that $\overline{u} \in \kappa$ is not a square. Then

$$k^*/{k^*}^2 = \{1, u, \pi, u\pi\}.$$

Since κ is finite, every 3-dimensional quadratic form over κ is isotropic. By Hensel's lemma, every 3-dimensional form $\langle u_1, u_2, u_3\rangle$ over k, with u_i units in k is isotropic. Since every form q in k has a diagonal representation

$$\langle u_1, \ldots, u_r\rangle \perp \pi\langle v_1, \ldots, v_s\rangle,$$

if r or s exceeds 3, q is isotropic. In particular every 5-dimensional quadratic form over k is isotropic. Further, up to isometry, there is a unique quadratic form in dimension 4 which is anisotropic, namely,

$$\langle 1, -u, -\pi, u\pi\rangle.$$

This is the norm form of the unique quaternion division algebra $H(u, \pi)$ over k (cf. Sect. 2.3).

2.3 Central Simple Algebras and the Brauer Group

Recall that a finite-dimensional algebra A over a field k is a *central simple algebra* over k if A is simple (has no two-sided ideals) and the center of A is k. Recall also that for a field k,

$$\mathrm{Br}(k) = \{\text{Isomorphism classes of central simple algebras over } k\} \,/ \sim$$

where the Brauer equivalence $\sim$ is given by: $A \sim B$ if and only if $M_n(A) \cong M_m(B)$ for some integers m, n. The pair $(\mathrm{Br}(k), \otimes)$ is a group. The inverse of $[A]$ is $[A^{\mathrm{op}}]$ where A^{op} is the *opposite algebra* of A: the multiplication structure, $*$, on A^{op} is given by $a * b = ba$. We have a k-algebra isomorphism $\phi\colon A \otimes A^{\mathrm{op}} \xrightarrow{\sim} \mathrm{End}_k(A)$ induced by $\phi(a \otimes b)(c) = acb$. The identity element in $\mathrm{Br}(k)$ is given by $[k]$. By Wedderburn's theorem on central simple algebras, the elements of $\mathrm{Br}(k)$ parametrize the isomorphism classes of finite-dimensional central division algebras over k.

For elements $a, b \in k^*$, we define the **quaternion algebra** $H(a, b)$ to be the 4-dimensional central simple algebra over k generated by $\{i, j\}$ with the relations $i^2 = a, j^2 = b, ij = -ji$. This is a generalization of Hamilton's quaternion algebra

$H(-1,-1)$ over the field of real numbers. The algebra $H(a,b)$ admits a canonical involution^{-}: $H(a,b) \to H(a,b)$ given by

$$\overline{\alpha + i\beta + j\gamma + ij\delta} = \alpha - i\beta - j\gamma - ij\delta$$

This involution gives an isomorphism $H(a,b) \cong H(a,b)^{\mathrm{op}}$; in particular, $H(a,b)$ has order 2 in $\mathrm{Br}(k)$. Let ${}_2\mathrm{Br}(k)$ denote the 2-torsion subgroup of the Brauer group of k. The norm form for this algebra is given by $N(x) = x\overline{x}$, which is a quadratic form on $H(a,b)$ represented with respect to the orthogonal basis $\{1, i, j, ij\}$ by $\langle 1, -a, -b, ab\rangle = \langle\langle -a, -b\rangle\rangle$.

2.4 Classical Invariants for Quadratic Forms

Let (V, q) be a regular quadratic form. We define $\dim(q) = \dim(V)$ and $\dim_2(q) = \dim(V)$ modulo 2. We have a ring homomorphism $\dim_2 \colon W(k) \to \mathbb{Z}/2\mathbb{Z}$. We note that $I(k)$ is the kernel of $\dim_2$. This gives an isomorphism

$$\dim_2 \colon W(k)/I(k) \xrightarrow{\sim} \mathbb{Z}/2\mathbb{Z}.$$

Let $\mathrm{disc}(q) = (-1)^{n(n-1)/2}[\det(A(q))] \in k^*/k^{*2}$. Since $A(q)$ is determined up to congruence, $\det(A(q))$ is determined modulo squares. We have $\mathrm{disc}(H) = 1$, where H is the hyperbolic plane. The discriminant induces a group homomorphism

$$\mathrm{disc} \colon I(k) \to k^*/k^{*2}$$

which is clearly onto. It is easy to verify that $\ker(\mathrm{disc}) = I^2(k)$. Thus the discriminant homomorphism induces an isomorphism $I(k)/I^2(k) \to k^*/k^{*2}$.

Example 2.6. Let $\langle a, b\rangle$ be a binary quadratic form. Then $\mathrm{disc}\langle a, b\rangle = -ab$. The discriminant is trivial if and only if $\langle a, b\rangle \cong \langle 1, -1\rangle$ is a hyperbolic plane. Further, if $\langle a, b\rangle$ represents a value $c \in k^*$, then $\langle a, b\rangle \cong \langle c, abc\rangle$.

The next invariant for quadratic forms is the Clifford invariant. To each quadratic form (V, q) we wish to construct a central simple algebra containing V whose multiplication on elements of V satisfies $v \cdot v = q(v)$. The smallest such algebra (defined by a universal property) will be the Clifford algebra.

Definition 2.7. The **Clifford algebra** $C(q)$ of the quadratic form (V, q) is $T(V)/I_q$, where I_q is the two-sided ideal in the tensor algebra $T(V)$ generated by $\{v \otimes v - q(v) \mid v \in V\}$.

The algebra $C(q)$ has a $\mathbb{Z}/2\mathbb{Z}$ gradation $C(q) = C_0(q) \oplus C_1(q)$ induced by the gradation $T(V) = T_0(V) \oplus T_1(V)$, where

$$T_0(V) = \bigoplus_{i \geq 0,\ i \text{ even}} V^{\otimes i} \qquad \text{and} \qquad T_1(V) = \bigoplus_{i \geq 1,\ i\text{odd}} V^{\otimes i}.$$

If $\dim(q)$ is even, then $C(q)$ is a central simple algebra over k. If $\dim(q)$ is odd, $C_0(q)$ is a central simple algebra over k. The Clifford algebra $C(q)$ comes equipped with an involution τ defined by $\tau(v) = -v$ for $v \in V$. Thus, if $\dim(q)$ is even, $C(q)$ determines a 2-torsion element in $\mathrm{Br}(k)$.

Definition 2.8. The **Clifford invariant** $c(q)$ of (V, q) in $\mathrm{Br}(k)$ is defined as

$$c(q) = \begin{cases} [C(q)], & \text{if } \dim(q) \text{ is even} \\ [C_0(q)], & \text{if } \dim(q) \text{ is odd} \end{cases}$$

Example 2.9. Let $q \cong \otimes_{i=1}^{n} \langle\langle -a_i, -b_i \rangle\rangle \in I^2(k)$. Then

$$c(q) = [\otimes_{1 \leq i \leq n} H_i]$$

where $H_i = H(a_i, b_i)$.

Exercise 2.10. Given $\bigotimes_{1 \leq i \leq n} H_i$, a tensor product of n quaternion algebras over k, show that there is a quadratic form q over k **of dimension** $\mathbf{2n + 2}$ such that $c(q) = [\bigotimes_{1 \leq i \leq n} H_i]$.

The Clifford invariant induces a homomorphism $c\colon I^2(k) \to_2 \mathrm{Br}(k)$, ${}_2\mathrm{Br}(k)$ denoting the 2-torsion in the Brauer group of k. The very first case of the Milnor conjecture (see Sect. 3) states: c is surjective and $\ker(c) = I^3(k)$.

Theorem 2.11 (Merkurjev [M1]). *The map c induces an isomorphism*

$$I^2(k)/I^3(k) \cong_2 \mathrm{Br}(k)$$

Thus the image of $I^2(q)$ in ${}_2\mathrm{Br}(k)$ is spanned by quaternion algebras. It was a longstanding question whether ${}_2\mathrm{Br}(k)$ is spanned by quaternion algebras. Merkurjev's theorem answers this question in the affirmative; further, it gives precise relations between quaternion algebras in ${}_2\mathrm{Br}(k)$.

3 Galois Cohomology and the Milnor Conjecture

Let $\bar{k}$ be a separable closure of k. Let $\Gamma_k = \mathrm{Gal}(\bar{k}|k)$ be the absolute Galois group of k. The group Γ_k is a profinite group:

$$\Gamma_k = \varprojlim_{L \subset \bar{k},\ L/k \text{ finite Galois}} \mathrm{Gal}(L/k).$$

A *discrete* Γ_k*-module* M is a continuous Γ_k-module for the discrete topology on M and the profinite topology on Γ_k. A Γ_k-module M is discrete if and only if the stabilizer of each $m \in M$ is an open subgroup, in particular, of finite index in Γ_k. For a discrete Γ_k-module M, we define $H^n(k, M)$ as the direct limit of the cohomology of the finite quotients

$$H^n(k, M) = \varinjlim_{L \subset \bar{k},\, L/k \text{ finite Galois}} H^n(\mathrm{Gal}(L/k), M^{\Gamma_L}).$$

Suppose $\mathrm{char}(k) \neq 2$ and $M = \mu_2$. The module μ_2 has trivial Γ_k action and is isomorphic to $\mathbb{Z}/2\mathbb{Z}$. We have

$$H^0(k, \mathbb{Z}/2\mathbb{Z}) = \mathbb{Z}/2\mathbb{Z}$$
$$H^1(k, \mathbb{Z}/2\mathbb{Z}) \cong k^*/k^{*2}$$
$$H^2(k, \mathbb{Z}/2\mathbb{Z}) \cong_2 \mathrm{Br}(k)$$

These can be seen from the Kummer exact sequence of Γ_k-modules:

$$0 \longrightarrow \mu_2 \longrightarrow \bar{k}^* \xrightarrow{\cdot 2} \bar{k}^* \longrightarrow 0$$

and noting that $H^1(\Gamma_k, \bar{k}^*) = 0$ (Hilbert's Theorem 90) and $H^2(\Gamma_k, \bar{k}^*) = \mathrm{Br}(k)$.

For an element $a \in k^*$, we denote by (a) its class in $H^1(k, \mathbb{Z}/2\mathbb{Z})$ and for $a_1, \dots, a_n \in k^*$, the cup product $(a_1) \cup \cdots \cup (a_n) \in H^n(k, \mathbb{Z}/2\mathbb{Z})$ is denoted by $(a_1) \cdot \cdots \cdot (a_n)$.

For $a, b \in k^*$, the element $(a).(b)$ represents the class of $H(a, b)$ in ${}_2\mathrm{Br}(k)$. The map

$$c\colon I^2(k) \to H^2(k, \mathbb{Z}/2\mathbb{Z})$$

sends $\langle 1, -a, -b, ab\rangle$ to the class of $H(a, b)$ in $H^2(k, \mathbb{Z}/2\mathbb{Z})$. The forms $\langle 1, -a, -b, ab\rangle$ additively generate $I^2(k)$. Merkurjev's theorem asserts that $H^2(k, \mathbb{Z}/2\mathbb{Z})$ is generated by $(a).(b)$, with $a, b \in k^*$. The Milnor conjecture (quadratic form version) proposes higher invariants $I^n(k) \to H^n(k, \mathbb{Z}/2\mathbb{Z})$ extending the classical invariants.

Milnor Conjecture. *The assignment*

$$\langle 1, a_1\rangle \otimes \cdots \otimes \langle 1, a_n\rangle \mapsto (a_1) \cdot \cdots \cdot (a_n)$$

yields a map $e_n\colon P_n(k) \to H^n(k, \mathbb{Z}/2\mathbb{Z})$. *This map extends to a homomorphism* $e_n\colon I^n(k) \to H^n(k, \mathbb{Z}/2\mathbb{Z})$ *which is onto and* $\ker(e_n) = I^{n+1}(k)$.

The maps *dimension mod* 2, *discriminant* and *Clifford invariant* coincide with e_0, e_1 and e_2. Unlike these classical invariants, which are defined on all quadratic forms, conjecturally e_n, $n \geq 3$, are defined only on elements in $I^n(k)$ on which the invariants e_i, $i \leq n - 1$, vanish. In 1975, Arason [Ar] proved that $e_3\colon I^3(k) \to H^3(k, \mathbb{Z}/2\mathbb{Z})$ is well defined and is one-one on $P_3(k)$. As we mentioned earlier, the first nontrivial case of the Milnor conjecture was proved by Merkurjev for $n = 2$. The Milnor conjecture (quadratic form version) is now a theorem due to Orlov–Vishik–Voevodsky [OVV].

The Milnor conjecture gives a classification of quadratic forms by their Galois cohomology invariants: Given anisotropic quadratic forms q_1 and q_2, suppose $e_i(q_1 \perp -q_2) = 0$ for $i \geq 0$. Then $q_1 = q_2$ in $W(k)$. We need only to verify $e_i(q_1 \perp -q_2) = 0$ for $i \leq N$ where $N \leq 2^n$ and $\dim(q_1 \perp -q_2) \leq 2^n$, by the following theorem of Arason and Pfister.

Theorem 3.1 (Arason–Pfister Hauptsatz). *Let k be a field. The dimension of an anisotropic quadratic form in $I^n(k)$ is at least 2^n.*

4 Pfister Forms

The theory of Pfister forms (or multiplicative forms, as Pfister called them) evolved from questions on classification of quadratic forms whose nonzero values form a group (hereditarily).

Definition 4.1. A regular quadratic form q over k is called **multiplicative** if the nonzero values of q over any extension field L over k form a group.

We have the following examples of quadratic forms which are multiplicative.

Example 4.2. $\langle 1 \rangle$: nonzero squares are multiplicatively closed in k^*.

Example 4.3. $\langle 1, -a \rangle$: $x^2 - ay^2$, $a \in k^*$ is the norm from the quadratic algebra $k[t]/(t^2 - a)$ over k and the norm is multiplicative.

Example 4.4. $\langle 1, -a \rangle \otimes \langle 1, -b \rangle$: $x^2 - ay^2 - bz^2 + abt^2$ is a norm form from the quaternion algebra $H(a, b)$: $N(\alpha + i\beta + j\gamma + ij\delta) = \alpha^2 - a\beta^2 - b\gamma^2 + ab\delta^2$. The norm once again is multiplicative.

Example 4.5. $\langle 1, -a \rangle \otimes \langle 1, -b \rangle \otimes \langle 1, -c \rangle$: $(x^2 - ay^2 - bz^2 + abt^2) - c(u^2 - av^2 - bw^2 + abs^2)$ is the norm form from an octonion algebra associated to the triple (a, b, c); it is a non-associative algebra obtained from the quaternion algebra $H(a, b)$ by a doubling process (see [J, Sect. 7.6]). The norm is once again multiplicative.

Theorem 4.6 (Pfister). *An anisotropic quadratic form q over k is multiplicative if and only if q is isomorphic to a Pfister form.*

We shall sketch a proof of this theorem. The main ingredients are the Cassels–Pfister Theorem 4.7 and the Subform Theorem 4.10, which will not be proved in the text. We refer to [L, Chap. IX, Theorems 1.3 and 2.8] for the proofs.

Theorem 4.7 (Cassels–Pfister). *Let $q = \langle a_1, \ldots, a_n \rangle$ be a regular quadratic form over k and $f(X) \in k[X]$, a polynomial over k which is a value of q over $k(X)$. Then there exist polynomials $g_1, \ldots, g_n \in k[X]$ such that $f(X) = a_1 g_1^2 + \cdots + a_n g_n^2$.*

Corollary 4.8 (Specialization Lemma). *Let $q = \langle a_1, \ldots, a_n \rangle$ be a quadratic form over k, $X = \{X_1, \ldots, X_n\}$, $p(X) \in k(X)$ a rational function represented by q over $k(X)$. Then for any $v \in k^n$ where $p(v)$ is defined, $p(v)$ is represented by q over k.*

Proof. We may assume, by multiplying $p(X)$ by a square, that $p(X) \in k[X]$. Let $p(X) = p_1(X_n)$, where p_1 is a polynomial in X_n with coefficients in $k[X_1, \ldots, X_{n-1}]$. By the Cassels–Pfister theorem, $p_1(X_n)$ is represented by q over $k(X_1, \ldots, X_{n-1})[X_n]$. Let $v = (v_1, \ldots, v_n)$. Then specializing X_n to v_n, we have $p_1(v_n) \in k[X_1, \ldots, X_{n-1}]$ is represented by q over $k(X_1, \ldots, X_{n-1})$. By an induction argument, one concludes that $p(v_1, \ldots, v_n)$ is a value of q over k. □

Corollary 4.9. *Let q be an anisotropic quadratic form over k of dimension n. Then q is multiplicative if and only if, for indeterminates $X = (X_1, \ldots, X_n)$, $Y = (Y_1, \ldots, Y_n)$, $q(X)\,q(Y)$ is a value of q over $k(X_1, \ldots, X_n, Y_1, \ldots, Y_n)$.*

Proof. The only non-obvious part is "if". Suppose L/k is a field extension and $v, w \in L^n$. Let $q(v) = c$ and $q(w) = d$. Since $q(X)\,q(Y)$ is a value of q over $k(X, Y)$, by the Specialization lemma, $q(X)\,q(w)$ is a value of q over $L(X)$ and by the same lemma, $q(v)\,q(w)$ is a value of q over L. □

Theorem 4.10 (Subform Theorem). *Let $q = \langle a_1, \ldots, a_n \rangle$, $\gamma = \langle b_1, \ldots, b_m \rangle$ be quadratic forms over k with q anisotropic. Then γ is a subform of q (i.e., $q \cong \gamma \perp \gamma'$ for some form γ' over k) if and only if $b_1X_1^2 + \cdots + b_mX_m^2$ is a value of q over $k(X_1, \ldots, X_m)$.*

Corollary 4.11. *Let q be an anisotropic quadratic form over k of dimension n. Let $X = \{X_1, \ldots, X_n\}$ be a list of n indeterminates. Then q is multiplicative if and only if $q \cong q(X)\,q$ over $k(X)$.*

Proof. Suppose $q \cong q(X)\,q$ over $k(X)$. Let A be the matrix representing q over k. There exists $W \in \mathrm{GL}_n(k(X))$ such that $q(X)A = WAW^t$. Let $Y = \{Y_1, \ldots, Y_n\}$ be a list of n indeterminates. Over $k(X, Y)$,

$$q(X)\,q(Y) = Y(q(X)A)Y^t = (YW)A(YW)^t = q(Z)$$

where $Z = YW$. Thus $q(X)\,q(Y)$ is a value of q over $k(X, Y)$ and by Corollary 4.9, q is multiplicative.

Suppose conversely that q is multiplicative. Then $q(X)\,q(Y)$ is a value of q over $k(X, Y)$. By the Subform theorem, $q(X)\,q$ is a subform of q. A dimension count yields $q \cong q(X)\,q$. □

Proof of Pfister's Theorem 4.6. Let $q = \langle 1, a_1 \rangle \otimes \cdots \otimes \langle 1, a_n \rangle$ be an anisotropic quadratic form over k. Over any field extension L/k, either q is an anisotropic Pfister form or isotropic in which case it is universal. Thus it suffices to show that the nonzero values of q form a subgroup of k^* for any anisotropic n-fold Pfister form q. The proof is by induction on n; for $n = 1$, q is the norm form from a quadratic extension of k (see Example 4.3) and we are done. Let $n \geq 2$. We have $q \cong q_1 \perp a_nq_1$, where $q_1 = \langle 1, a_1 \rangle \otimes \cdots \otimes \langle 1, a_{n-1} \rangle$ is an anisotropic $(n-1)$-fold Pfister form. Let $X = \{X_1, \ldots, X_{2^{n-1}}\}$, $Y = \{Y_1, \ldots, Y_{2^{n-1}}\}$ be two lists of 2^{n-1} indeterminates. Since q_1 is multiplicative, by Corollary 4.11, $q_1(X)\,q_1 \cong q_1$ over $k(X)$ and $q_1(Y)\,q_1 \cong q_1$ over $k(Y)$. We have, over $k(X, Y)$,

$$q \cong q_1(X)\, q_1 \perp a_n q_1(Y)\, q_1 \cong \langle q_1(X), a_n q_1(Y)\rangle \otimes q_1 .$$

Since $q(X, Y) = q_1(X) + a_n q_1(Y)$, $\langle q_1(X), a_n q_1(Y)\rangle$ represents $q(X, Y)$. Therefore, by a comparison of discriminants,

$$\begin{aligned}\langle q_1(X), a_n q_1(Y)\rangle &\cong \langle q(X, Y), a_n q(X, Y) q_1(X) q_1(Y)\rangle \\ &\cong q(X, Y)(1 \perp a_n q_1(X) q_1(Y))\end{aligned}$$

In particular,

$$\begin{aligned}q &\cong q(X, Y)\langle 1, a_n q_1(X) q_1(Y)\rangle \otimes q_1 \\ &\cong q(X, Y)(q_1 \perp a_n q_1) \\ &\cong q(X, Y)\, q\end{aligned}$$

Thus by Corollary 4.11, q is multiplicative.

Conversely, let q be an anisotropic quadratic form over k which is multiplicative. Let n be the largest integer such that q contains an n-fold Pfister form $q_1 = \langle 1, a_1\rangle \otimes \cdots \otimes \langle 1, a_n\rangle$ as a subform. Suppose $q \cong q_1 \perp \gamma$, $\gamma = \langle b_1, \ldots, b_m\rangle$, with $m \geq 1$. Let $Z = \{Z_1, \ldots, Z_{2^n}\}$. Over $k(Z)$,

$$q \cong q(Z, 0)\, q \cong q_1(Z)(q_1 \perp \gamma) \cong q_1(Z)\, q_1 \perp q_1(Z)\, \gamma \cong q_1 \perp q_1(Z)\, \gamma .$$

By Witt's cancellation, $\gamma \cong q_1(Z)\, \gamma$ over $k(Z)$. Thus γ represents $b_1 q_1(Z)$ over $k(Z)$ and by the Subform theorem, $\gamma \cong b_1\, q_1 \perp \gamma_1$. Then $q \cong q_1 \perp b_1\, q_1 \perp \gamma_1 \cong \langle 1, b_1\rangle \otimes q_1 \perp \gamma_1$ contains an $(n + 1)$-fold Pfister form $\langle 1, b_1\rangle \otimes q_1$, leading to a contradiction to the maximality of n. Thus $q \cong q_1$. □

An important property of Pfister forms is stated in the following.

Proposition 4.12. *Let ϕ be an n-fold Pfister form. If ϕ is isotropic then ϕ is hyperbolic.*

Proof. Let $\phi = r\,\langle 1, -1\rangle \perp \phi_0$, with ϕ_0 anisotropic, $\dim(\phi_0) \geq 1$ and $r \geq 1$. Let $\dim(\phi) = m$ and $X = \{X_1, \ldots, X_m\}$ be a list of m indeterminates. Over $k(X_1, \ldots, X_m)$

$$r\,\langle 1, -1\rangle \perp \phi_0 = \phi \cong \phi(X_1, \ldots, X_m)\, \phi \cong r\,\langle 1, -1\rangle \perp \phi(X_1, \ldots, X_m)\, \phi_0 .$$

By Witt's cancellation theorem

$$\phi_0 \cong \phi(X_1, \ldots, X_m)\, \phi_0 .$$

If b is a value of ϕ_0, $b\phi(X_1, \ldots, X_m)$ is a value of ϕ_0 and by the Subform theorem, $b\phi$ is a subform of ϕ_0 contradicting $\dim(\phi_0) < \dim(\phi)$. Thus $\phi \cong r\,\langle 1, -1\rangle$ is hyperbolic. □

Corollary 4.13. *The only integers n such that a product of sums of n squares is again a sum of n squares over every field of characteristic zero are $n = 2^m$ for all $m \geq 0$.*

Proof. Consider the quadratic form $\phi_n = x_1^2 + x_2^2 + \cdots + x_n^2$ over $\mathbb{Q}$. The form ϕ_n is anisotropic. The condition that a product of sums of n squares is again a sum of n squares over any field of characteristic zero is equivalent to ϕ_n being a Pfister form. Thus $\dim(\phi_n) = n = 2^m$ for some m. □

5 Level of a Field

Definition 5.1. The **level** of a field k is the least positive integer n such that -1 is a sum of n squares in k. We denote the level of k by $s(k)$.

If the field is formally real (i.e., -1 is not a sum of squares), then the level is defined to be infinite. It was a longstanding open question since the 1950s whether the level of a field, if finite, is always a power of 2. Pfister's theory of quadratic forms leads to an affirmative answer to this question.

Theorem 5.2 ([Pf1]). *The level of a field is a power of 2 if it is finite.*

Proof. Let $n = s(k)$. We choose an integer m such that $2^m \leq n < 2^{m+1}$. Suppose

$$-1 = (u_1^2 + u_2^2 + \cdots + u_{2^m}^2) + (u_{2^m+1}^2 + \cdots + u_n^2) \tag{5.3}$$

The element $u_1^2 + u_2^2 + \cdots + u_{2^m}^2 \neq 0$ since $s(k) \geq 2^m$. Every ratio of sums of 2^m squares is again a sum of 2^m squares since $\langle 1, 1\rangle^{\otimes m}$ is a multiplicative form. Thus, from (5.3) we see that

$$\begin{aligned} 0 &= 1 + \frac{u_{2^m+1}^2 + \cdots + u_n^2 + 1}{u_1^2 + \cdots + u_{2^m}^2} \\ &= 1 + (v_1^2 + \cdots + v_{2^m}^2) \end{aligned}$$

Therefore, $-1 = v_1^2 + \cdots + v_{2^m}^2$ and $s(k) = 2^m$. □

Remark 5.4. There exist fields with level 2^n for any $n \geq 1$. For instance, $\mathbb{R}(X_1, \ldots, X_{2^n})(\sqrt{-(X_1^2 + \cdots + X_{2^n}^2)})$ is a field of level 2^n (cf. [L], Sect. XI.2).

Exercise 5.5. Let k be a p-adic field with $p \neq 2$ and with residue field $\mathbb{F}_q$. Prove the following:

(1) $s(k) = 1$ if $q \equiv 1 \pmod 4$.
(2) $s(k) = 2$ if $q \equiv -1 \pmod 4$.

6 The u-Invariant

Definition 6.1. The **u-invariant** of a field k, denoted by $u(k)$, is defined to be the largest integer n such that every $(n+1)$-dimensional quadratic form over k is isotropic and there is an anisotropic form in dimension n over k; if no such integer exists, the u-invariant is said to be infinite. In other words,

$$u(k) = \max\{\dim(q)\ :\ q \text{ anisotropic form over } k\}.$$

If k admits an ordering, then sums of nonzero squares are never zero and there is a refined u-invariant for fields with orderings, due to Elman–Lam [EL]. In this article, we do not discuss this refined invariant.

Example 6.2. (1) $u(\mathbb{F}_q) = 2$, if q is odd.
(2) $u(k(X)) = 2$, if k is algebraically closed and X is an integral curve over k (Tsen's theorem).
(3) $u(k) = 4$ for k a p-adic field. For $p \neq 2$, see Sect. 2.2. For $p = 2$, see [L, Sect. XI.6].
(4) $u(k) = 4$ for k a totally imaginary number field. This follows from the Hasse–Minkowski theorem.
(5) Suppose $u(k) = n < \infty$. Let $k((t))$ denote the field of Laurent series over k. Then $u(k((t))) = 2n$. In fact, the square classes in $k((t))^*$ are $\{u_\alpha, tu_\alpha\}_{\alpha\in I}$ where $\{u_\alpha\}_{\alpha\in I}$ are the square classes in k^*. As in the p-adic field case, every form over $k((t))$ is isometric to $\langle u_1,\dots,u_r\rangle \perp t\langle v_1,\dots,v_s\rangle$, $u_i, v_i \in k^*$ and this form is anisotropic if and only if $\langle u_1,\dots,u_r\rangle$ and $\langle v_1,\dots,v_s\rangle$ are anisotropic.
(6) More generally, if K is a complete discrete valuated field with residue field κ of u-invariant n, then $u(K) = 2n$. For the case $\operatorname{char}(\kappa) = 2$, we refer to [Sp].

Definition 6.3. A field k is $\boldsymbol{C_i}$ if every homogeneous polynomial in N variables of degree d with $N > d^i$ has a nontrivial zero.

Example 6.4. Finite fields and function fields in one variable over algebraically closed fields are C_1.

If k is a C_i field, $u(k) \leq 2^i$. Further, the property C_i behaves well with respect to function field extensions. If l/k is finite and k is C_i then l is C_i; further, if $t_1,\dots,t_n$ are indeterminates, $k(t_1,\dots,t_n)$ is C_{i+n}.

Example 6.5. The u-invariant of transcendental extensions:

(1) $u(k(t_1,\dots,t_n)) = 2^n$ if k is algebraically closed. In fact,

$$u(k(t_1,\dots,t_n)) \leq 2^n$$

since $k(t_1, \ldots, t_n)$ is a C_n field. Further, the form

$$\langle\langle t_1, \ldots, t_n \rangle\rangle = \langle 1, t_1 \rangle \otimes \cdots \otimes \langle 1, t_n \rangle$$

is anisotropic over $k((t_1))((t_2)) \ldots ((t_n))$ and hence also over $k(t_1, \ldots, t_n)$.

(2) $u(\mathbb{F}_q(t_1, \ldots, t_n)) = 2^{n+1}$ if q is odd.

All fields of known u-invariant in the 1950s happened to have u-invariant a power of 2. Kaplansky raised the question whether the u-invariant of a field is always a power of 2.

Proposition 6.6. *The u-invariant does not take the values* $3, 5, 7$.

Proof. Let q be an anisotropic form of dimension 3. By scaling, we may assume that $q \cong \langle 1, a, b \rangle$. Then the form $\langle 1, a, b, ab \rangle$ is anisotropic; if $\langle 1, a, b, ab \rangle$ is isotropic, it is hyperbolic and Witt's cancellation yields $\langle a, b, ab \rangle \cong \langle 1, -1, -1 \rangle$ which is isotropic and $q \cong a\langle a, b, ab \rangle$ is isotropic leading to a contradiction. Thus $u(k) \neq 3$.

Let $u(k) < 8$. Every three-fold Pfister form (which has dimension 8) is isotropic and hence hyperbolic. Thus $I^3(k)$ which is generated by three-fold Pfister forms is zero. Let $q \in I^2(k)$ be any quadratic form. For any $c \in k^*$, $\langle 1, -c \rangle\, q \in I^3(k)$ is zero and cq is Witt equivalent to q, hence isometric to q by Witt's cancellation. We conclude that every quadratic form whose class is in $I^2(k)$ is universal.

Suppose $u(k) = 5$ or 7. Let q be an anisotropic form of dimension $u(k)$. Since every form in dimension $u(k) + 1$ is isotropic, if $\operatorname{disc}(q) = d$, $q \perp -d$ is isotropic and therefore q represents d. We may write $q \cong q_1 \perp \langle d \rangle$ where q_1 is even-dimensional with trivial discriminant. Hence $[q_1] \in I^2(k)$ so that q_1 is universal. This in turn implies that $q_1 \perp \langle d \rangle \cong q$ is isotropic, leading to a contradiction. □

In the 1990s Merkurjev [M2] constructed examples of fields k with $u(k) = 2n$ for any $n \geq 1$, $n = 3$ being the first open case, answering Kaplansky's question in the negative. Since then, it has been shown that the u-invariant could be odd. In [I], Izhboldin proves that there exist fields k with $u(k) = 9$ and in [V] Vishik has shown that there exist fields k with $u(k) = 2^r + 1$ for all $r \geq 3$.

Merkurjev's construction yields fields k which are not of arithmetic type, i.e., not finitely generated over a number field or a p-adic field. It is still an interesting question whether $u(k)$ is a power of 2 if k is of arithmetic type.

The behavior of the u-invariant is very little understood under rational function field extensions. For instance, it is an open question if $u(k) < \infty$ implies $u(k(t)) < \infty$ for the rational function field in one variable over k. This was unknown for $k = \mathbb{Q}_p$ until the late 1990s. Conjecturally, $u(\mathbb{Q}_p(t)) = 8$, in analogy with the positive characteristic local field case; the field $\mathbb{F}_p((X))(t)$ is C_3 (see [G]) so that $u(\mathbb{F}_p((X))(t)) \leq 8$ for p odd. If u is a nonsquare in $\mathbb{F}_p$, $\langle 1, -u \rangle \otimes \langle 1, -X \rangle \otimes \langle 1, -t \rangle$ is anisotropic over $\mathbb{F}_p((X))(t)$, so that $u(\mathbb{F}_p((X))(t)) = 8$.

We indicate some ways of bounding the u-invariant of a field k once we know how efficiently the Galois cohomology groups $H^n(k, \mathbb{Z}/2\mathbb{Z})$ are generated by symbols for all n.

We set

$$H^n_{\text{dec}}(k, \mathbb{Z}/2\mathbb{Z}) = \{(a_1) \cdot \cdots \cdot (a_n) : a_i \in k^*\}$$

and call elements in this set *symbols*. By Voevodsky's theorem on the Milnor conjecture, $H^n(k, \mathbb{Z}/2\mathbb{Z})$ is additively generated by $H^n_{\text{dec}}(k, \mathbb{Z}/2\mathbb{Z})$.

Proposition 6.7. *Let k be a field such that $H^{n+1}(k, \mathbb{Z}/2\mathbb{Z}) = 0$ and for $2 \leq i \leq n$, there exist integers N_i such that every element in $H^i(k, \mathbb{Z}/2\mathbb{Z})$ is a sum of N_i symbols. Then $u(k)$ is finite.*

Proof. Let q be a quadratic form over k of dimension m and discriminant d. Let $q_1 = \langle d \rangle$ if m is odd and $\langle 1, -d \rangle$ if m is even. Then $q \perp -q_1$ has even dimension and trivial discriminant. Hence $q \perp -q_1 \in I^2(k)$. Let $e_2(q \perp -q_1) = \sum_{j \leq N_2} \xi_{2j}$ where $\xi_{2j} \in H^2_{\text{dec}}(k, \mathbb{Z}/2\mathbb{Z})$. Let ϕ_{2j} be two-fold Pfister forms such that $e_2(\phi_{2j}) = \xi_{2j}$. Then $q_2 = \sum_{j \leq N_2} \phi_{2j}$ has dimension at most $4N_2$ and $e_2(q \perp -q_1 \perp -q_2) = 0$ and $q \perp -q_1 \perp -q_2 \in I^3(k)$, by Merkurjev's theorem. Repeating this process and using the Milnor conjecture, we get $q_i \in I^i(k)$ which is a sum of N_i i-fold Pfister forms and $q - \sum_{1 \leq i \leq n} q_i \in I^{n+1}(k) = 0$, since $H^{n+1}(k, \mathbb{Z}/2\mathbb{Z}) = 0$. Thus $[q] = \sum_{1 \leq i \leq n} q_i$ and $\dim(q_{an}) \leq \sum_{1 \leq i \leq n} 2^i N_i$. Thus $u(k) \leq \sum_{1 \leq i \leq n} 2^i N_i$. □

Definition 6.8. A field k is said to have **cohomological dimension at most n** (in symbols, $\text{cd}(k) \leq n$) if $H^i(k, M) = 0$ for $i \geq n+1$ for all finite discrete Γ_k-modules M (cf. [Se, §3]).

Example 6.9. Finite fields and function fields in one variable over algebraically closed fields have cohomological dimension 1. Totally imaginary number fields and p-adic fields are of cohomological dimension 2. If k is a p-adic field, and $k(X)$ a function field in one variable over k, $\text{cd}(k(X)) \leq 3$. In particular, $H^4(k(X), \mathbb{Z}/2\mathbb{Z}) = 0$.

Theorem 6.10 (Saltman [Sa]). *Let k be a non-dyadic p-adic field and $k(X)$ a function field in one variable over k. Every element in $H^2(k(X), \mathbb{Z}/2\mathbb{Z})$ is a sum of two symbols.*

Theorem 6.11 (Parimala–Suresh [PS1]). *Let $k(X)$ be as in the previous theorem. Then every element in $H^3(k(X), \mathbb{Z}/2\mathbb{Z})$ is a symbol.*

Corollary 6.12. *For $k(X)$ as above, $u(k(X)) \leq 2 + 8 + 8 = 18$.*

It is not hard to show from the above theorems that $u(k(X)) \leq 12$. With some further work it was proved in [PS1] that $u(k(X)) \leq 10$. More recently in [PS2] the estimated value $u(k(X)) = 8$ was proved. For an alternate approach to $u(k(X)) = 8$, we refer to [HH, HHK, CTPS]. More recently, Heath-Brown and Leep [HB] have proved the following spectacular theorem: If k is *any* p-adic field and $k(X)$ the function field in n variables over k, then $u(k(X)) = 2^{n+2}$.

7 Hilbert's Seventeenth Problem

An additional reference for sums of squares is [C].

Definition 7.1. An element $f \in \mathbb{R}(X_1, \ldots, X_n)$ is called **positive semi-definite** if $f(a) \geq 0$ for all $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$ where f is defined.

Hilbert's seventeenth problem:
Let $\mathbb{R}(X_1, \ldots, X_n)$ be the rational function field in n variables over the field $\mathbb{R}$ of real numbers. Hilbert's seventeenth problem asks whether every positive semi-definite $f \in \mathbb{R}(X_1, \ldots, X_n)$ is a sum of squares in $\mathbb{R}(X_1, \ldots, X_n)$. E. Artin settled this question in the affirmative and Pfister gave an effective version of Artin's result (cf. [Pf, Chap. 6]).

Theorem 7.2 (Artin, Pfister). *Every positive semi-definite function $f \in \mathbb{R}(X_1, \ldots, X_n)$ can be written as a sum of 2^n squares in $\mathbb{R}(X_1, \ldots, X_n)$.*

For $n \leq 2$ the above was due to Hilbert himself. If one asks for expressions of positive definite polynomials in $\mathbb{R}[X_1, \ldots, X_n]$ as sums of 2^n squares in $\mathbb{R}[X_1, \ldots, X_n]$, there are counterexamples for $n = 2$; the Motzkin polynomial

$$f(X_1, X_2) = 1 - 3X_1^2X_2^2 + X_1^4X_2^2 + X_1^2X_2^4$$

is positive semi-definite but not a sum of squares in $\mathbb{R}[X_1, X_2]$. In fact, Pfister's result has the following precise formulation.

Theorem 7.3 (Pfister). *Let $\mathbb{R}(X)$ be a function field in n variables over $\mathbb{R}$. Then every n-fold Pfister form in $\mathbb{R}(X)$ represents every sum of squares in $\mathbb{R}(X)$.*

We sketch a proof of this theorem below.

Definition 7.4. Let ϕ be an n-fold Pfister form with $\phi = 1 \perp \phi'$. The form ϕ' is called the **pure subform** of ϕ.

Proposition 7.5 (Pure Subform Theorem). *Let k be any field of characteristic not 2, ϕ an anisotropic n-fold Pfister form over k and ϕ' its pure subform. If b_1 is any value of ϕ', then $\phi \cong \langle\langle b_1, \ldots, b_n \rangle\rangle$ for some $b_2, \ldots, b_n \in k^*$.*

Proof. The proof is by induction on n; for $n = 1$ the statement is clear. Let $n > 1$. We assume the statement holds for all $(n-1)$-fold Pfister forms. Let $\phi = \langle\langle a_1, \ldots, a_n \rangle\rangle$, $\psi = \langle\langle a_1, \ldots, a_{n-1} \rangle\rangle$, and let ϕ', ψ' denote the pure subforms of ϕ and ψ respectively. We have $\phi = \psi \perp a_n\psi$, $\phi' = \psi' \perp a_n\psi$. Let b_1 be a value of ϕ'. We may write $b_1 = b_1' + a_nb$, with b_1' a value of ψ' and b a value of ψ. The only nontrivial case to discuss is when $b \neq 0$ and $b_1' \neq 0$. By induction, $\psi \cong \langle\langle b_1', b_2, \ldots, b_{n-1} \rangle\rangle$ and $b\psi \cong \psi$. We thus have

$$\begin{aligned} \phi &\cong \langle\langle b_1', b_2, \ldots, b_{n-1}, a_n \rangle\rangle \cong \langle\langle b_1', b_2, \ldots, b_{n-1}, a_nb \rangle\rangle \\ &\cong \langle\langle b_1', a_nb \rangle\rangle \otimes \langle\langle b_2, \ldots, b_{n-1} \rangle\rangle \end{aligned}$$

Since $b_1 = b_1' + a_n b$, $\langle b_1', a_n b\rangle \cong \langle b_1, b_1 b_1' a_n b\rangle$ and we have

$$\begin{aligned}\langle\langle b_1', a_n b\rangle\rangle &= \langle 1, b_1', a_n b, a_n b b_1'\rangle \\ &= \langle 1, b_1, b_1 b_1' a_n b, a_n b b_1'\rangle \\ &= \langle\langle b_1, c_1\rangle\rangle,\end{aligned}$$

where $c_1 = b_1 b_1' a_n b$. Thus,

$$\phi \cong \langle\langle b_1, c_1, b_2, \cdots, b_{n-1}\rangle\rangle. \qquad \square$$

Proof of Pfister's Theorem 7.3. Let ϕ be an anisotropic n-fold Pfister form over $K = \mathbb{R}(X)$. Let $b = b_1^2 + \cdots + b_m^2$, $b_i \in K^*$. We show that ϕ represents b by induction on m. For $m = 1$, b is a square and is represented by ϕ. Suppose $m = 2$, $b = b_1^2 + b_2^2$, $b_1 \neq 0, b_2 \neq 0$. The field $K(\sqrt{-1})$ is a function field in n variables over $\mathbb{C}$ and is C_n. Then ϕ is universal over $K(\sqrt{-1})$ and hence represents $\beta = b_1 + ib_2$. Let $v, w \in K^{2^n}$ such that $\phi_{K(\sqrt{-1})}(v + \beta w) = \beta$. Hence

$$\phi(v) + \beta^2\phi(w) + \beta(2\phi(v, w) - 1) = 0.$$

The irreducible polynomial of β over K is

$$\phi(w)X^2 + (2\phi(v, w) - 1)X + \phi(v)$$

and hence $N(\beta) = b = \frac{\phi(v)}{\phi(w)}$ is a value of ϕ since ϕ is multiplicative.

Suppose $m > 2$. We argue by induction on m. Suppose ϕ represents all sums of $m - 1$ squares. Let b be a sum of m squares. After scaling b by a square, we may assume that $b = 1 + c$, $c = c_1^2 + \cdots + c_{m-1}^2$, $c \neq 0$. Let $\phi \cong 1 \perp \phi'$. By induction hypothesis, ϕ represents c. Let $c = c_0^2 + c'$, c' a value of ϕ'. Let $\psi = \phi \otimes \langle 1, -b\rangle$ and $\psi = 1 \perp \psi'$ with $\psi' = \langle -b\rangle \perp \phi' \perp -b\phi'$. The form ψ' represents $c' - b = (c - c_0^2) - (1 + c) = -1 - c_0^2$. Thus, by the Pure Subform theorem,

$$\psi \cong \langle\langle -1 - c_0^2, d_1, \ldots, d_n\rangle\rangle = \langle 1, -1 - c_0^2\rangle \otimes \langle\langle d_1, \ldots, d_n\rangle\rangle.$$

By induction, the n-fold Pfister form $\langle\langle d_1, \ldots, d_n\rangle\rangle$ represents $1 + c_0^2$ which is a sum of 2 squares; thus ψ is isotropic, hence hyperbolic. Thus $\phi \cong b\phi$ represents b. □

Corollary 7.6. *Let $K = \mathbb{R}(X)$ be a function field in n variables over $\mathbb{R}$. Then every sum of squares in K is a sum of 2^n squares.*

Proof. Set $\phi = \langle 1, 1\rangle^{\otimes n}$ in the above theorem. □

8 Pythagoras Number

Definition 8.1. The **Pythagoras number** $p(k)$ of a field k is the least positive integer n such that every sum of squares in k^* is a sum of at most n squares; if no such n exists, $p(k)$ is defined to be infinity.

Example 8.2. If $\mathbb{R}$ is the field of real numbers, $p(\mathbb{R}) = 1$.

Example 8.3. If $\mathbb{R}(X_1, \dots, X_n)$ is a function field in n variables over $\mathbb{R}$, by Pfister's theorem (Corollary 7.6), $p(\mathbb{R}(X_1, \dots, X_n)) \leq 2^n$.

Let $K = \mathbb{R}(X_1, \dots, X_n)$ be the rational function field in n variables over $\mathbb{R}$. We discuss the effectiveness of the bound $p(K) \leq 2^n$. For $n = 1$ the bound is sharp. For $n = 2$ the Motzkin polynomial

$$f(X_1, X_2) = 1 - 3X_1^2X_2^2 + X_1^4X_2^2 + X_1^2X_2^4$$

is positive semi-definite; Cassels–Ellison–Pfister [CEP] show that this polynomial is not a sum of three squares in $\mathbb{R}(X_1, X_2)$ (see also [CT]). Therefore $p(\mathbb{R}(X_1, X_2)) = 4$.

Lemma 8.4 (Key Lemma). *Let k be a field and $n = 2^m$. Let $u = (u_1, \dots, u_n)$ and $v = (v_1, \dots, v_n) \in k^n$ be such that $u \cdot v = \sum_{1 \leq i \leq n} u_iv_i = 0$. Then there exist $w_j \in k$, $1 \leq j \leq n-1$ such that*

$$\Bigg(\sum_{1 \leq i \leq n} u_i^2\Bigg)\Bigg(\sum_{1 \leq i \leq n} v_i^2\Bigg) = \sum_{1 \leq j \leq n-1} w_j^2.$$

Proof. Let $\lambda = \sum_{1 \leq i \leq n} u_i^2$, $\mu = \sum_{1 \leq i \leq n} v_i^2$. We may assume without loss of generality that $u \neq 0$ and $v \neq 0$. The elements λ and μ are values of $\phi_m = \langle 1, 1\rangle^{\otimes m}$ and $\lambda\phi_m \cong \phi_m$, $\mu\phi_m \cong \phi_m$. We choose isometries $f\colon \lambda\phi_m \cong \phi_m$, $g\colon \mu\phi_m \cong \phi_m$ such that $f(1, 0, \dots, 0) = u$ and $g(1, 0, \dots, 0) = v$. If U and V are matrices representing f, g respectively, we have

$$UU^t = \lambda^{-1},\ \ VV^t = \mu^{-1},\ \ \lambda^{-1}\mu^{-1} = \lambda^{-1}VV^t = (VU^t)(VU^t)^t.$$

The first row of VU^t is of the form $(0, w_2, \dots, w_n)$ since $u \cdot v = 0$. Thus $\lambda^{-1}\mu^{-1} = \sum_{2 \leq i \leq n} w_i^2$. □

Corollary 8.5. *Let k be an ordered field with $p(k) = n$. Then $p(k(t)) \geq n + 1$.*

Proof. Let $\lambda \in k^*$ be such that λ is a sum of n squares and not a sum of less than n squares. Suppose $\lambda + t^2$ is a sum of n squares in $k(t)$. By the Cassels–Pfister theorem,

$$\lambda + t^2 = (\mu_1 + \nu_1 t)^2 + \cdots + (\mu_n + \nu_n t)^2$$

with $\mu_i, \nu_i \in k^*$. If $u = (\mu_1, \ldots, \mu_n)$, $v = (\nu_1, \ldots, \nu_n)$, then $u \cdot v = 0$, $\sum_{1 \leq i \leq n} \mu_i^2 = \lambda$, $\sum_{1 \leq i \leq n} \nu_i^2 = 1$. Thus $\lambda = (\sum_{1 \leq i \leq n} \mu_i^2)(\sum_{1 \leq i \leq n} \nu_i^2)$ is a sum of $n - 1$ squares by the Key Lemma 8.4, contradicting the choice of λ. □

Corollary 8.6. *For $n \geq 2$,*

$$n + 2 \leq p(\mathbb{R}(X_1, \ldots, X_n)) \leq 2^n.$$

Proof. By [CEP], we know that $p(\mathbb{R}(X_1, X_2)) = 4$. The fact that $n + 2 \leq p(\mathbb{R}(X_1, \ldots, X_n))$ now follows by Corollary 8.5 and induction. □

Remark 8.7. It is open whether $p(\mathbb{R}(X_1, X_2, X_3)) = 5, 6, 7$ or 8.

Remark 8.8. The possible values of the Pythagoras number of a field have all been listed ([H], [Pf, p. 97]).

Proposition 8.9. *If k is a non-formally real field, $p(k) = s(k)$ or $s(k) + 1$.*

Proof. If $s(k) = n$, then -1 is not a sum of less than n squares, so that $p(k) \geq s(k)$. For $a \in k^*$,

$$a = \left(\frac{a+1}{2}\right)^2 + (-1)\left(\frac{a-1}{2}\right)^2$$

is a sum of $n + 1$ squares if -1 is a sum of n squares. Thus $p(k) \leq s(k) + 1$. □

Let k be a p-adic field and $K = k(X_1, \ldots, X_n)$ a rational function field in n variables over k. Then $s(k) = 1, 2$ or 4 so that $s(K) = 1, 2$ or 4. Thus $p(K) \leq 5$. (In fact it is easy to see that if $s(k) = s$, $p(K) = s + 1$.)

Thus we have bounds for $p(k(X_1, \ldots, X_n))$ if k is the field of real or complex numbers or the field of p-adic numbers. The natural questions concern a number field k.

9 Function Fields Over Number Fields

Let k be a number field and $F = k(t)$ the rational function field in one variable over k. In this case $p(k(t)) = 5$ is a theorem [La]. The fact that $p(k(t)) \leq 8$ can be easily deduced from the following injectivity in the Witt groups [CTCS, Proposition 1.1]:

$$W(k(t)) \longrightarrow \prod_{w \in \Omega(k)} W(k_w(t)),$$

with $\Omega(k)$ denoting the set of places of k. In fact, if $f \in k(t)$ is a sum of squares, f is a sum of at most two squares in $k_w(t)$ for a real place w, by Pfister's theorem (which in the case of function fields of curves goes back to Witt). Further, for a finite place w of k or a complex place, $\langle 1, 1 \rangle^{\otimes 3} = 0$ in $W(k_w)$. Thus $\langle 1, 1 \rangle^{\otimes 3} \otimes \langle 1, -f \rangle$ is hyperbolic over $k_w(t)$ for all $w \in \Omega(k)$.

By the above injectivity, this form is hyperbolic over $k(t)$, leading to the fact that f is a sum of at most eight squares in $k(t)$.

We have the following conjecture due to Pfister for function fields over number fields.

Conjecture (Pfister). *Let k be a number field and $F = k(X)$ a function field in d variables over k. Then*

(1) *For $d = 1$, $p(F) \leq 5$.*
(2) *For $d \geq 2$, $p(F) \leq 2^{d+1}$.*

For a function field $k(X)$ in one variable over k, ($d = 1$), the best known result is due to F. Pop, $p(F) \leq 6$ [P]. For $d = 2$, the conjecture is settled in [CTJ]. We sketch some results and conjectures from the arithmetic side which imply Pfister's conjecture for $d \geq 3$ (see Colliot-Thélène and Jannsen [CTJ] for more details).

For any field k, by Voevodsky's theorem, we have an injection

$$e_n \colon P_n(k) \to H^n(k, \mathbb{Z}/2\mathbb{Z}).$$

In fact, for any field k, if ϕ_1, $\phi_2 \in P_n(k)$ have the same image under e_n then $\phi_1 \perp -\phi_2 \in \ker(e_n) = I^{n+1}(k)$. In $W(k)$, $\phi_1 \perp -\phi_2 = \phi_1' \perp -\phi_2'$ where ϕ_1' and ϕ_2' are the pure subforms of ϕ_1 and ϕ_2. Moreover, $\dim(\phi_1' \perp -\phi_2')_{\mathrm{an}} \leq 2^{n+1} - 2 < 2^{n+1}$. By the Arason–Pfister Hauptsatz, (Theorem 3.1), anisotropic forms in $I^{n+1}(k)$ must have dimension at least 2^{n+1}. Therefore $\phi_1 = \phi_2$.

Let k be a number field and $F = k(X)$ be a function field in d variables over k. Let $f \in F$ be a function which is a sum of squares in F. One would like to show that f is a sum of 2^{d+1} squares. Let $\phi_{d+1} = \langle 1, 1\rangle^{\otimes(d+1)}$ and $q = \phi_{d+1} \otimes \langle 1, -f\rangle$. This is a $(d+2)$-fold Pfister form and ϕ_{d+1} represents f if and only if q is hyperbolic or equivalently, by the injectivity of e_n above, $e_{d+2}(\phi_{d+1} \otimes \langle 1, -f\rangle) = 0$.

We look at this condition locally at all completions k_v at places v of k. Let $k_v(X)$ denote the function field of X over k_v. (We may assume that X is geometrically integral.) Let v be a complex place. The field $k_v(X)$ has cohomological dimension d so that $H^m(k_v(X), \mathbb{Z}/2\mathbb{Z}) = 0$ for $m \geq d + 1$. Hence $e_{d+2}(\phi_{d+1} \otimes \langle 1, -f\rangle) = 0$ over $k_v(X)$. Let v be a real place. Over $k_v(X)$, f is a sum of squares, hence a sum of at most 2^d squares (by Pfister's Theorem 7.3) so that $\phi_{d+1} \otimes \langle 1, -f\rangle$ is hyperbolic over $k_v(X)$. Hence $e_{d+2}(\phi_{d+1} \otimes \langle 1, -f\rangle) = 0$.

Let v be a non-dyadic p-adic place of k. Then ϕ_2 is hyperbolic over k_v so that $\phi_{d+1} \otimes \langle 1, -f\rangle = 0$ and $e_{d+2}(\phi_{d+1} \otimes \langle 1, -f\rangle) = 0$.

Let v be a dyadic place of k. Over k_v, ϕ_3 is hyperbolic so that $e_{d+2}(\phi_{d+1} \otimes \langle 1, -f\rangle) = 0$. Thus for all completions v of k, $e_{d+2}(\phi_{d+1} \otimes \langle 1, -f\rangle)$ is zero. The following conjecture of Kato implies Pfister's conjecture for $d \geq 2$.

Conjecture (Kato). *Let k be a number field, X a geometrically integral variety over k of dimension d. Then the map*

$$H^{d+2}(k(X), \mathbb{Z}/2\mathbb{Z}) \to \prod_{v \in \Omega_k} H^{d+2}(k_v(X), \mathbb{Z}/2\mathbb{Z})$$

has trivial kernel.

The above conjecture is the classical Hasse–Brauer–Noether theorem if the dimension of X is zero, i.e., the injectivity of the Brauer group map:

$$\mathrm{Br}(k) \hookrightarrow \bigoplus_{v \in \Omega_k} \mathrm{Br}(k_v).$$

For $\dim X = 1$, the conjecture is a theorem of Kato [Ka]. Kato's conjecture is now a theorem due to Jannsen [Ja1, Ja2] for $\dim X \geq 2$. Thus for every function field $k(X)$ in d variables over a number field k, $d \geq 2$, we have $p(k(X)) \leq 2^{d+1}$.

We now explain how Kato's theorem was used by Colliot-Thélène to derive $p(k(X)) \leq 7$ for a curve X over a number field. We note that this bound is weaker than the bound established by F. Pop.

Suppose $K = k(X)$ has no ordering. We claim that $s(K) \leq 4$. To show this it suffices to show that $\langle 1,1\rangle^{\otimes 3}$ is zero over $k_v(X)$ for every place v of k. At finite places v, $\langle 1,1\rangle^{\otimes 3}$ is already zero in k_v. If v is a real place of k, $k_v(X)$ is the function field of a real curve over the field of real numbers which has no orderings. By a theorem of Witt, $\mathrm{Br}(k_v(X)) = 0$ and every sum of squares is a sum of two squares in $k_v(X)$. Thus -1 is a sum of two squares in $k_v(X)$ and $\langle 1,1\rangle^{\otimes 3} = 0$ over $k_v(X)$. Since $H^3(k(X), \mathbb{Z}/2\mathbb{Z}) \to \prod_{v\in\Omega_k} H^3(k_v(X), \mathbb{Z}/2\mathbb{Z})$ is injective by Kato's theorem, $e_3(\langle 1,1\rangle^{\otimes 3}) = 0$ in $H^3(k(X), \mathbb{Z}/2\mathbb{Z})$. Since e_3 is injective on three-fold Pfister forms, $\langle 1,1\rangle^{\otimes 3} = 0$ in $k(X)$. Thus $s(k(X)) \leq 4$. In this case, $p(k(X)) \leq 5$.

Suppose K has an ordering. Let $f \in K^*$ be a sum of squares in K. Then $K(\sqrt{-f})$ has no orderings and hence -1 is a sum of 4 squares in $K(\sqrt{-f})$. Let $a_i, b_i \in K$ be such that

$$-1 = \sum_{1\leq i\leq 4} (a_i + b_i\sqrt{-f})^2, \quad a_i, b_i \in K.$$

Then

$$1 + \sum_{1\leq i\leq 4} a_i^2 = f\Big(\sum_{1\leq i\leq 4} b_i^2\Big), \quad \sum_{1\leq i\leq 4} a_i b_i = 0.$$

By the Key Lemma 8.4, $(1 + \sum_{1\leq i\leq 4} a_i^2)\sum_{1\leq i\leq 4} b_i^2$ is a sum of at most 7 squares.

References

[Ar] Arason, J.K., *Cohomologische Invarianten quadratischer Formen, J. Algebra*, **36** (1975), 448–491.

[C] Cohen, H., *Représentations comme sommes de carrés, Séminaire de Théorie des Nombres, 1971–1972 (Univ. Bordeaux I, Talence), Exp. No. 21*, (1972). English translation available at http://www.math.u-bordeaux1.fr/~cohen/Cohensquares.pdf

[CEP] Cassels, J.W.S., Ellison, W.J., Pfister, A., *On sums of squares and on elliptic curves over function fields, J. Number Th.* **3** (1971), 125–149.

[CT] Colliot-Thélène, J.-L., *The Noether-Lefschetz theorem and sums of 4 squares in the rational function field* $\mathbb{R}(x,y)$*, Compos.Math.* **86** (1993), 235–243.

[CTCS] Colliot-Thélène, J.-L., Coray, D., Sansuc, J.-J., *Descente et principe de Hasse pour certaines variétés rationnelles, J. Reine Angew. Math.*, **320** (1980), 150–191.

[CTJ] Colliot-Thélène, J.-L., Jannsen, U., *Sommes de carrès dans les corps de fonctions, C. R. Acad. Paris, Sèr. I* **312** (1991), 759–762.

[CTPS] Colliot-Thélène, J.-L., Parimala, R., Suresh, V., *Patching and local-global principles for homogeneous spaces over function fields of p-adic curves*, preprint arXiv:0812.3099.

[EKM] Elman, R., Karpenko, N., Merkurjev, A., *The algebraic and geometric theory of quadratic forms, American Mathematical Society Colloquium Publications*, **56** (2008).

[EL] Elman, R. and Lam, T.-Y., *Quadratic forms and the u-invariant I, Math. Z.* **131** (1973), 238–304.

[G] Greenberg, M.J., *Rational points in Henselian discrete valuation rings, Inst. Hautes Études Sci. Publ. Math.*, **31** (1966), 59–64.

[H] Hoffmann, D.W., *Pythagoras numbers of fields, J. Amer. Math. Soc.*, **12** (1999), 3, 839–848.

[HB] Heath-Brown, D.R., *Artin's Conjecture on Zeros of p-Adic Forms*, Proceedings of the International Congress of Mathematicians, Volume II, 249–257, Hindustan Book Agency, New Delhi, 2010.

[HH] Harbater, D., Hartmann, J., *Patching over fields, Israel J. Math.*, **176** 2010, 61–107.

[HHK] Harbater, D., Hartmann, J., Krashen, D., *Applications of patching to quadratic forms and central simple algebras, Invent. Math.*, **178** 2009, 2, 231–263.

[I] Izhboldin, Oleg T., *Fields of u-invariant* 9. *Ann. of Math. (2)* **154** (2001).

[J] Jacobson, N., *Basic algebra I, W. H. Freeman and Company* (1985).

[Ja1] Jannsen, U., *Principe de Hasse cohomologique, Séminaire de Théorie des Nombres, Paris, 1989–90, Progr. Math., Birkhäuser Boston*, **102** 121–140 (1992).

[Ja2] Jannsen, U., *Hasse principles for higher-dimensional fields*, preprint arXiv:0910.2803v1.

[K] Kahn, B., *Formes quadratiques sur un corps, Cours Spécialisés, Société Mathématique de France* **15** 2008.

[Ka] Kato, Kazuya, *A Hasse principle for two-dimensional global fields, with an appendix by Jean-Louis Colliot-Thélène, J. Reine Angew. Math.* **366** (1986).

[L] Lam, T.-Y., *Introduction to quadratic forms over fields, GSM, AMS* **67** (2004).

[La] Landau, E., *Über die Darstellung definiter Funktionen durch Quadrate, Math. Ann.*, **62** (1906), 2, 272–285.

[M1] Merkurjev, A.S. *On the norm residue symbol of degree* 2*, Dokl. Akad. Nauk SSSR* **261** (1981), 542–547.

[M2] Merkurjev, A.S., *Simple algebras and quadratic forms, Izv. Akad. Nauk SSSR Ser. Mat.* **55** (1991), 218–224; *translation in Math. USSR-Izv. 38 (1992), 1, 215–221.*

[OVV] Orlov, D., Vishik, A., and Voevodsky, V., *An exact sequence for $K_*^M/2$ with applications to quadratic forms, Annals of Math.* **165** (2007), 1–13.

[P] Pop, F., *Summen von Quadraten in arithmetischen Funktionenkorpern*, preprint (http://www.math.upenn.edu/~pop/Research/Papers.html).

[Pf1] Pfister, A., *Zur Darstellung von ? 1 als Summe von Quadraten in einem Krper*, *J. London Math. Soc.* **40** (1965), 159–165.

[Pf] Pfister, A., *Quadratic forms with applications to algebraic geometry and topology, London Mathematical Society Lecture Note Series, Cambridge University Press*, **217** (1995).

[PS1] Parimala, R. and Suresh, V., *Isotropy of quadratic forms over function fields in one variable over p-adic fields, Publ. de I.H.É.S.* **88** (1998), 129–150.

[PS2] Parimala, R. and Suresh, V., *The u-invariant of the function fields of p-adic curves, Ann. of Math. (2)* **172** (2010), 2, 1391–1405.

[S] Scharlau, W., *Quadratic and Hermitian forms, Springer-Verlag* **270** (1985).

[Sa] Saltman, D., *Division algebras over p-adic curves with an appendix by William Jacob and J.-P. Tignol, Journal of Ramanujan Math. Soc.* **12** (1997), 25–47.

[Se] Serre, J-P., *Galois cohomology, Springer-Verlag* (1997).

[Sp] Springer, T. A., *Quadratic forms over fields with a discrete valuation. I. Equivalence classes of definite forms, Nederl. Akad. Wetensch. Proc. Ser. A.* ***58*** *= Indag. Math.* **17** (1955), 352–362.

[V] Vishik, A., *Fields of u-invariant* $2^r + 1$. *Algebra, arithmetic and geometry: in honor of Yu. I. Manin. Vol. II, Progr. Math., Birkhäuser Boston*, **270** 661–685 (2009).

On the Length of Binary Forms

Bruce Reznick

Abstract The K-length of a form f in $K[x_1,\dots,x_n]$, $K \subset \mathbb{C}$, is the smallest number of d-th powers of linear forms of which f is a K-linear combination. We present many results, old and new, about K-length, mainly for $n = 2$, and often about the length of the same form over different fields. For example, the K-length of $3x^5 - 20x^3y^2 + 10xy^4$ is three for $K = \mathbb{Q}(\sqrt{-1})$, four for $K = \mathbb{Q}(\sqrt{-2})$ and five for $K = \mathbb{R}$.

Key words and Phrases Length • Sums of powers • Binary forms • Sylvester

Mathematics Subject Classification (2000): Primary: 11E76, 11P05, 14N10

1 Introduction and Overview

Suppose $f(x_1,\dots,x_n)$ is a form of degree d with coefficients in a field $K \subseteq \mathbb{C}$. The *K-length of f*, $L_K(f)$, is the smallest r for which there exist $\lambda_j, \alpha_{jk} \in K$ so that

$$f(x_1,\dots,x_n) = \sum_{j=1}^{r} \lambda_j \left(\alpha_{j1}x_1 + \cdots + \alpha_{jn}x_n\right)^d. \tag{1.1}$$

In this paper, we consider the K-length of a fixed form f as K varies; this is apparently an open question in the literature, even for binary forms ($n = 2$).

B. Reznick (✉)
Department of Mathematics, University of Illinois at Urbana-Champaign,
Urbana, IL 61801, USA
e-mail: reznick@math.uiuc.edu

K. Alladi et al. (eds.), *Quadratic and Higher Degree Forms*, Developments
in Mathematics 31, DOI 10.1007/978-1-4614-7488-3_8,

Sylvester [53, 54] explained how to compute $L_{\mathbb{C}}(f)$ for binary forms in 1851 and gave a lower bound for $L_{\mathbb{R}}(f)$ for binary forms in 1864. Except for a few remarks, we shall restrict our attention to binary forms.

It is trivially true that $L_K(f) = 1$ for linear f and for $d = 2$, $L_K(f)$ equals the rank of f: a representation over K can be found by completing the square, and this length cannot be shortened by enlarging the field. Accordingly, we shall also assume that $d \ge 3$.

When $K = \mathbb{C}$, the λ_j's in (1.1) are superfluous. The computation of $L_{\mathbb{C}}(f)$ is a huge, venerable and active subject, and difficult when $n \ge 3$. The interested reader is directed to [2, 7, 8, 14, 17, 18, 22, 25, 28, 34, 44–46] as representative recent works. Even for small $n, d \ge 3$, there are still many open questions. Landsberg and Teitler [34] complete a classification of $L_{\mathbb{C}}(f)$ for ternary cubics f and also discuss $L_{\mathbb{C}}(x_1 x_2 \cdots x_n)$, among other topics. Historically, much attention has centered on the $\mathbb{C}$-length of a *general* form of degree d. In 1995, Alexander and Hirschowitz [1] (see also [5, 36]) established that for $n, d \ge 3$, this length is $\lceil \frac{1}{n}\binom{n+d-1}{n-1} \rceil$, the constant-counting value, with the four exceptions known since the nineteenth century – $(n, d) = (3, 5), (4, 3), (4, 4), (4, 5)$ – in which the length is $\lceil \frac{1}{n}\binom{n+d-1}{n-1} \rceil + 1$. There has been a recent series of papers studying $L_{\mathbb{R}}(f)$ [4, 9, 15]; these study the length in a greater depth than we do here.

Two central examples illustrate the phenomenon of multiple lengths over different fields.

Example 1.1. Suppose $f(x, y) = (x + \sqrt{2}y)^d + (x - \sqrt{2}y)^d \in \mathbb{Q}[x, y]$. Then $L_K(f)$ is 2 (if $\sqrt{2} \in K$) and d (otherwise). This example first appeared in [47, p. 137]. (See Theorem 4.6 for a generalization.)

Example 1.2. If $\phi(x, y) = 3x^5 - 20x^3y^2 + 10xy^4$, then $L_K(\phi) = 3$ if and only if $\sqrt{-1} \in K$, $L_K(\phi) = 4$ for $K = \mathbb{Q}(\sqrt{-2}), \mathbb{Q}(\sqrt{-3}), \mathbb{Q}(\sqrt{-5}), \mathbb{Q}(\sqrt{-6})$ (at least) and $L_{\mathbb{R}}(\phi) = 5$. (We give proofs of these assertions in Examples 2.1 and 3.1.)

The following simple definitions and remarks apply in the obvious way to forms in $n \ge 3$ variables, but for simplicity are given for binary forms. A representation such as (1.1) is called *K-minimal* if $r = L_K(f)$. Two linear forms are called *distinct* if they (or their d-th powers) are not proportional. A representation is *honest* if the summands are pairwise distinct. Any minimal representation is honest. Two honest representations are *different* if the ordered sets of summands are not rearrangements of each other; we shall not distinguish between ℓ^d and $(\zeta_d^k \ell)^d$ where $\zeta_d = e^{2\pi i/d}$. If g is obtained from f by an invertible linear change of variables over K, then $L_K(f) = L_K(g)$.

Given a form $f \in \mathbb{C}[x, y]$, we let E_f denote the field generated by the coefficients of f over $\mathbb{C}$; $L_K(f)$ is defined for fields K satisfying $E_f \subseteq K \subseteq \mathbb{C}$. The following implication is immediate:

$$K_1 \subset K_2 \implies L_{K_1}(f) \ge L_{K_2}(f). \tag{1.2}$$

Strict inequality in (1.2) is possible, as shown by the two examples. Finally, we define the *cabinet* of f, $\mathcal{C}(f)$, to be the set of all possible lengths for f.

There is a natural alternative definition of length in which sums of powers are considered without coefficients. This makes no difference when $K = \mathbb{C}$ or $K = \mathbb{R}$ and d is odd, but in other cases, a form might not even be a sum of powers. For example, $\sqrt{2}$ is not totally positive in $K = \mathbb{Q}(\sqrt{2})$, so $\sqrt{2} \cdot x^2$ is not a sum of squares in $K[x]$, and $x^4+\lambda x^2 y^2+y^4$ is a sum of fourth powers of real linear forms if and only if $0 \le \lambda \le 6$. This alternative definition was studied by Ellison [19] in the special cases $K = \mathbb{C}, \mathbb{R}, \mathbb{Q}$. Newman and Slater [39] do not restrict to homogeneous polynomials. They write x as a sum of d d-th powers of linear polynomials; by substitution, any polynomial is a sum of at most d d-th powers of polynomials. They also show that the minimum number of d-th powers in this formulation is $\ge \sqrt{d}$. Because of the degrees of the summands, these methods do not homogenize to forms. Mordell [37] showed that a polynomial that is a sum of cubes of linear forms over $\mathbb{Z}$ is also a sum of at most eight such cubes. More generally, if R is a commutative ring, then its *d-Pythagoras number*, $P_d(R)$, is the smallest integer k so that any sum of d-th powers in R is a sum of k d-th powers. Helmke [25] uses both definitions for length for forms, but is mainly concerned with the alternative definition in the case when K is an algebraically closed (or a real closed) field of characteristic zero, not necessarily a subset of $\mathbb{C}$. This subject is closely related to Hilbert's 17th Problem; see [10–12]. In [47], a principal object of study is $Q_{n,2k}$, the (closed convex) cone of real forms which are a sum of $2k$-th powers of real linear forms.

We now outline the remainder of the paper.

In Sect. 2, we give a self-contained proof of Sylvester's 1851 Theorem (Theorem 2.1). Although originally given over $\mathbb{C}$, it adapts easily to any $K \subset \mathbb{C}$ (Corollary 2.2). If f is a binary form, then $L_K(f) \le r$ if and only if a certain subspace of the binary forms of degree r (a subspace determined by f) contains a form that splits into distinct factors over K. We illustrate the algorithm by proving the assertions of lengths 3 and 4 for ϕ in Example 1.2.

In Sect. 3, we prove (Theorem 3.2) a homogenized version of Sylvester's 1864 Theorem (Theorem 3.1), which implies that if real f has r linear factors over $\mathbb{R}$ (counting multiplicity), then $L_{\mathbb{R}}(f) \ge r$. As far as we can tell, Sylvester did not connect his two theorems: perhaps because he presented the second one for non-homogeneous polynomials in one variable.

We apply these theorems and some other simple observations in Sects. 4 and 5. We first show that if $L_{\mathbb{C}}(f) = 1$, then $L_{E_f}(f) = 1$ as well (Theorem 4.1). Any set of $d+1$ d-th powers of pairwise distinct linear forms is linearly independent (Theorem 4.2). It follows quickly that if $f(x, y)$ has two different honest representations of length r and s, then $r + s \ge d + 2$ (Corollary 4.3), and so if $L_{E_f}(f) = r \le \frac{d+1}{2}$, then the representation over E_f is the unique minimal $\mathbb{C}$-representation (Corollary 4.4). We show that Example 1.1 gives a template for forms f satisfying $L_{\mathbb{C}}(f) = 2 < L_{E_f}(f)$ (see Theorem 4.6), and give two generalizations which provide other types of constructions (Corollaries 4.7 and 4.8) of forms with multiple lengths. We apply Sylvester's 1851 Theorem to give an easy proof of the known result that $L_{\mathbb{C}}(f) \le d$ (Theorem 4.9) and a slightly trickier proof of the probably-known result that $L_K(f) \le d$ as well (Theorem 4.10). Theorem 4.10 combines with Theorem 3.2 into Corollary 4.11: if $f \in \mathbb{R}[x, y]$ is a product of d linear factors and not a d-th power, then $L_{\mathbb{R}}(f) = d$. Conjecture 4.12 asserts that $f \in \mathbb{R}[x, y]$ is a product of d linear factors if and only if $L_{\mathbb{R}}(f) = d$. This conjecture has recently been proven by Comon-Ottaviani-Causa-Re [9, 15] when the factors of f are distinct.

In Corollary 5.1, we discuss the various possible cabinets when $d = 3, 4$; and give examples for each one not already ruled out. We then completely classify binary cubics; the key point of Theorem 5.2 is that if the cubic f has no repeated factors, then $L_k(f) = 2$ if and only if $E_f(\sqrt{-3\Delta(f)}) \subseteq K$; this significance of the discriminant $\Delta(f)$ can already be found in Salmon [52, Sect. 167]. This proves Conjecture 4.12 for $d = 3$. In Theorem 5.3, we show that Conjecture 4.12 also holds for $d = 4$. We then show (Theorem 5.4) that $L_{\mathbb{C}}(f) = d$ if and only if there are distinct linear forms ℓ, ℓ' so that $f = \ell^{d-1}\ell'$. (One direction of this result is well-known; the other has recently been proved by Białynicki-Birula and Schinzel [2].) The minimal representations of $x^k y^k$ are parameterized (Theorem 5.5), and in Corollary 5.6, we show that $L_K((x^2 + y^2)^k) \ge k + 1$, with equality if and only if $\tan \frac{\pi}{k+1} \in K$. In particular, $L_{\mathbb{Q}}((x^2 + y^2)^2) = 4$. Theorem 5.7 shows that $L_{\mathbb{Q}}(x^4 + 6\lambda x^2 y^2 + y^4) = 3$ if and only if a certain quartic diophantine equation over $\mathbb{Z}$ has a non-zero solution.

Section 6 lists some open questions.

We would like to express our appreciation to the organizers of the *Higher Degree Forms* conference in Gainesville in May 2009 for offering the opportunities to speak on these topics, and to write this article for its Proceedings. We also thank Mike Bennett, Tony Geramita, Giorgio Ottaviani, Joe Rotman and Zach Teitler for helpful conversations and correspondence.

2 Sylvester's 1851 Theorem

Modern proofs of Theorem 2.1 can be found in the work of Kung and Rota: [33, Sect. 5], with further discussion in [30–32, 49]. We present here a very elementary proof showing the connection with constant coefficient linear recurrences, in the hopes that this remarkable theorem might become better known to the modern reader.

Theorem 2.1 (Sylvester). *Suppose*

$$f(x, y) = \sum_{j=0}^{d} \binom{d}{j} a_j x^{d-j} y^j \tag{2.1}$$

and suppose

$$h(x,y) = \sum_{t=0}^{r} c_t x^{r-t} y^t = \prod_{j=1}^{r} (-\beta_j x + \alpha_j y) \tag{2.2}$$

is a product of pairwise distinct linear factors. Then there exist $\lambda_k \in \mathbb{C}$ *so that*

$$f(x,y) = \sum_{k=1}^{r} \lambda_k (\alpha_k x + \beta_k y)^d \tag{2.3}$$

if and only if

$$\begin{pmatrix} a_0 & a_1 & \cdots & a_r \\ a_1 & a_2 & \cdots & a_{r+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{d-r} & a_{d-r+1} & \cdots & a_d \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_r \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}; \tag{2.4}$$

that is, if and only if

$$\sum_{t=0}^{r} a_{\ell+t} c_t = 0, \qquad \ell = 0, 1, \ldots, d-r. \tag{2.5}$$

Proof. First suppose that (2.3) holds. Then for $0 \le j \le d$,

$$a_j = \sum_{k=1}^{r} \lambda_k \alpha_k^{d-j} \beta_k^j \implies \sum_{t=0}^{r} a_{\ell+t} c_t = \sum_{k=1}^{r} \sum_{t=0}^{r} \lambda_k \alpha_k^{d-\ell-t} \beta_k^{\ell+t} c_t$$

$$= \sum_{k=1}^{r} \lambda_k \alpha_k^{d-\ell-r} \beta_k^{\ell} \sum_{t=0}^{r} \alpha_k^{r-t} \beta_k^t c_t = \sum_{k=1}^{r} \lambda_k \alpha_k^{d-\ell-r} \beta_k^{\ell}\, h(\alpha_k, \beta_k) = 0.$$

Now suppose that (2.4) holds and suppose first that $c_r \neq 0$. We may assume without loss of generality that $c_r = 1$ and that $\alpha_j = 1$ in (2.2), so that the β_j's are distinct. Define the *infinite* sequence $(\tilde{a}_j)$, $j \ge 0$, by:

$$\tilde{a}_j = a_j \quad \text{if} \quad 0 \le j \le r-1; \quad \tilde{a}_{r+\ell} = -\sum_{t=0}^{r-1} \tilde{a}_{t+\ell} c_t \quad \text{for} \quad \ell \ge 0. \tag{2.6}$$

This sequence satisfies the recurrence of (2.5), so that

$$\tilde{a}_j = a_j \quad \text{for} \quad j \le d. \tag{2.7}$$

Since $|\tilde{a}_j| \le \gamma \cdot M^j$ for suitable γ, M by induction, the generating function

$$\Phi(T) = \sum_{j=0}^{\infty} \tilde{a}_j T^j$$

converges in a neighborhood of 0. We have

$$\left(\sum_{t=0}^{r} c_{r-t}T^t\right)\Phi(T) = \sum_{n=0}^{r-1}\left(\sum_{j=0}^{n} c_{r-(n-j)}\tilde{a}_j\right)T^n + \sum_{n=r}^{\infty}\left(\sum_{t=0}^{r} c_{r-t}\tilde{a}_{n-t}\right)T^n.$$

It follows from (2.6) that the second sum vanishes, and hence $\Phi(T)$ is a rational function with denominator

$$\sum_{t=0}^{r} c_{r-t}T^t = h(T,1) = \prod_{j=1}^{r}(1-\beta_j T).$$

By the method of partial fractions, there exist $\lambda_k \in \mathbb{C}$ so that

$$\sum_{j=0}^{\infty}\tilde{a}_jT^j = \Phi(T) = \sum_{k=1}^{r}\frac{\lambda_k}{1-\beta_kT} \implies \tilde{a}_j = \sum_{k=1}^{r}\lambda_k\beta_k^j. \tag{2.8}$$

A comparison of (2.8) and (2.7) with (2.1) shows that

$$f(x,y) = \sum_{j=0}^{d}\binom{d}{j}a_jx^{d-j}y^j = \sum_{k=1}^{r}\lambda_k\sum_{j=0}^{d}\binom{d}{j}\beta_k^jx^{d-j}y^j = \sum_{k=1}^{r}\lambda_k(x+\beta_ky)^d, \tag{2.9}$$

as claimed in (2.3).

If $c_r = 0$, then $c_{r-1} \neq 0$, because h has distinct factors. We may proceed as before, replacing r by $r-1$ and taking $c_{r-1} = 1$, so that (2.2) becomes

$$h(x,y) = \sum_{t=0}^{r-1} c_tx^{r-t}y^t = x\prod_{j=1}^{r-1}(y-\beta_jx). \tag{2.10}$$

Since $c_r = 0$, the system (2.4) can be rewritten as

$$\begin{pmatrix} a_0 & a_1 & \cdots & a_{r-1} \\ a_1 & a_2 & \cdots & a_r \\ \vdots & \vdots & \ddots & \vdots \\ a_{d-r} & a_{d-r+1} & \cdots & a_{d-1} \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{r-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

We may now argue as before, except that (2.7) becomes

$$\tilde{a}_j = a_j \quad \text{for} \quad j \le d-1, \quad a_d = \tilde{a}_d + \lambda_r \tag{2.11}$$

for some λ_r, and (2.9) becomes

$$\begin{aligned} f(x,y) = \sum_{j=0}^{d} \binom{d}{j} a_j x^{d-j} y^j &= \lambda_r y^d + \sum_{k=1}^{r-1} \lambda_k \sum_{j=0}^{d} \binom{d}{j} \beta_k^j x^{d-j} y^j \\ &= \lambda_r y^d + \sum_{k=1}^{r-1} \lambda_k (x + \beta_k y)^d. \end{aligned} \tag{2.12}$$

By (2.10), (2.12) meets the description of (2.3), completing the proof. □

The $(d-r+1) \times (r+1)$ Hankel matrix in (2.4) will be denoted $H_r(f)$. If (f,h) satisfy the criterion of this theorem, we shall say that h is a *Sylvester form for* f. If the only Sylvester forms of degree r are λh for $\lambda \in \mathbb{C}$, we say that h is the *unique* Sylvester form for f. Any polynomial multiple of a Sylvester form that has no repeated factors is also a Sylvester form, since there is no requirement that $\lambda_k \neq 0$ in (2.3). If f has a unique Sylvester form of degree r, then $L_{\mathbb{C}}(f) = r$ and $L_K(f) \geq r$.

The proof of Theorem 2.1 in [49] is based on apolarity. If f and h are given by (2.1) and (2.2), and $h(D) = \prod_{j=1}^{r}(-\beta_j \frac{\partial}{\partial x} + \alpha_j \frac{\partial}{\partial y})$, then

$$h(D)f = \sum_{m=0}^{d-r} \frac{d!}{(d-r-m)!m!} \left(\sum_{i=0}^{d-r} a_{i+m} c_i \right) x^{d-r-m} y^m.$$

Thus, (2.4) is equivalent to $h(D)f = 0$. One can then argue that each linear factor in $h(D)$ kills a different summand, and dimension counting takes care of the rest. In particular, if $\deg h > d$, then $h(D)f = 0$ automatically, and this implies that $L_{\mathbb{C}}(f) \leq d+1$. Theorem 4.2 gives another explanation of this fact.

If h has repeated factors, a condition of interest in [30–33, 49], then Gundelfinger's Theorem [23], first proved in 1886, shows that a factor $(-\beta x + \alpha y)^\ell$ of h corresponds to a summand $q(x,y)(\alpha x + \beta y)^{d+1-\ell}$ in f, where q is an arbitrary form of degree $\ell - 1$. (We are not interested in such summands when $\ell > 1$. For more discussion of this case, see [49].)

If $d = 2s-1$ and $r = s$, then $H_s(f)$ is $s \times (s+1)$ and has a non-trivial null-vector; for a general f, the resulting form h has distinct factors, and so is a unique Sylvester form. (The coefficients of h, and its discriminant, are polynomials in the coefficients of f.) This is how Sylvester proved that a general binary form of degree $2s-1$ is a sum of s powers of linear forms over $\mathbb{C}$, and the minimal representation is unique.

If $d = 2s$ and $r = s$, then $H_s(f)$ is square; $\det(H_s(f))$ is the *catalecticant* of f. (For more on the term "catalecticant", see [47, pp. 49–50] and [22, pp. 104–105].) In general, there exists λ so that the catalecticant of $f(x,y) - \lambda x^{2s}$ vanishes, and the resulting non-trivial null vector is generally a Sylvester form (no repeated factors). Thus, a general binary form of degree $2s$ is a sum of λx^{2s} plus s powers of linear forms over $\mathbb{C}$.

Sylvester's Theorem can also be adapted to compute K-length when $K \subsetneq \mathbb{C}$, with the understanding that a Sylvester form of minimal degree might not split over K.

Corollary 2.2. *Given $f \in K[x,y]$, $L_K(f)$ is the minimal degree of a Sylvester form for f which splits completely over K.*

Proof. If (2.3) is a minimal representation for f over K, where $\lambda_k, \alpha_k, \beta_k \in K$, then $h(x,y) \in K[x,y]$ splits over K by (2.2). Conversely, if h is a Sylvester form for f satisfying (2.2) with $\alpha_k, \beta_k \in K$, then (2.3) holds for some $\lambda_k \in \mathbb{C}$. This is equivalent to saying that the linear system

$$a_j = \sum_{k=1}^{r} \alpha_k^{d-j} \beta_k^j X_k, \quad (0 \le j \le d) \tag{2.13}$$

has a solution $\{X_k = \lambda_k\}$ over $\mathbb{C}$. Since $a_j, \alpha_k^{d-j}\beta_k^j \in K$ and (2.13) has a solution over $\mathbb{C}$, it also has a solution over K. Thus, f has a K-representation of length r.

□

Example 2.1 (Continuing Example 1.2). Note that

$$\phi(x,y) = 3x^5 - 20x^3y^2 + 10xy^4 = \binom{5}{0} \cdot 3\, x^5 + \binom{5}{1} \cdot 0\, x^4y$$
$$+ \binom{5}{2} \cdot (-2)\, x^3y^2 + \binom{5}{3} \cdot 0\, x^2y^3 + \binom{5}{4} \cdot 2\, xy^4 + \binom{5}{5} \cdot 0\, y^5.$$

Since

$$\begin{pmatrix} 3 & 0 & -2 & 0 \\ 0 & -2 & 0 & 2 \\ -2 & 0 & 2 & 0 \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff (c_0, c_1, c_2, c_3) = r(0,1,0,1),$$

ϕ has a *unique* Sylvester form of degree 3: $h(x,y) = y(x^2+y^2) = y(y-ix)(y+ix)$. Accordingly, there exist $\lambda_k \in \mathbb{C}$ so that

$$\phi(x,y) = \lambda_1 x^5 + \lambda_2(x+iy)^5 + \lambda_3(x-iy)^5.$$

Indeed, $\lambda_1 = \lambda_2 = \lambda_3 = 1$, as may be checked. It follows that $L_K(\phi) = 3$ if and only if $i \in K$. (There is no representation of length two.)

To find representations for ϕ of length 4, we consider (2.4) for ϕ with $r = 4$:

$$H_4(\phi) \cdot (c_0, c_1, c_2, c_3, c_4)^t = (0,0)^t \iff 3c_0 - 2c_2 + 2c_4 = -2c_1 + 2c_3 = 0$$
$$\iff (c_0, c_1, c_2, c_3, c_4) = r_1(2,0,3,0,0) + r_2(0,1,0,1,0) + r_3(0,0,1,0,1),$$

hence $h(x,y) = r_1x^2(2x^2+3y^2) + y(x^2+y^2)(r_2x+r_3y)$. Given a field K, it is unclear whether there exist $\{r_\ell\}$ so that h splits into distinct factors over K. We have found such $\{r_\ell\}$ for small imaginary quadratic fields.

The choice $(r_1,r_2,r_3) = (1,0,2)$ gives $h(x,y) = (2x^2+y^2)(x^2+2y^2)$ and

$$24\phi(x,y) = 4(x+\sqrt{-2}y)^5 + 4(x-\sqrt{-2}y)^5 + (2x+\sqrt{-2}y)^5 + (2x-\sqrt{-2}y)^5.$$

Similarly, $(r_1,r_2,r_3) = (2,0,9)$ and $(2,0,-5)$ give $h(x,y) = (x^2+3y^2)(4x^2+3y^2)$ and $(x^2-y^2)(4x^2+5y^2)$, leading to representations for ϕ of length 4 over $\mathbb{Q}(\sqrt{-3})$ and $\mathbb{Q}(\sqrt{-5})$. The simplest such representation we have found for $\mathbb{Q}(\sqrt{-6})$ uses $(r_1,r_2,r_3) = (12{,}675, 0, -156{,}816)$ and

$$h(x,y) = (5x+12y)(5x-12y)(6\cdot 13^2x^2+33^2y^2).$$

We conjecture that $L_{\mathbb{Q}(\sqrt{-m})}(\phi) = 4$ for all non-square $m \geq 2$. In Example 3.1, we shall show that there is no choice of (r_1,r_2,r_3) for which h splits into distinct factors over any subfield of $\mathbb{R}$.

3 Sylvester's 1864 Theorem

Theorem 3.1 was discovered by Sylvester [55] in 1864 while proving Isaac Newton's conjectural variation on Descartes' Rule of Signs, see [27, 56]. This theorem appeared in Pólya-Szegö [42, Chap. 5, Problem 79], and has been used by Pólya and Schoenberg [41] and Karlin [29, p. 466]. The (dehomogenized) version proved in [42] is:

Theorem 3.1 (Sylvester). *Suppose* $0 \neq \lambda_k$ *for all* k *and* $\gamma_1 < \cdots < \gamma_r$, $r \geq 2$, *are real numbers such that*

$$Q(t) = \sum_{k=1}^{r} \lambda_k (t-\gamma_k)^d$$

does not vanish identically. Suppose the sequence $(\lambda_1,\dots,\lambda_r,(-1)^d\lambda_1)$ *has* C *changes of sign and* Q *has* Z *zeros, counting multiplicity. Then* $Z \leq C$.

We shall prove an equivalent version which exploits the homogeneity of f to avoid discussion of zeros at infinity in the proof. (The equivalence is discussed in [50].)

Theorem 3.2. *Suppose* $f(x,y)$ *is a non-zero real form of degree* d *with* τ *real linear factors (counting multiplicity) and*

$$f(x,y) = \sum_{j=1}^{r} \lambda_j (\cos\theta_j x + \sin\theta_j y)^d \tag{3.1}$$

where $-\frac{\pi}{2} < \theta_1 < \cdots < \theta_r \leq \frac{\pi}{2}$, $r \geq 2$ *and* $\lambda_j \neq 0$. *If there are* σ *sign changes in the tuple* $(\lambda_1,\lambda_2,\dots,\lambda_r,(-1)^d\lambda_1)$, *then* $\tau \leq \sigma$. *In particular,* $\tau \leq r$.

Example 3.1 (Examples 1.2 and 2.1 concluded). Since

$$\phi(x,y) = 3x\big(x^2 - \tfrac{10-\sqrt{70}}{3}y^2\big)\big(x^2 - \tfrac{10+\sqrt{70}}{3}y^2\big)$$

is a product of five linear factors over $\mathbb{R}$, $L_{\mathbb{R}}(\phi) \ge 5$. The representation

$$6\phi(x,y) = 36x^5 - 10(x+y)^5 - 10(x-y)^5 + (x+2y)^5 + (x-2y)^5$$

over $\mathbb{Q}$ implies that $L_{\mathbb{R}}(\phi) = 5$. It will follow from Theorem 4.10 that $\mathcal{C}(\phi) = \{3,4,5\}$.

Proof of Theorem 3.2. We first rewrite (3.1):

$$2f(x,y) = \sum_{j=1}^{r} \lambda_j (\cos\theta_j x + \sin\theta_j y)^d + \sum_{j=1}^{r} (-1)^d \lambda_j (\cos(\theta_j + \pi)x + \sin(\theta_j + \pi)y)^d. \tag{3.2}$$

View the sequence $(\lambda_1, \lambda_2, \dots, \lambda_r, (-1)^d\lambda_1, (-1)^d\lambda_2, \dots, (-1)^d\lambda_r, \lambda_1)$ cyclically, identifying the first and last term. There are 2σ pairs of consecutive terms with a negative product. This count is independent of the starting point, so if we make any invertible change of variables $(x,y) \mapsto (\cos\theta x + \sin\theta y, -\sin\theta x + \cos\theta y)$ in (3.1) (which doesn't affect τ, and which "dials" the angles by θ), and reorder the "main" angles to $(-\frac{\pi}{2}, \frac{\pi}{2}]$, the value of σ is unchanged. We may therefore assume that neither x nor y divide f, that x^d and y^d are not summands in (3.2) (i.e., θ_j is not a multiple of $\frac{\pi}{2}$), and that if there is a sign change in $(\lambda_1, \lambda_2, \dots, \lambda_r)$, then $\theta_u < 0 < \theta_{u+1}$ implies $\lambda_u \lambda_{u+1} < 0$. Under these hypotheses, we may safely dehomogenize f by setting either $x = 1$ or $y = 1$ and avoid zeros at infinity and know that τ is the number of zeros of the resulting polynomial. The rest of the proof generally follows [42].

Let $\bar{\sigma}$ denote the number of sign changes in $(\lambda_1, \lambda_2, \dots, \lambda_r)$. We induct on $\bar{\sigma}$. The base case is $\bar{\sigma} = 0$ (and $\lambda_j > 0$ without loss of generality). If d is even, then $\sigma = 0$ and

$$f(x,y) = \sum_{j=1}^{r} \lambda_j (\cos\theta_j x + \sin\theta_j y)^d$$

is definite, so $\tau = 0$. If d is odd, then $\sigma = 1$. Let $g(t) = f(t,1)$, so that

$$g'(t) = \sum_{j=1}^{r} d\,(\lambda_j \cos\theta_j)\,(\cos\theta_j t + \sin\theta_j)^{d-1}.$$

Since $d-1$ is even, $\cos\theta_j > 0$ and $\lambda_j > 0$, g' is definite and $g' \neq 0$. Rolle's Theorem implies that g has at most one zero; that is, $\tau \le 1 = \sigma$.

Suppose the theorem is valid for $\bar{\sigma} = m \geq 0$ and suppose that $\bar{\sigma} = m + 1$ in (3.1). Now let $h(t) = f(1,t)$. We have

$$h'(t) = \sum_{j=1}^{r} d\left(\lambda_j \sin\theta_j\right)\left(\cos\theta_j + \sin\theta_j t\right)^{d-1}.$$

Note that $h'(t) = q(1,t)$, where

$$q(x,y) = \sum_{j=1}^{r} d\left(\lambda_j \sin\theta_j\right)\left(\cos\theta_j x + \sin\theta_j y\right)^{d-1}.$$

Since $\bar{\sigma} \geq 1$, $\theta_u < 0 < \theta_{u+1}$ implies that $\lambda_u\lambda_{u+1} < 0$, so that the number of sign changes in $(d\lambda_1 \sin\theta_1, d\lambda_2 \sin\theta_2, \dots, d\lambda_r \sin\theta_r)$ is m, as the sign change at the u-th consecutive pair has been removed, and no other possible sign changes are introduced. The induction hypothesis implies that $q(x,y)$ has at most m linear factors, hence $q(1,t) = h'(t)$ has $\leq m$ zeros (counting multiplicity) and Rolle's Theorem implies that h has $\leq m+1$ zeros, completing the induction. □

4 Applications to Forms of General Degree

We begin with a folklore result: the vector space of complex forms f in n variables of degree d is spanned by the set of linear forms taken to the d-th power. It follows from a 1903 theorem of Biermann (see [47, Proposition 2.11] or [51] for a proof) that a canonical set of the "correct" number of d-th powers over $\mathbb{Z}$ forms a basis:

$$\left\{(i_1x_1 + \ldots + i_nx_n)^d \ : \ 0 \leq i_k \in \mathbb{Z},\ i_1 + \cdots + i_n = d\right\}. \qquad (4.1)$$

If $f \in K[x_1,\dots,x_n]$, then f is a K-linear combination of these forms and so $L_K(f) \leq \binom{n+d-1}{n-1}$. We show below (Theorems 4.10 and 5.4) that when $n = 2$, the bound for $L_K(f)$ can be improved from $d+1$ to d, but this is best possible.

The first two results are presented explicitly for completeness.

Theorem 4.1. *If $f \in K[x,y]$, then $L_K(f) = 1$ if and only if $L_{\mathbb{C}}(f) = 1$.*

Proof. One direction is immediate from (1.2). For the other, suppose $f(x,y) = (\alpha x + \beta y)^d$ with $\alpha, \beta \in \mathbb{C}$. If $\alpha = 0$, then $f(x,y) = \beta^d y^d$, with $\beta^d \in K$. If $\alpha \neq 0$, then $f(x,y) = \alpha^d(x + (\beta/\alpha)y)^d$. Since the coefficients of x^d and $dx^{d-1}y$ in f are α^d and $\alpha^{d-1}\beta$, it follows that α^d and $\beta/\alpha = (\alpha^{d-1}\beta)/\alpha^d$ are both in K. □

Theorem 4.2. *Any set $\{(\alpha_j x + \beta_j y)^d : 0 \leq j \leq d\}$ of pairwise distinct d-th powers is linearly independent and spans the binary forms of degree d.*

Proof. The matrix of this set with respect to the basis $\binom{d}{i}x^{d-i}y^i$ is $[\alpha_j^{d-i}\beta_j^i]$, whose determinant is Vandermonde:

$$\prod_{0\le j<k\le d}\begin{vmatrix}\alpha_j & \beta_j\\ \alpha_k & \beta_k\end{vmatrix}.$$

This determinant is a product of non-zero terms by hypothesis. □

By considering the difference of two representations of a given form, we obtain an immediate corollary about different representations of the same form. Trivial counterexamples, formed by splitting summands, occur in non-honest representations.

Corollary 4.3. *If f has two different honest representations:*

$$f(x,y)=\sum_{i=1}^{s}\lambda_i(\alpha_i x+\beta_i y)^d=\sum_{j=1}^{t}\mu_j(\gamma_j x+\delta_j y)^d, \tag{4.2}$$

then $s+t\ge d+2$. If $s+t=d+2$ in (4.2), then the combined set of linear forms, $\{\alpha_i x+\beta_i y, \gamma_j x+\delta_j y\}$, is pairwise distinct.

The next result collects some consequences of Corollary 4.3.

Corollary 4.4. *Let $E=E_f$.*

(1) *If $L_E(f)=r\le\frac{d}{2}+1$, then $L_{\mathbb{C}}(f)=r$, so $\mathcal{C}(f)=\{r\}$.*
(2) *If, further, $L_E(f)=r\le\frac{d}{2}+\frac{1}{2}$, then f has a unique $\mathbb{C}$-minimal representation.*
(3) *If $d=2s-1$ and $H_s(f)$ has full rank, f has a unique Sylvester form h of degree s and $E_f\subseteq K$, then $L_K(f)\ge s$, with equality if and only if h splits in K.*

Proof. We take the parts in turn.

(1) A different representation of f over $\mathbb{C}$ must have length $\ge d+2-r\ge\frac{d}{2}+1\ge r$ by Corollary 4.3, and so $L_{\mathbb{C}}(f)=r$.
(2) If $r\le\frac{d}{2}+\frac{1}{2}$, then any other representation has length $\ge\frac{d}{2}+\frac{3}{2}>r$, and so cannot be minimal.
(3) If $d=2s-1$ and $r=s$, then the last case applies, so f has a unique $\mathbb{C}$-minimal representation, and by Corollary 2.2, this representation can be expressed in K if and only if the Sylvester form splits over K. □

We now give some more explicit constructions of forms with multiple lengths. We first need a lemma about cubics.

Lemma 4.5. *If f is a cubic given by (2.1) and $H_2(f)=\begin{pmatrix}a_0 & a_1 & a_2\\ a_1 & a_2 & a_3\end{pmatrix}$ has rank ≤ 1, then f is a cube.*

Proof. If $a_0 = 0$, then $a_1 = 0$, so $a_2 = 0$ and f is a cube. If $a_0 \neq 0$, then $a_2 = a_1^2/a_0$ and $a_3 = a_1 a_2/a_0 = a_1^3/a_0^2$ and $f(x,y) = a_0(x + \frac{a_1}{a_0}y)^3$ is again a cube. □

Theorem 4.6. *Suppose $d \geq 3$ and there exist $\alpha_i, \beta_i \in \mathbb{C}$ so that*

$$f(x,y) = \sum_{i=0}^{d} \binom{d}{i} a_i x^{d-i} y^i = (\alpha_1 x + \beta_1 y)^d + (\alpha_2 x + \beta_2 y)^d \in K[x,y]. \quad (4.3)$$

If (4.3) *is honest and $L_K(f) > 2$, then there exists $u \in K$ with $\sqrt{u} \notin K$ so that $L_{K(\sqrt{u})}(f) = 2$. The summands in* (4.3) *are conjugates of each other in $K(\sqrt{u})$.*

Proof. First observe that if $\alpha_2 = 0$, then $\alpha_2\beta_1 \neq \alpha_1\beta_2$ implies that $\alpha_1 \neq 0$. But then $a_0 = \alpha_1^d \neq 0$ and $a_1 = \alpha_1^{d-1}\beta_1$ imply that $\alpha_1^d, \beta_1/\alpha_1 \in K$ as in Theorem 4.1, and so

$$f(x,y) - \alpha_1^d(x + (\beta_1/\alpha_1)y)^d = (\beta_2 y)^d = \beta_2^d y^d \in K[x,y].$$

This contradicts $L_K(f) > 2$, so $\alpha_2 \neq 0$; similarly, $\alpha_1 \neq 0$. Let $\lambda_i = \alpha_i^d$ and $\gamma_i = \beta_i/\alpha_i$ for $i = 1, 2$, so $\lambda_1\lambda_2 \neq 0$ and $\gamma_1 \neq \gamma_2$. We have

$$f(x,y) = \lambda_1(x + \gamma_1 y)^d + \lambda_2(x + \gamma_2 y)^d \implies a_i = \lambda_1\gamma_1^i + \lambda_2\gamma_2^i.$$

Now let

$$g(x,y) = \lambda_1(x + \gamma_1 y)^3 + \lambda_2(x + \gamma_2 y)^3 = a_0x^3 + 3a_1x^2y + 3a_2xy^2 + a_3y^3.$$

Since $\lambda_i \neq 0$ and (4.3) is honest, Corollary 4.3 implies that $L_{\mathbb{C}}(g) = 2$, so $H_2(g)$ has full rank by Lemma 4.5. It can be checked directly that

$$\begin{pmatrix} a_0 & a_1 & a_2 \\ a_1 & a_2 & a_3 \end{pmatrix} \cdot \begin{pmatrix} \gamma_1\gamma_2 \\ -(\gamma_1 + \gamma_2) \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

and this gives $h(x,y) = (y - \gamma_1 x)(y - \gamma_2 x)$ as the unique Sylvester form for g. Since $H_2(g)$ has entries in K and hence has a null vector in K, we must have $h \in K[x,y]$. By hypothesis, h does not split over K; it must do so over $K(\sqrt{u})$, where $u = (\gamma_1 - \gamma_2)^2 = (\gamma_1 + \gamma_2)^2 - 4\gamma_1\gamma_2 \in K$. Moreover, if σ denotes conjugation with respect to $\sqrt{u}$, then $\gamma_2 = \sigma(\gamma_1)$ and since $\lambda_1 + \lambda_2 \in K$, $\lambda_2 = \sigma(\lambda_1)$ as well. Note that $\lambda_i = \alpha_i^d$ and $\gamma_i = \beta_i/\alpha_i \in K(\sqrt{u})$, but this is not necessarily true for α_i and β_i themselves. □

Corollary 4.7. *Suppose $g \in E[x,y]$ does not split over E, but factors into distinct linear factors $g(x,y) = \prod_{j=1}^{r}(x + \alpha_j y)$ over an extension field K of E. If $d > 2r - 1$, then for each $\ell \geq 0$,*

$$f_\ell(x,y) = \sum_{j=1}^{r} \alpha_j^\ell (x+\alpha_j y)^d \in E[x,y],$$

and $L_K(f_\ell) = r < d+2-r \le L_E(f_\ell)$.

Proof. The coefficient of $\binom{d}{k}x^{d-k}y^k$ in f_ℓ is $\sum_{j=1}^{r}\alpha_j^{\ell+k}$. Each such power-sum belongs to E by Newton's Theorem on Symmetric Forms. If $\alpha_s \notin E$ (which must hold for at least one $\alpha_s \ne 0$), then $\alpha_s^\ell(x+\alpha_s y)^d \notin E[x,y]$. Apply Corollary 4.3. □

Corollary 4.8. *Suppose K is an extension field of E_f, $r \le \frac{d+1}{2}$, and*

$$f(x,y) = \sum_{i=1}^{r} \lambda_i(\alpha_i x + \beta_i y)^d$$

with $\lambda_i, \alpha_i, \beta_i \in K$. Then every automorphism of K which fixes E_f permutes the summands of the representation of f.

Proof. We interpret $\sigma(\lambda(\alpha x+\beta y)^d) = \sigma(\lambda)(\sigma(\alpha)x+\sigma(\beta)y)^d$. Since $\sigma(f)=f$, the action of σ is to give another representation of f. Corollary 4.4(2) implies that this is the same representation, perhaps reordered. □

Theorem 4.9. *If $f \in K[x,y]$, then $L_{\mathbb{C}}(f) \le \deg f$.*

Proof. By a change of variables, which does not affect the length, we may assume that neither x nor y divide f, hence $a_0a_d \ne 0$ and $h = a_dx^d - a_0y^d$ is a Sylvester form which splits over $\mathbb{C}$. □

We have been unable to find an "original" citation for Theorem 4.9. It appears as an exercise in Harris [24, Exercise 11.35], with the (dehomogenized) maximal length occurring at $x^{d-1}(x+1)$ (see Theorem 5.4). Landsberg and Teitler [34, Corollary 5.2] prove that $L_{\mathbb{C}}(f) \le \binom{n+d-1}{n-1}-(n-1)$, which reduces to Theorem 4.9 for $n=2$. The proof of Theorem 4.9 will not apply to $L_K(f)$ for $K \ne \mathbb{C}$, because $a_dx^d - a_0y^d$ usually does not split over K. A more careful argument is required, constructing an explicit Sylvester form of degree d for f which splits over K.

Theorem 4.10. *If $f \in K[x,y]$, then $L_K(f) \le \deg f$.*

Proof. Write f as in (2.1). If f is identically zero, there is nothing to prove. Otherwise, we may assume that $f(1,0) = a_0 \ne 0$ after a change of variables if necessary. By Corollary 2.2, it suffices to find $h(x,y) = \sum_{k=0}^{d} c_kx^{d-k}y^k$ which splits into distinct linear factors over K and satisfies $\sum_{k=0}^{d} a_kc_k = 0$.

Let $e_0 = 1$ and $e_k(t_1,\dots,t_{d-1})$ denote the usual k-th elementary symmetric functions. We make a number of definitions:

$$h_0(t_1,\dots,t_{d-1};x,y) := \sum_{k=0}^{d-1} e_k(t_1,\dots,t_{d-1})x^{d-1-k}y^k = \prod_{j=1}^{d-1}(x+t_jy),$$

$$\beta(t_1,\dots,t_{d-1}) := -\sum_{k=0}^{d-1} a_ke_k(t_1,\dots,t_{d-1}),$$

$$\alpha(t_1,\dots,t_{d-1}) := \sum_{k=0}^{d-1} a_{k+1}e_k(t_1,\dots,t_{d-1}),$$

$$\Phi(t_1,\dots,t_{d-1}) := \prod_{j=1}^{d-1}(\alpha(t_1,\dots,t_{d-1})t_j - \beta(t_1,\dots,t_{d-1})),$$

$$\Psi(t_1,\dots,t_{d-1}) := \Phi(t_1,\dots,t_{d-1}) \times \prod_{1\le i<j\le d-1}(t_i-t_j).$$

Then $\beta(0,\dots,0) = -a_0e_0 = -a_0 \neq 0$, so $\Phi(0,\dots,0) = a_0^{d-1} \neq 0$ and Φ is not the zero polynomial, and thus neither is Ψ. Choose $\gamma_j \in K$, $1 \le j \le d-1$, so that $\Psi(\gamma_1,\dots,\gamma_{d-1}) \neq 0$. It follows that the γ_j's are distinct, and $\alpha\gamma_j \neq \beta$, where $\alpha = \alpha(\gamma_1,\dots,\gamma_{d-1})$ and $\beta = \beta(\gamma_1,\dots,\gamma_{d-1})$. Let $\tilde{e}_k = e_k(\gamma_1,\dots,\gamma_{d-1})$. We claim that

$$h(x,y) = \sum_{i=0}^{d} c_ix^{d-1}y^i := (\alpha x+\beta y)h_0(\gamma_1,\dots,\gamma_{d-1};x,y) = (\alpha x+\beta y)\prod_{j=1}^{d-1}(x+\gamma_jy)$$

$$= (\alpha x+\beta y)\sum_{k=0}^{d-1}\tilde{e}_kx^{d-1-k}y^k = \alpha\tilde{e}_0x^d + \sum_{k=1}^{d-1}(\alpha\tilde{e}_k+\beta\tilde{e}_{k-1})x^{d-k}y^k + \beta\tilde{e}_{d-1}y^d$$

is a Sylvester form for f. Note that the γ_j's are distinct and $\alpha\gamma_j \neq \beta$, $1 \le j \le d-1$, so that h is a product of distinct linear factors. Finally,

$$\sum_{k=0}^{d} a_kc_k = \alpha\tilde{e}_0a_0 + \sum_{k=1}^{d-1}(\alpha\tilde{e}_k+\beta\tilde{e}_{k-1})a_k + \beta\tilde{e}_{d-1}a_k =$$

$$\alpha\sum_{k=0}^{d-1}\tilde{e}_ka_k + \beta\sum_{k=0}^{d-1}\tilde{e}_ka_{k+1} = \alpha(-\beta)+\beta\alpha = 0.$$

This completes the proof. □

Corollary 4.11. *If f is a product of d real linear forms and not a d-th power, then $L_{\mathbb{R}}(f) = d$.*

Proof. Write f as a sum of $L_{\mathbb{R}}(f) = r \le d$ d-th powers and rescale into the shape (3.1). Taking $\tau = d$ in Theorem 3.2, we see that $d \le \sigma \le r$. □

Conjecture 4.12. *If $f \in \mathbb{R}[x,y]$ is a form of degree $d \ge 3$, then $L_{\mathbb{R}}(f) = d$ if and only if f is a product of d linear forms.*

We shall see in Theorems 5.2 and 5.3 that this conjecture is true for $d = 3, 4$.

After a preprint of this paper was distributed, Giorgio Ottaviani pointed out that in the case that the roots of f are distinct, Conjecture 4.12 has been proved very recently by Comon and Ottaviani [15] and by Causa and Re [9].

5 Applications to Forms of Particular Degree

Corollary 4.3 and Theorem 4.10 impose some immediate restrictions on the possible cabinets of a form of degree d.

Corollary 5.1. *Suppose* $\deg f = d$.

(1) *If $L_{\mathbb{C}}(f) = r$, then $\mathcal{C}(f) \subseteq \{r\} \cup \{d - i : 0 \le i \le r - 2\}$.*
(2) *If $L_{\mathbb{C}}(f) = 2$, then $\mathcal{C}(f)$ is either $\{2\}$ or $\{2, d\}$.*
(3) *If f has k different lengths, then $d \ge 2k - 1$.*
(4) *If f is cubic, then $\mathcal{C}(f) = \{1\}, \{2\}, \{3\}$ or $\{2, 3\}$.*
(5) *If f is quartic, then $\mathcal{C}(f) = \{1\}, \{2\}, \{3\}, \{4\}, \{2, 4\}$ or $\{3, 4\}$.*

We now completely classify $L_K(f)$ when f is a binary cubic.

Theorem 5.2. *Suppose $f(x,y) \in E_f[x,y]$ is a cubic form with discriminant Δ and suppose $E_f \subseteq K \subseteq \mathbb{C}$.*

(1) *If f is a cube, then $L_K(f) = 1$ and $\mathcal{C}(f) = \{1\}$.*
(2) *If f has a repeated linear factor, but is not a cube, then $L_K(f) = 3$ and $\mathcal{C}(f) = \{3\}$.*
(3) *If f does not have a repeated factor, then $L_K(p) = 2$ if $\sqrt{-3\Delta} \in K$ and $L_K(p) = 3$ otherwise, so either $\mathcal{C}(f) = \{2\}$ or $\mathcal{C}(f) = \{2, 3\}$.*

Proof. The first case follows from Theorem 4.1. In the second case, after an invertible linear change of variables, we may assume that $f(x,y) = 3x^2y$, and apply Theorem 2.1 to test for representations of length 2. But

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies c_0 = c_1 = 0, \tag{5.1}$$

so h has repeated factors. Hence $L_K(x^2y) \ge 3$ and by Theorem 4.10, $L_K(x^2y) = 3$.

Finally, suppose

$$f(x,y) = a_0x^3 + 3a_1x^2y + 3a_2xy^2 + a_3y^3 = \prod_{j=1}^{3}(r_jx + s_jy)$$

does not have repeated factors, so that

$$0 \neq \Delta(f) = \prod_{j<k} (r_j s_k - r_k s_j)^2,$$

and consider the system:

$$\begin{pmatrix} a_0 & a_1 & a_2 \\ a_1 & a_2 & a_3 \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

By Lemma 4.5, this system has rank 2; the unique Sylvester form is

$$h(x,y) = (a_1a_3 - a_2^2)x^2 + (a_1a_2 - a_0a_3)xy + (a_0a_2 - a_1^2)y^2,$$

which happens to be the Hessian of f. Since $h \in E_f[x,y] \subseteq K[x,y]$, it splits over K if and only if its discriminant is a square in K. A computation shows that

$$(a_1a_2 - a_0a_3)^2 - 4(a_1a_3 - a_2^2)(a_0a_2 - a_1^2) = -\frac{\Delta(f)}{27} = -\frac{3\Delta(f)}{9^2}.$$

Thus, $L_K(f) = 2$ if and only if $\sqrt{-3\Delta(f)} \in K$. If h does not split over F, then $L_F(f) = 3$ by Theorem 4.10. □

In particular, x^3, $x^3 + y^3$, x^2y and $(x+iy)^3 + (x-iy)^3$ have the cabinets enumerated in Corollary 5.1(4). If f has three distinct real linear factors, then $\Delta(f) > 0$, so $\sqrt{-3\Delta(f)} \notin \mathbb{R}$ and $L_{\mathbb{R}}(f) = 3$. If f is real and has one real and two conjugate complex linear factors, then $\Delta(f) < 0$, so $L_{\mathbb{R}}(f) = 2$. Counting repeated roots, we see that if f is a real cubic, and not a cube, then $L_{\mathbb{R}}(f) = 3$ if and only if it has three real factors, thus proving Conjecture 4.12 when $d = 3$.

Example 5.1. We find all representations of $3x^2y$ of length 3. Note that

$$H_3(f) \cdot (c_0, c_1, c_2, c_3)^t = (0) \iff c_1 = 0 \iff h(x,y) = c_0x^3 + c_2xy^2 + c_3y^3.$$

If $c_0 = 0$, then $y^2 \mid h$, which is to be avoided, so we scale and assume $c_0 = 1$. We can parameterize the Sylvester forms as $h(x,y) = (x-ay)(x-by)(x+(a+b)y)$ with $a, b, -(a+b)$ distinct. This leads to an easily checked general formula

$$\begin{gathered} 3(a-b)(a+2b)(2a+b)x^2y = \\ (a+2b)(ax+y)^3 - (2a+b)(bx+y)^3 + (a-b)(-(a+b)x+y)^3. \end{gathered} \tag{5.2}$$

Białynicki-Birula and Schinzel [2, Lemma 7.1] give the general formula for $dx^{d-1}y$ as a sum of d d-th powers of linear forms.

Theorem 5.3. *If f is a real quartic form, then $L_{\mathbb{R}}(f) = 4$ if and only if f is a product of four linear factors.*

Proof. Factor $\pm f$ as a product of k positive definite quadratic forms and $4-2k$ linear forms. If $k=0$, then Corollary 4.11 implies that $L_{\mathbb{R}}(f)=4$. We must show that if $k=1$ or $k=2$, then f has a representation over $\mathbb{R}$ as a sum of ≤ 3 fourth powers.

If $k=2$, then f is positive definite and by [43, Theorem 6], after an invertible linear change of variables, $f(x,y)=x^4+6\lambda x^2y^2+y^4$, with $6\lambda\in(-2,2]$. (This is also proved in [51].) If $r\ne 1$, then

$$\begin{aligned}&(rx+y)^4+(x+ry)^4-(r^3+r)(x+y)^4\\ &=(r-1)^2(r^2+r+1)\left(x^4-\left(\tfrac{6r}{r^2+r+1}\right)x^2y^2+y^4\right).\end{aligned}\tag{5.3}$$

Let $\phi(r)=-\frac{6r}{r^2+r+1}$. Then $\phi(-2+\sqrt{3})=2$ and $\phi(1)=-2$, and since ϕ is continuous, it maps $[-2+\sqrt{3},1)$ onto $(-2,2]$, and (5.3) shows that $L_{\mathbb{R}}(f)\le 3$.

If $k=1$, there are two cases, depending on whether the linear factors are distinct. Suppose that after a linear change, $f(x,y)=x^2h(x,y)$, where h is positive definite, and so for some $\lambda>0$ and linear ℓ, $h(x,y)=\lambda x^2+\ell^2$. After another linear change,

$$f(x,y)=x^2(2x^2+12y^2)=(x+y)^4+(x-y)^4-2y^4,\tag{5.4}$$

and (5.4) shows that $L_{\mathbb{R}}(f)\le 3$.

If the linear factors are distinct, then after a linear change,

$$f(x,y)=xy(ax^2+2bxy+cy^2),$$

where $a>0, c>0, b^2<ac$. After a scaling, $f(x,y)=xy(x^2+dxy+y^2)$, $|d|<2$, and by taking $\pm f(x,\pm y)$, we may assume $d\in[0,2)$. If $r\ne 1$, then

$$\begin{aligned}&(r^4+1)(x+y)^4-(rx+y)^4-(x+ry)^4\\ &=4(r-1)^2(r^2+r+1)\left(x^3y+\left(\tfrac{3(1+r)^2}{2(r^2+r+1)}\right)x^2y^2+xy^3\right).\end{aligned}\tag{5.5}$$

Let $\psi(r)=\frac{3(1+r)^2}{2(r^2+r+1)}$. Since $\psi(-1)=0$, $\psi(1)=2$ and ψ is continuous, it maps $[-1,1)$ onto $[0,2)$, and (5.5) shows that $L_{\mathbb{R}}(f)\le 3$. □

The next result may be very old; $L_{\mathbb{C}}(x^{d-1}y)=d$ seems well known, but the only reference we have seen for the converse is the very recent [2, Corollary 3]. Białynicki-Birula and Schinzel also classify all binary p with $\deg p=d$ and $L_{\mathbb{C}}(p)=d-k$ for $1\le k\le 3$ and sufficiently large d. Landsberg and Teitler [34, Corollary 4.5] and Boij, Carlini and Geramita [4] have both shown that $L_{\mathbb{C}}(x^ay^b)=\max(a+1,b+1)$ if $a,b\ge 1$.

Theorem 5.4. *If $d\ge 3$, then $L_{\mathbb{C}}(f)=d$ if and only if there are two distinct linear forms ℓ and ℓ' so that $f=\ell^{d-1}\ell'$.*

Proof. If $f = \ell^{d-1}\ell'$, then after an invertible linear change, we may assume that $f(x,y) = dx^{d-1}y$. If $L_{\mathbb{C}}(dx^{d-1}y) \le d-1$, then f would have a Sylvester form of degree $d-1$. But then, as in (5.1), (2.4) becomes

$$\begin{pmatrix} 0 \; 1 \cdots 0 \\ 1 \; 0 \cdots 0 \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{d-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies c_0 = c_1 = 0,$$

so $h(x,y) = \sum_{t=0}^{d-1} c_t x^{d-1-t} y^t$ does not have distinct factors. Thus, $L_{\mathbb{C}}(dx^{d-1}y) = d$.

Conversely, suppose $L_{\mathbb{C}}(f) = d$. Factor $f = \prod \ell_j^{m_j}$ as a product of pairwise distinct linear forms, with $\sum m_j = d$, $m_1 \ge m_2 \cdots \ge m_s \ge 1$, and $s > 1$ (otherwise, $L_{\mathbb{C}}(f) = 1$.) Make an invertible linear change taking $(\ell_1, \ell_2) \mapsto (x,y)$, and call the new form g; $L_{\mathbb{C}}(g) = d$ as well. If $g(x,y) = \sum_{\ell=0}^{d} \binom{d}{\ell} b_\ell x^{d-\ell} y^\ell$, then $b_0 = b_d = 0$. By hypothesis, there does not exist a Sylvester form of degree $d-1$ for g. Consider Theorem 2.1 in this case. We have

$$\begin{pmatrix} 0 & b_1 & \cdots & b_{d-2} & b_{d-1} \\ b_1 & b_2 & \cdots & b_{d-1} & 0 \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{d-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

If $m_1 \ge m_2 \ge 2$, then $x^2, y^2 \mid g(x,y)$ and $b_1 = b_{d-1} = 0$ and $x^{d-1} - y^{d-1}$ is a Sylvester form of degree $d-1$ for f. Thus $m_2 = 1$ and so y^2 does not divide g and $b_1 \ne 0$. Let $q(t) = \sum_{i=0}^{d-2} b_{i+1} t^i$ (note the absence of binomial coefficients!) and suppose q is not the constant polynomial. Then there exists t_0 so that $q(t_0) = 0$. Since $q(0) = b_1$, $t_0 \ne 0$. We have

$$\begin{pmatrix} 0 & b_1 & \cdots & b_{d-2} & b_{d-1} \\ b_1 & b_2 & \cdots & b_{d-1} & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ t_0 \\ \vdots \\ t_0^{d-1} \end{pmatrix} = \begin{pmatrix} t_0 q(t_0) \\ q(t_0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Since

$$h(x,y) = \sum_{i=0}^{d-1} t_0^i x^{d-1-i} y^i = \frac{x^d - t_0^d y^d}{x - t_0 y} = \prod_{k=1}^{d-1} (x - \zeta_{d-1}^k t_0 y)$$

has distinct linear factors, it is a Sylvester form for g, and $L_{\mathbb{C}}(g) \le d-1$. This contradiction implies that q has no zeros, so $q(t) = b_1$ must be a constant. It follows that $g(x,y) = db_1 x^{d-1} y$, as promised. □

By Corollaries 4.4 and 5.1, instances of the first five cabinets in Corollary 5.1(5) are: x^4, x^4+y^4, $x^4+y^4+(x+y)^4$, x^3y and $(x+iy)^4+(x-iy)^4$. It will follow from the next results that $\mathcal{C}((x^2+y^2)^2)=\{3,4\}$.

Theorem 5.5. *If $d=2k$ and $f(x,y)=\binom{2k}{k}x^ky^k$, then $L_{\mathbb{C}}(f)=k+1$. The minimal $\mathbb{C}$-representations of f are given by*

$$(k+1)\binom{2k}{k}x^ky^k=\sum_{j=0}^{k}(\zeta_{2k+2}^{j}wx+\zeta_{2k+2}^{-j}w^{-1}y)^{2k},\qquad 0\neq w\in\mathbb{C}. \tag{5.6}$$

Proof. We first evaluate the right-hand side of (5.6) by expanding the powers:

$$\begin{aligned}\sum_{j=0}^{k}(\zeta_{2k+2}^{j}wx+\zeta_{2k+2}^{-j}w^{-1}y)^{2k}&=\sum_{j=0}^{k}\sum_{t=0}^{2k}\binom{2k}{t}\zeta_{2k+2}^{j(2k-t)-jt}w^{(2k-t)-t}x^{2k-t}y^{t}\\&=\sum_{t=0}^{2k}\binom{2k}{t}w^{2k-2t}x^{2k-t}y^{t}\left(\sum_{j=0}^{k}\zeta_{k+1}^{j(k-t)}\right).\end{aligned}\tag{5.7}$$

But $\sum_{j=0}^{m-1}\zeta_m^{rj}=0$ unless $m\mid r$, in which case it equals m. Since the only multiple of $k+1$ in the set $\{k-t:0\le t\le 2k\}$ occurs for $t=k$, (5.7) reduces to the left-hand side of (5.6). We now show that these are *all* the minimal $\mathbb{C}$-representations of f.

Since $H_k(x^ky^k)$ has 1's on the NE-SW diagonal, it is non-singular, so $L_{\mathbb{C}}(x^ky^k)>k$, and $L_{\mathbb{C}}(x^ky^k)=k+1$ by (5.6). By Corollary 4.3, any minimal $\mathbb{C}$-representation *not* given by (5.6) can only use powers of forms which are distinct from *any* $wx+w^{-1}y$. If $ab=c^2\neq 0$, then $ax+by$ is a multiple of $\frac{a}{c}x+\frac{c}{a}y$. This leaves only x^{2k} and y^{2k}, and there is no linear combination of these giving x^ky^k. □

The representations in (5.6) arise because the null-vectors of $H_{k+1}(x^ky^k)$ can only be $(c_0,0,\dots,0,c_{k+1})^t$ and $c_0x^{k+1}+c_{k+1}y^{k+1}$ is a Sylvester form when $c_0c_{k+1}\neq 0$.

Corollary 5.6. *For $k\ge 2$, $L_{\mathbb{C}}((x^2+y^2)^k)=k+1$, and $L_K((x^2+y^2)^k)=k+1$ iff $\tan\frac{\pi}{k+1}\in K$. The $\mathbb{C}$-minimal representations of $(x^2+y^2)^k$ are given by*

$$\binom{2k}{k}(x^2+y^2)^k=\frac{1}{k+1}\sum_{j=0}^{k}\left(\cos(\tfrac{j\pi}{k+1}+\theta)x+\sin(\tfrac{j\pi}{k+1}+\theta)y\right)^{2k},\quad \theta\in\mathbb{C}. \tag{5.8}$$

Proof. The invertible map $(x,y)\mapsto(x-iy,x+iy)$ takes x^ky^k into $(x^2+y^2)^k$. Setting $0\neq w=e^{i\theta}$ in (5.6) gives (5.8). If $\tan\alpha\neq 0$, then

$$(\cos\alpha\, x+\sin\alpha\, y)^{2k} = \cos^{2k}\alpha\cdot(x+\tan\alpha\, y)^{2k} = (1+\tan^2\alpha)^{-k}(x+\tan\alpha\, y)^{2k}.$$

Thus, if $\tan\alpha \in K$, then $(\cos\alpha\ x + \sin\alpha\ y)^{2k} \in K[x,y]$. Further, if $\cos\alpha = 0$, then $(\cos\alpha\, x + \sin\alpha\, y)^{2k} = y^{2k} \in K[x,y]$. Conversely, if $(\cos\alpha\ x + \sin\alpha\ y)^{2k} \in K[x,y]$ and $\cos\alpha \neq 0$, then the ratio of the coefficients of $x^{2k-1}y$ and x^{2k} equals $2k\tan\alpha$, which must be in K. It follows that $L_K((x^2+y^2)^k) = k+1$ if and only if there exists $\theta \in \mathbb{C}$ so that for each $0 \le j \le k$, either $\cos(\frac{j\pi}{k+1}+\theta) = 0$ or $\tan(\frac{j\pi}{k+1}+\theta) \in K$. Since $\tan\alpha, \tan\beta \in K$ imply $\tan(\alpha-\beta) \in K$ and $k \ge 2$, we see that (5.8) is a representation over K if and only if $\tan\frac{\pi}{k+1} \in K$. □

In particular, since $\tan\frac{\pi}{3} = \sqrt{3} \notin \mathbb{Q}$, $L_{\mathbb{Q}}((x^2+y^2)^2) > 3$ and so must equal 4. Thus, $\mathcal{C}((x^2+y^2)^2) = \{3,4\}$, as promised. Since $\tan\frac{\pi}{m}$ is irrational for $m \ge 5$ (see e.g. [40, Corollary 3.12]), it follows that $L_{\mathbb{Q}}((x^2+y^2)^k) = k+1$ only for $k = 1, 3$.

It is worth remarking that $x^k y^k$ is a highly singular complex form, as is $(x^2+y^2)^k$. However, as a *real* form, $(x^2+y^2)^k$ is interior to the real convex cone $Q_{2,2k}$. For real θ, the formula in (5.8) goes back at least to Friedman [21] in 1957. It was shown in [47] that all minimal *real* representations of $(x^2+y^2)^k$ have this shape. There is an equivalence between representations of $(x^2+y^2)^k$ as a real sum of $2k$-th powers and quadrature formulas on the circle – see [47]. In this sense, (5.8) can be traced back to Mehler [35] in 1864.

A real representation (1.1) of $(\sum x_i^2)^k$ (with positive real coefficients λ_j) is called a *Hilbert Identity*; Hilbert [20, 26] used such representations with rational coefficients to solve Waring's problem. Hilbert Identities have been important in studying quadrature problems on S^{n-1}, the Delsarte-Goethals-Seidel theory of spherical designs in combinatorics and for embedding questions in Banach spaces [47, Chaps. 8 and 9], as well as for explicit computations in Hilbert's 17th problem [48]. It can be shown that any such representation requires at least $\binom{n+k-1}{n-1}$ summands, and this bound also applies if negative coefficients λ_j are allowed. It is not known whether allowing negative coefficients can reduce the total number of summands. However, Blekherman [3] has recently constructed $f \in Q_{6,4}$ which has a smaller length if one allows negative λ_j in a real representation. When $(\sum x_i^2)^k$ is a sum of exactly $\binom{n+k-1}{n-1}$ $2k$-th powers, the coordinates of minimal representations can be used to produce tight spherical designs. Such representations exist when $n = 2$, $2k = 2$, $(n,2k) = (3,4)$, $(n,2k) = (u^2-2,4)$ $(u = 3,5)$, $(n,2k) = (3v^2-4,6)$ $(v = 2,3)$, $(n,2k) = (24,10)$. It has been proved that they do not exist otherwise, unless possibly $(n,2k) = (u^2-2,4)$ for some odd integer $u \ge 7$ or $(n,2k) = (3v^2-4,6)$ for some integer $v \ge 4$. These questions have been largely open for more than thirty years. It is also not known whether there exist (k,n) so that $L_{\mathbb{R}}((\sum x_i^2)^k) > L_{\mathbb{C}}((\sum x_i^2)^k)$, although this cannot happen for $n = 2$. For that matter, it is not known whether there exists any $f \in Q_{n,d}$ so that $L_{\mathbb{R}}(f) > L_{\mathbb{C}}(f)$.

We conclude this section with a related question: if $f_\lambda(x,y) = x^4 + 6\lambda x^2y^2 + y^4$ for $\lambda \in \mathbb{Q}$, what is $L_{\mathbb{Q}}(f_\lambda)$? If $\lambda \le -\frac{1}{3}$, then f_λ has four real factors, so $L_{\mathbb{Q}}(f_\lambda) = 4$. Since $\det H_2(f_\lambda) = \lambda - \lambda^3$, $L_{\mathbb{C}}(f_\lambda) = 2$ for $\lambda = 0, 1, -1$. The formula

$$(x^4 + 6\lambda x^2y^2 + y^4) = \tfrac{\lambda}{2}\left((x+y)^4 + (x-y)^4\right) + (1-\lambda)(x^4+y^4)$$

shows that $L_{\mathbb{Q}}(f_0) = L_{\mathbb{Q}}(f_1) = 2$; $2f_{-1}(x,y) = (x+iy)^4 + (x-iy)^4$ has $\mathbb{Q}$-length 4.

Theorem 5.7. *Suppose $\lambda = \frac{a}{b} \in \mathbb{Q}, \lambda^3 \ne \lambda$. Then $L_{\mathbb{Q}}(x^4 + 6\lambda x^2y^2 + y^4) = 3$ if and only if there exist integers $(m,n) \ne (0,0)$ so that*

$$\Gamma(a,b,m,n) = 4a^3b\,m^4 + (b^4 - 6a^2b^2 - 3a^4)m^2n^2 + 4a^3b\,n^4 \tag{5.9}$$

is a non-zero square.

Proof. By Corollary 2.2, such a representation occurs if and only if there is a cubic $h(x,y) = \sum_{i=0}^{3} c_i x^{3-i} y^i$ which splits over $\mathbb{Q}$ and satisfies

$$c_0 + \lambda c_2 = \lambda c_1 + c_3 = 0. \tag{5.10}$$

Assume that $h(x,y) = (mx + ny)g(x,y)$, $(m,n) \ne (0,0)$ with $m,n \in \mathbb{Z}$. If $g(x,y) = rx^2 + sxy + ty^2$, then $c_0 = mr, c_1 = ms + nr, c_2 = mt + ns, c_3 = nt$ and (5.10) becomes

$$\begin{pmatrix} m & \lambda n & \lambda m \\ \lambda n & \lambda m & n \end{pmatrix} \cdot \begin{pmatrix} r \\ s \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{5.11}$$

If $m = 0$, then the general solution to (5.11) is $(r,s,t) = (r, 0, -\lambda r)$ and $rx^2 - \lambda ry^2$ splits over $\mathbb{Q}$ into distinct factors iff λ is a non-zero square; that is, iff ab is a square, and similarly if $n = 0$. Otherwise, the system has full rank since $\lambda^2 \ne 1$ and any solution is a multiple of

$$rx^2 + sxy + ty^2 = (\lambda n^2 - \lambda^2 m^2)x^2 + (\lambda^2 - 1)mnxy + (\lambda m^2 - \lambda^2 n^2)y^2. \tag{5.12}$$

The quadratic in (5.12) splits over $\mathbb{Q}$ into distinct factors iff its discriminant

$$4\lambda^3 m^4 + (1 - 6\lambda^2 - 3\lambda^4)m^2n^2 + 4\lambda^3 n^4 = b^{-4}\Gamma(a,b,m,n) \tag{5.13}$$

is a non-zero square in $\mathbb{Q}$. □

In particular, we have the following identities: $\Gamma(u^2, v^2, v, u) = (u^5v - uv^5)^2$ and $\Gamma(uv, u^2 - uv + v^2, 1, 1) = (u-v)^6(u+v)^2$, hence $L_{\mathbb{Q}}(f_\lambda) = 3$ for $\lambda = \tau^2$ and $\lambda = \frac{\tau}{\tau^2 - \tau + 1}$, where $\tau = \frac{u}{v} \in \mathbb{Q}$, $\tau \ne \pm 1$. These show that $L_{\mathbb{Q}}(f_\lambda) = 3$ for a dense set of rationals in $[-\frac{1}{3}, \infty)$. These families do not exhaust the possibilities. If $\lambda = \frac{38}{3}$, so $f_\lambda(x,y) = x^4 + 76x^2y^2 + y^4$, then λ is expressible neither as τ^2 nor $\frac{\tau}{\tau^2-\tau+1}$ for $\tau \in \mathbb{Q}$, but $\Gamma(38, 3, 2, 19) = 276{,}906^2$.

We mention two negative cases: if $\lambda = \frac{1}{3}$, $\Gamma(1,3,m,n) = 12(m^2+n^2)^2$, which is never a square, giving another proof that $L_{\mathbb{Q}}((x^2+y^2)^2) = 4$. If $\lambda = \frac{1}{2}$, then

$$\Gamma(1,2,m,n) = 8m^4 - 11m^2n^2 + 8n^4 = \tfrac{27}{4}(m^2-n^2)^2 + \tfrac{5}{4}(m^2+n^2)^2,$$

hence if $L_{\mathbb{Q}}(x^4 + 3x^2y^2 + y^4) = 3$, then there is a solution to the Diophantine equation $27X^2 + 5Y^2 = Z^2$. A simple descent shows that this has no non-zero solutions: working mod 5, we see that $2X^2 = Z^2$; since 2 is not a quadratic residue mod 5, it follows that $5 \mid X, Z$, and these imply that $5 \mid Y$ as well. It follows that, $L_{\mathbb{Q}}(x^4 + 3x^2y^2 + y^4) = 4$.

Solutions of the Diophantine equation $Am^4 + Bm^2n^2 + Cn^4 = r^2$ were first studied by Euler; see [16, pp. 634–639] and [38, pp. 16–29] for more on this topic. This equation has not yet been completely solved; see [6, 13]. We hope to return to the analysis of (5.9) in a future publication.

6 Open Questions

We are confident that Conjecture 4.12 can be completely settled. This raises the question of whether there exist other fields besides $\mathbb{C}$ (and possibly $\mathbb{R}$) for which there is a simple description of $\{f : L_K(f) = \deg f\}$.

Which cabinets are possible for binary forms? Are there other restrictions beyond Corollary 5.1(1)? How many different lengths are possible? If $|\mathcal{C}(f)| \geq 4$, then $d \geq 7$. Can anything more be said about forms in $n \geq 3$ variables?

Can f have more than one, but a finite number, of K-minimal representations, where K is not necessarily equal to E_f? Theorem 5.7 might be a way to find such examples.

Length is generic over $\mathbb{C}$, but not over $\mathbb{R}$. For $d = 2r$, the $\mathbb{R}$-length of a real form is always $2r$ in a small neighborhood of $\prod_{j=1}^{d}(x - jy)$, but the $\mathbb{R}$-length is always $r+1$ in a small neighborhood of $(x^2+y^2)^r$ [47]. Which combinations of degrees and lengths have interior? Does the parity of d matter? This question is explored in much greater detail in [15].

References

1. J. Alexander and A. Hirschowitz, *Polynomial interpolation in several variables*, J. Algebraic Geom., **4** (1995), 201–222, MR1311347 (96f:14065).
2. A. Białynicki-Birula and A. Schinzel, *Extreme binary forms*, Acta Arith. **142** (2010), 219–249, MR2606966 (2011d:12001).
3. G. Blekherman, personal correspondence.
4. M. Boij, E. Carlini, A. Geramita, *Monomials as sums of powers: the real binary case*, Proc. Amer. Math. Soc. **139** (2011), 3039–3043, arXiv:1005.3050, MR2811260 (2012e:11070).

5. M. Brambilla and G. Ottaviani, *On the Alexander-Hirschowitz theorem*, J. Pure Appl. Algebra, **212** (2008), 1229–1251, arXiv:math/0701409, MR2387598 (2008m:14104).
6. E. Brown, $x^4 + dx^2y^2 + y^4 = z^2$*: some cases with only trivial solutions—and a solution Euler missed*, Glasgow Math. J., 31 (1989), 297–307, MR1021805 (91d:11026).
7. E. Carlini, *Varieties of simultaneous sums of power for binary forms*, Matematiche (Catania), **57** (2002), 83–97, arXiv:math.AG/0202050, MR2075735 (2005d:11058).
8. E. Carlini and J. Chipalkatti, *On Waring's problem for several algebraic forms*, Comment. Math. Helv., **78** (2003), 494–517, arXiv:math.AG/0112110, MR1998391 (2005b:14097).
9. A. Causa and R. Re, *On the maximum rank of a real binary form*, Annali di Matematica Pura et Applicata (4) **190** (2011), no. 1, 55–59, arXiv:1006.5127, MR2747463 (2011k:12001).
10. M. D. Choi, Z. D. Dai, T. Y. Lam and B. Reznick, *The Pythagoras number of some affine algebras and local algebras*, J. Reine Angew. Math., **336** (1982), 45–82, MR0671321 (84f:12012).
11. M. D. Choi, T. Y. Lam, A. Prestel and B. Reznick, *Sums of* $2m$*th powers of rational functions in one variable over real closed fields*, Math. Z., **221** (1996), 93–112, MR1369464 (96k:12003).
12. M. D. Choi, T. Y. Lam and B. Reznick, *Sums of squares of real polynomials*, K-theory and algebraic geometry: connections with quadratic forms and division algebras (Santa Barbara, CA, 1992), 103–126, Proc. Sympos. Pure Math., **58**, Part 2, Amer. Math. Soc., Providence, RI, 1995, MR1327293 (96f:11058).
13. J. H. E. Cohn, *On the Diophantine equation* $z^2 = x^4 + Dx^2y^2 + y^4$, Glasgow Math. J., **36** (1994), 283–285, MR1295501 (95k:11035).
14. P. Comon and B. Mourrain, *Decomposition of quantics in sums of powers of linear forms*, Signal Processing, **53** (1996), 93–107.
15. P. Comon and G. Ottaviani, *On the typical rank of real binary forms*, Linear Multilinear Algebra, **60** (2012), 657–667, MR 2929176, (arXiv:0909.4865).
16. L. E. Dickson, *History of the Theory of Numbers, vol II: Diophantine Analysis*, Carnegie Institute, Washington 1920, reprinted by Chelsea, New York, 1971, MR0245500 (39 #6807b).
17. I. Dolgachev and V. Kanev, *Polar covariants of plane cubics and quartics*, Adv. Math., **98** (1993), 216–301, MR1213725 (94g:14029).
18. R. Ehrenborg and G.-C. Rota, *Apolarity and canonical forms for homogeneous polynomials*, European J. Combin., **14** (1993), 157–181, MR1215329 (94e:15062).
19. W. J. Ellison, *A 'Waring's problem' for homogeneous forms*, Proc. Cambridge Philos. Soc., **65** (1969), 663–672, MR0237450 (38 #5732).
20. W. J. Ellison, *Waring's problem*, Amer. Math. Monthly, **78** (1971), 10–36, MR0414510 (54 #2611).
21. A. Friedman, *Mean-values and polyharmonic polynomials*, Michigan Math. J., **4** (1957), 67–74, MR0084045 (18,799b).
22. A. Geramita, *Inverse systems of fat points: Waring's problem, secant varieties of Veronese varieties and parameter spaces for Gorenstein ideals*, The Curves Seminar at Queen's, Vol. X (Kingston, ON, 1995), 2–114, Queen's Papers in Pure and Appl. Math., **102**, Queen's Univ., Kingston, ON, 1996, MR1381732 (97h:13012).
23. S. Gundelfinger, *Zur Theorie der binären Formen*, J. Reine Angew. Math., **100** (1886), 413–424.
24. J. Harris, *Algebraic geometry. A first course*, Graduate Texts in Mathematics, **133**, Springer-Verlag, New York, 1992, MR1182558 (93j:14001).
25. U. Helmke, *Waring's problem for binary forms*, J. Pure Appl. Algebra, **80** (1992), 29–45, MR1167385 (93e:11057).
26. D. Hilbert, *Beweis für die Darstellbarkeit der ganzen Zahlen durch eine feste Anzahl* n*-ter Potenzen (Waringsches Problem)*, Math. Ann., **67** (1909), 281–300, Ges. Abh. 1, 510–527, Springer, Berlin, 1932, reprinted by Chelsea, New York, 1981.
27. P. Holgate, *Studies in the history of probability and statistics. XLI. Waring and Sylvester on random algebraic equations*, Biometrika, **73** (1986), 228–231, MR0836453 (87m:01026).
28. A. Iarrobino and V. Kanev, *Power Sums, Gorenstein algebras, and determinantal loci*, Lecture Notes in Mathematics, **1721** (1999), MR1735271 (2001d:14056).

29. S. Karlin, *Total Positivity, vol. 1*, Stanford University Press, Stanford, 1968, MR0230102 (37 #5667).
30. J. P. S. Kung, *Gundelfinger's theorem on binary forms*, Stud. Appl. Math., **75** (1986), 163–169, MR0859177 (87m:11020).
31. J. P. S. Kung, *Canonical forms for binary forms of even degree*, in *Invariant theory*, Lecture Notes in Mathematics, **1278**, 52–61, Springer, Berlin, 1987, MR0924165 (89h:15037).
32. J. P. S. Kung, *Canonical forms of binary forms: variations on a theme of Sylvester*, in *Invariant theory and tableaux (Minnesota, MN, 1988)*, 46–58, IMA Vol. Math. Appl., **19**, Springer, New York, 1990, MR1035488 (91b:11046).
33. J. P. S. Kung and G.-C. Rota, *The invariant theory of binary forms*, Bull. Amer. Math. Soc. (N. S.), **10** (1984), 27–85, MR0722856 (85g:05002).
34. J. M. Landsberg and Z. Teitler, *On the ranks and border ranks of symmetric tensors*, Found. Comp. Math., **10** (2010), 339–366, arXiv:0901.0487, MR2628829 (2011d:14095).
35. G. Mehler, *Bemerkungen zur Theorie der mechanischen Quadraturen*, J. Reine Angew. Math., **63**, (1864), 152–157.
36. R. Miranda, *Linear systems of plane curves*, Notices Amer. Math. Soc., **46** (1999), 192–202, MR1673756 (99m:14012).
37. L. J. Mordell, *Binary cubic forms expressed as a sum of seven cubes of linear forms*, J. London. Math. Soc., **42** (1967), 646–651, MR0249355 (40 #2600).
38. L. J. Mordell, *Diophantine equations*, Academic Press, London-New York 1969, MR0249355 (40 #2600).
39. D. J. Newman and M. Slater, *Waring's problem for the ring of polynomials*, J. Number Theory, **11** (1979), 477–487, MR0544895 (80m:10016).
40. I. Niven, *Irrational numbers*, Carus Mathematical Monographs, No. 11. Math. Assoc. Amer., New York, 1956, MR0080123 (18,195c).
41. G. Pólya and I. J. Schoenberg, *Remarks on de la Vallée Poussin means and convex conformal maps of the circle*, Pacific J. Math., **8** (1958), 295–234, MR0100753 (20 #7181).
42. G. Pólya and G. Szegö, *Problems and theorems in analysis, II*, Springer-Verlag, New York-Heidelberg 1976, MR0465631 (57 #5529).
43. V. Powers and B. Reznick, *Notes towards a constructive proof of Hilbert's theorem on ternary quartics*, Quadratic forms and their applications (Dublin, 1999), 209–227, Contemp. Math., **272**, Amer. Math. Soc., Providence, RI, 2000, MR1803369 (2001h:11049).
44. K. Ranestad and F.-O. Schreyer, *Varieties of sums of powers*, J. Reine Angew. Math., **525** (2000), 147–181, MR1780430 (2001m:14009).
45. B. Reichstein, *On expressing a cubic form as a sum of cubes of linear forms*, Linear Algebra Appl., **86** (1987), 91–122, MR0870934 (88e:11022).
46. B. Reichstein, *On Waring's problem for cubic forms*, Linear Algebra Appl., **160** (1992), 1–61, MR1137842 (93b:11048).
47. B. Reznick, *Sums of even powers of real linear forms*, Mem. Amer. Math. Soc., **96** (1992), no. 463, MR1096187 (93h:11043).
48. B. Reznick, *Uniform denominators in Hilbert's Seventeenth Problem*, Math. Z., **220** (1995), 75–97, MR1347159 (96e:11056).
49. B. Reznick, *Homogeneous polynomial solutions to constant coefficient PDE's*, Adv. Math., **117** (1996), 179–192, MR1371648 (97a:12006).
50. B. Reznick, *Laws of inertia in higher degree binary forms*, Proc. Amer. Math. Soc., **138** (2010), 815–826, arXiv:0906.5559, MR26566547 (2011e:11074).
51. B. Reznick, *Blenders*, in Notions of Positivity and the Geometry of Polynomials (P. Branden, M. Passare, M. Putinar, editors), Trends in Math., Birkhuser, Basel, 2011, pp. 345–373, arXiv:1008.4533v1.
52. G. Salmon, *Lesson introductory to the modern higher algebra*, fifth edition, Chelsea, New York, 1964.

53. J.J. Sylvester, *An Essay on Canonical Forms, Supplement to a Sketch of a Memoir on Elimination, Transformation and Canonical Forms*, originally published by George Bell, Fleet Street, London, 1851; Paper 34 in *Mathematical Papers*, Vol. 1, Chelsea, New York, 1973. Originally published by Cambridge University Press in 1904.
54. J. J. Sylvester, *On a remarkable discovery in the theory of canonical forms and of hyperdeterminants*, originally in Phiosophical Magazine, vol. 2, 1851; Paper 42 in *Mathematical Papers*, Vol. 1, Chelsea, New York, 1973. Originally published by Cambridge University Press in 1904.
55. J. J. Sylvester, *On an elementary proof and demonstration of Sir Isaac Newton's hitherto undemonstrated rule for the discovery of imaginary roots*, Proc. Lond. Math. Soc. **1** (1865/1866), 1–16; Paper 84 in *Mathematical Papers*, Vol.2, Chelsea, New York, 1973. Originally published by Cambridge University Press in 1908.
56. J.-C. Yakoubsohn, *On Newton's rule and Sylvester's theorems*, J. Pure Appl. Algebra, **65** (1990), 293–309, MR1072286 (91j:12002).

Representation of Quadratic Forms by Integral Quadratic Forms

Rainer Schulze-Pillot

Abstract The number of representations of a positive definite integral quadratic form of rank n by another positive definite integral quadratic form of rank $m \geq n$ has been studied by arithmetic, analytic, and ergodic methods. We survey and compare in this article the results obtained by these methods.

Key words Representation of quadratic forms • Integral quadratic forms

Mathematics Subject Classification (2000): Primary 11E12, Secondary 11E04

1 Introduction

It is a classical problem to study the solvability and the number of integral solutions of the quadratic diophantine equation

$$\sum_{i,j=1}^{m} a_{ij} x_i x_j = t$$

for an integral symmetric matrix $A = (a_{ij})$ and an integer t. Already Gauß studied more generally systems of such equations of the form

$${}^{t}XAX = T$$

R. Schulze-Pillot (✉)
Fachrichtung 6.1 Mathematik, Universität des Saarlandes, Saarbrücken, Germany
e-mail: schulzep@math.uni-sb.de

K. Alladi et al. (eds.), *Quadratic and Higher Degree Forms*, Developments in Mathematics 31, DOI 10.1007/978-1-4614-7488-3_9,

where now T is another (half-) integral symmetric matrix of size $n \leq m$. If one looks for rational instead of integral solutions, the Hasse-Minkowski theorem states the validity of the local-global principle for this problem, i.e., a rational solution exists if and only if solutions exist over $\mathbb{R}$ and over all p-adic fields $\mathbb{Q}_p$. That the local-global principle fails for integral solutions is already seen in simple examples like $Q(x_1, x_2) = 5x_1^2 + 11x_2^2$ which represents 1 over $\mathbb{R}$ and over all $\mathbb{Z}_p$ but not over $\mathbb{Z}$. Whereas the integral local-global principle can be saved with some modifications for indefinite A by the theory of spinor genera, the best possible results in the definite case prove that local representability implies global representability for T that are sufficiently large in a suitable sense and yield asymptotic formulas for T which are locally represented. The case of one equation ($n = 1$) is already classical, and considerable effort has been spent in the last 60 years on the case of $n > 1$, using both analytic and purely arithmetic methods. The introduction of ergodic theory as a new tool in [16] by Ellenberg and Venkatesh in 2008 has brought dramatic progress, it builds on the arithmetic approach of Eichler and Kneser and is also inspired by work of Linnik on representation of integers by ternary quadratic forms.

In this survey we sketch all three approaches (arithmetic, analytic, ergodic) and compare their results. At present each of the methods gives results which cannot be achieved by one of the others.

I thank G. Harcos and the referee for some helpful remarks.

2 Statement of the Problem and Notations

An integral valued quadratic form $Q = Q_A$ on $\mathbb{Z}^m$ is given as $Q_A(\mathbf{x}) = \frac{1}{2}{}^t\mathbf{x}A\mathbf{x}$, where $A \in M_m^{\text{sym}}(\mathbb{Z})$ is an integral symmetric matrix with even diagonal. Associated to it are the symmetric bilinear forms $b(\mathbf{x}, \mathbf{y}) = Q(\mathbf{x} + \mathbf{y}) - Q(\mathbf{x}) - Q(\mathbf{y}) = {}^t\mathbf{x}A\mathbf{y}$ and $B = \frac{1}{2}b$ with $B(\mathbf{x}, \mathbf{x}) = Q(\mathbf{x})$. One says that the symmetric matrix T of size n is represented by Q_A or that Q_{2T} is represented by Q_A over $\mathbb{Z}$ (resp. over $\mathbb{Q}$) if there is $X \in M_{m,n}(\mathbb{Z})$ (resp. $\in M_{m,n}(\mathbb{Q})$) with $Q_A(X) = \frac{1}{2}{}^t XAX = T$.

An integral representation is called primitive if all elementary divisors of X are 1, in particular for $n = 1$ this says that the coefficients $x_1, \dots, x_m$ of the representing vector $\mathbf{x} \in \mathbb{Z}^m$ are relatively prime. If the matrix A is positive definite the matrix equation $Q_A(\mathbf{x}) = T$ has only finitely many solutions over $\mathbb{Z}$ and one calls

$$r(A, T) := \#\{X \in M_{m,n}(\mathbb{Z}) \mid Q_A(X) = T\}$$

the number of representations of T by A. The matrices $A_1, A_2 \in M_m^{\text{sym}}(\mathbb{Z})$ and their associated quadratic forms Q_{A_1}, Q_{A_2} are called rationally resp. integrally equivalent if the equations $A_1 =^t X_1 A_2 X_1$, $A_2 =^t X_2 A_1 X_2$ are solvable with $X_1, X_2 \in \mathrm{GL}_m(\mathbb{Q})$ resp. in $\mathrm{GL}_m(\mathbb{Z})$. Clearly, integrally equivalent forms represent the same numbers and matrices and have the same representation numbers. The forms Q_{A_1}, Q_{A_2} (or their associated symmetric matrices) are said to be locally everywhere integrally (resp. rationally) equivalent, if ${}^tXA_1X = A_2$ is solvable with $X \in \mathrm{GL}_m(\mathbb{Z}_p)$ (resp. $\in \mathrm{GL}_m(\mathbb{Q}_p)$) for all primes p and A_1 and A_2 have the same signature (i.e., ${}^tXA_1X = A_2$ is solvable in $\mathrm{GL}_m(\mathbb{R})$). Forms which are locally everywhere integrally equivalent are said to belong to the same genus. Analogously one defines the notion that T is locally everywhere representable by Q_A (integrally or rationally). The Hasse-Minkowski theorem [17] asserts that rational representation locally everywhere is equivalent to representation (over $\mathbb{Q}$); this is not true for integral representation, where representation over $\mathbb{Z}$ is stronger than representation locally everywhere.

Since it is in principle easy to determine the numbers or matrices which are represented locally everywhere by determining the solvability of finitely many congruences, the problem to determine all T which are represented by Q_A is reduced to

Problem 1. *Given $A \in M_m^{\mathrm{sym}}(\mathbb{Z})$ determine conditions on $T \in M_n^{\mathrm{sym}}(\mathbb{Z})$ (with $n \leq m$) such that T meeting these conditions is represented integrally by Q_A if it is represented locally everywhere integrally by Q_A.*

Similarly, determine conditions under which primitive representability locally everywhere implies (primitive) representability over $\mathbb{Z}$.

For many purposes it is convenient to use the equivalent but slightly more flexible language of quadratic spaces and lattices in them which has been introduced by Witt; in particular for generalizations to forms over number fields and their integers it is the more natural framework:

Definition 1. *Let F be a field. A quadratic space (V, Q) over F is a finite dimensional vector space V over F equipped with a map $Q : V \to F$ satisfying*

(a) $Q(ax) = a^2Q(x)$ *for all* $x \in V$, $a \in F$
(b) $b(x, y) := Q(x + y) - Q(x) - Q(y)$ *defines a symmetric bilinear form on* V.

The map Q is called the quadratic form on V and b is called its associated bilinear form.

If $\mathcal{B} = (e_1, \dots, e_m)$ is a basis of V we call the matrix $M_{\mathcal{B}}(Q) := (b(e_i, e_j)) \in M_m^{\mathrm{sym}}(F)$ the Gram matrix of (V, Q) (or just of Q) with respect to $\mathcal{B}$.

If (V', Q') is another quadratic space over F a linear isomorphism $f : V \to V'$ is called an isometry if $Q'(f(x)) = Q(x)$ holds for all $x \in V$. If an isometry $f : (V, Q) \to (V', Q')$ exists one says that the spaces are isometric or equivalent or that they belong to the same class.

If the mapping f above is just injective but may fail to be surjective it is called an isometric embedding of (V, Q) into (V', Q') and one says that (V, Q) is represented by (V', Q').

The geometric formulation of integral quadratic forms is obtained by considering lattices on quadratic spaces. In the most classical case we have:

Definition and Lemma 2. *Let (V, Q) be a quadratic space over the field $\mathbb{Q}$ of rational numbers.*

A $\mathbb{Z}$- lattice (or simply lattice) on V is a finitely generated $\mathbb{Z}$-submodule L of V which generates V over $\mathbb{Q}$.

Equivalently, $L = \bigoplus_{i=1}^{m} \mathbb{Z}e_i$ for some basis $(e_1, \ldots, e_m)$ of the space V (which is then also a basis of the $\mathbb{Z}$-module L). We also call (L, Q) a quadratic lattice.

The lattice L is called integral if $Q(L) \subseteq \mathbb{Z}$.

If (L, Q) is an integral $\mathbb{Z}$-lattice and $\mathcal{B} = (e_1, \ldots, e_m)$ is a basis of L, the quadratic polynomial $P_{Q,\mathcal{B}}(x_1, \ldots, x_m) = Q(\sum_{i=1}^{m} x_i e_i)$ has integral coefficients; this polynomial is then what is usually called an integral valued quadratic form (see [9]). One obtains a classically integral quadratic form in the sense of [9] if in addition the bilinear form $B = b/2$ assumes integral values.

Since we will also consider the number field situation we need the following more general definition:

Definition 3. *Let F be a number field and R its ring of integers or let F be the completion of a number field at a non-archimedean place and again R its ring of integers.*

A finitely generated R-submodule L of V is called an R-lattice on V if L generates V over F. We also call (L, Q) a quadratic lattice over R. The lattice L is called integral (with respect to R) if $Q(L) \subseteq R$.

If the lattice L is free with basis $\mathcal{B} = (e_1, \ldots, e_m)$ over R the matrix $A = (b(e_i, e_j)) \in M_m^{\mathrm{sym}}(\mathbb{F})$ is called its Gram matrix with respect to $\mathcal{B}$.

Lattices L on (V, Q) and L' on (V', Q') are called isometric or equivalent if there is an isometry $f : (V', Q') \to (V, Q)$ with $f(L') = L$; one writes $L' \in \mathrm{cls}(L)$ and also says that L and L' belong to the same class.

The lattice L' is said to be represented by L if there is an isometric embedding $f : (V', Q') \to (V, Q)$ with $f(L') \subseteq L$. We write $r(L, L')$ for the number of such representations if this number is finite.

The representation f is called primitive if $f(L') = L \cap f(V')$. For $a \in R$ it is called of imprimitivity bounded by a if $a(L \cap f(V')) \subseteq f(L')$.

If F is totally real and L is (totally) positive definite we denote by $\min(L) := \min\{N_{\mathbb{Q}}^{F}(Q(x)) \mid x \in L, x \neq \mathbf{0}\}$ the minimum of the lattice L. (For the question which lattices have large minimum it does not matter whether we chose this definition or $\min\{Tr_{\mathbb{Q}}^{F}(Q(x)) \mid x \in L, x \neq \mathbf{0}\}$ instead, see the remark in [20, p.139]*.)*

Remark. (a) If one wants to use the language of matrices instead of that of lattices and R is no principal ideal domain, one has to consider Gram matrices with respect to linearly dependent generating sets (see [48]); this is one of the reasons why lattices give the more convenient framework.

(b) An equivalent characterization of a lattice on V is: L is an R-submodule of F, and for some basis $(x_1, \ldots, x_m)$ of V and some $c \in R$ one has $cL \subseteq Rx_1 + \ldots + Rx_m \subseteq L$. If R is a principal ideal domain we can instead require $L = Rx_1 + \ldots + Rx_m$ for some basis of V as before. If R is no PID one admits non free lattices as well.

(c) If R is the ring of integers of the number field F and $S = R_v$ its completion at some place v we write (L_v, Q) for the extension of (L, Q) to R_v and call it the completion of L at v; if v is archimedean we have $R_v = F_v$ and $L_v = V_v$.

In the sequel we let F be a number field with ring of integers R.

The R-lattices Λ, Λ' are in the same genus ($\Lambda' \in \mathrm{gen}(\Lambda)$) if Λ_v is isometric to Λ'_v for all places v of F. The R-lattice N is represented by Λ locally everywhere if N_v is represented by Λ_v for all places v of F. If the lattices in question are free and have Gram matrices A, T with respect to bases $\mathcal{B}$ of Λ and $\mathcal{B}'$ of Λ' resp. N, the notions of equivalence, genus, (primitive) representation (locally everywhere) for lattices given above translate into those for symmetric matrices described earlier in this section.

Problem 1 from above becomes

Problem 1'. Given an R-lattice Λ of rank m describe conditions on an R-lattice N of rank $n \leq m$ such that N satisfying these conditions is represented by Λ (primitively) if it is represented by Λ (primitively) locally everywhere. If possible give (approximate) formulas for the numbers or measures of representations.

3 Siegel's Theorem

Although a strict local-global principle is not valid for representation of numbers or forms by integral quadratic forms, the Hasse-Minkowski theorem for quadratic forms over a number field has the consequence that a lattice N which is represented by the lattice Λ locally everywhere (primitively) is represented by some lattice Λ' in the genus of Λ (primitively), see [9,38]. Siegel's celebrated theorem in fact gives the so-called mass formula (German: Maßformel, verbal translation to English: measure formula) for the average number of representations of K by Λ.

Theorem 4. *Let $\{L_1, \ldots, L_h\}$ be a set of representatives of the classes of lattices in the genus of Λ. If Q is definite put $w = \sum_{i=1}^{h} \frac{1}{|O(L_i)|}$ (where $O(L_i)$ is the group of isometries of L onto itself with respect to Q) and write*

$$r(\mathrm{gen}\,\Lambda, N) = \frac{1}{w} \sum_{i=1}^{h} \frac{r(L_i, N)}{|O(L_i)|}$$

for Siegel's weighted average of the representation numbers of N by the lattices L_i in the genus of Λ.

If Q is not definite and neither the space $F\Lambda$ nor the orthogonal complement of a representation of FN in $F\Lambda$ is a hyperbolic plane the measure (mass) $w = \mu(\Lambda)$ of Λ and the representation measures $\mu(L_i, N)$ of N by the L_i can be defined as in [48] and one puts

$$r(\mathrm{gen}\,\Lambda, N) = \frac{1}{w} \sum_{i=1}^{h} \mu(L_i, N).$$

Then $r(\operatorname{gen} L, N)$ can be expressed as a product of local densities over the non-archimedean places v of F,

$$r(\operatorname{gen} \Lambda, N) = c \cdot (N_{\mathbb{Q}}^F(\det N))^{\frac{m-n-1}{2}} (N_{\mathbb{Q}}^F(\det \Lambda))^{-\frac{n}{2}} \prod_v \alpha_v(\Lambda, N)$$

with some constant c.

Here by $N_{\mathbb{Q}}^F(\det(\Lambda))$ resp. $N_{\mathbb{Q}}^F(\det(N))$ we denote the norm of the ideal generated by the determinants of the Gram matrices with respect to linearly independent subsets of the respective lattice, the local density $\alpha_v(\Lambda, N)$ is for a non-archimedean place v of F with residue field of order q_v and local prime element $\omega_v \in R_v$ given as

$$\begin{aligned}\alpha_v(\Lambda, N) &= \alpha_v(S_v, T_v)\\ &= q_v^{j\cdot(\frac{n\cdot(n+1)}{2} - mn)} \# \mathcal{A}_j(S_v, T_v),\end{aligned}$$

for sufficiently large j with an additional factor $\frac{1}{2}$ if $m = n$, where S_v, T_v denote Gram matrices of the local lattices Λ_v, N_v and where we write

$$\mathcal{A}_j(S_v, T_v) = \{X = (x_{ij}) \in M_{m,n}(R_v)/\omega_v^j M_{m,n}(R_v) \mid^t X S_v X \equiv T \bmod \omega_v^j\}$$

with the congruence being required modulo integral symmetric matrices with even diagonal.

An analogous formula holds for averaged primitive representation numbers resp. measures and primitive local densities $\alpha_v^*(\Lambda, N)$ counting congruence solutions as above but with the representing matrix X being primitive.

The (primitive) local densities in Siegel's theorem are nonzero if N is represented (primitively) locally by Λ and their product converges, so the theorem implies that (as mentioned above) such an N is represented (primitively) by at least one class of lattices in the genus of Λ; it gives a quantitative version of this qualitative result.

If Λ happens to be in a genus of one class, as is the case e.g. for the lattice over $\mathbb{Z}$ corresponding to the sum of k integral squares with $k \leq 8$, Siegel's theorem gives an exact formula for $r(\Lambda, N)$ resp. the measure $\mu(\Lambda, N)$. Since one can give closed formulae for $\alpha_v(\Lambda, N)$ for almost all v (where the exceptional set depends on Λ) the average representation numbers or measures can be determined explicitly for given Λ by determining the numbers of solutions of finitely many congruence systems.

In the asymptotic formulas to be discussed later the average representation number $r(\operatorname{gen} \Lambda, N)$ will be the main term.

4 The Indefinite Case

For the rest of this article we restrict attention to quadratic spaces and lattices with non-degenerate quadratic form Q, as usual we will often suppress the quadratic form Q in the notation.

The case that Λ_v is isotropic (i.e., represents zero nontrivially) for at least one archimedean place v of F has been solved as completely as possible:

Theorem 5 (Eichler [15], Kneser [35], Weil [51], Hsia [19]). *Let Λ be a non-degenerate quadratic R-lattice of rank m such that Λ_v is isotropic for at least one archimedean place v of F.*

(a) *If the non-degenerate quadratic R-lattice N of rank $n \leq m-3$ is represented by Λ locally everywhere it is represented by Λ, and the measure of representations (Darstellungsmaß) of N by Λ' is the same for all lattices Λ' in the genus of Λ.*

(b) *If N is as above with $n = m-2$ then either the measure of representations of N by a lattice Λ' is the same for all Λ' in the genus of Λ or the genus of Λ splits into two half genera consisting of equally many classes such that the measure of representations of N by Λ', Λ'' is the same if Λ', Λ'' belong to the same half genus.*

The latter case occurs only for N for which the discriminant of the space FN (i.e., the determinant of a Gram matrix of that quadratic space) belongs to one of finitely many square classes depending on Λ which can be explicitly determined.

Remark. (a) The proof uses the theory of spinor genera, a modified version of it plays a role in the arithmetic and the ergodic approach to problem 1' for definite lattices, see Lemma 13 below.

(b) The measure of representations has been defined by Siegel in [47]; an equivalent definition using measures on adelic orthogonal groups is given e.g. in [35].

(c) In the case $n = 1$, $m = 3$, the difference between the representation measures of the half genera occurring in part (b) of the theorem has been calculated in [43]. The integers represented (primitively) locally everywhere but not globally by all classes in the genus have been determined explicitly in [42] without the primitivity condition and in [14] for the primitive case; they are called (primitive) spinor exceptions.

(d) Some further results for the case $n = m-2$ have been obtained in [52, 53]. The cases $n = m-1$ and $m = n$ do not admit clean solutions; what can be done is shown in [52, 53].

(e) The determination of the square classes in part (b) of the theorem is achieved by computing the spinor norms of the local orthogonal groups of the lattices N_v. This computation is given in [13, 18, 34].

(f) An analytic proof of the result for $n = 1$, $m = 4$ has been given by Siegel in [49].

5 Representation of Integers ($n = 1$)

By the results of the previous section we can restrict attention from now on to the case that F is totally real and Λ is totally (positive) definite. In this case the representation numbers

$$r(\Lambda, N) = \#\{\varphi : N \longrightarrow \Lambda \mid \varphi \text{ linear isometry}\}$$

are finite. The first general result here is the following theorem. It has been proven with the help of the Hardy Littlewood circle method in [32] by Kloosterman for diagonal forms and in [33] for general forms using both modular forms and the Hardy Littlewood method in 1927; Kloosterman's first proof has been generalized in [50] by Tartakovskii to general forms in 1929.

Theorem 6 (Kloosterman, Tartakovskii). *Let Λ be a positive definite $\mathbb{Z}$-lattice of rank $m \geq 5$. Then Λ represents all sufficiently large numbers t which are represented by it locally everywhere. The same is true for $m = 4$ if one restricts attention to t which are represented locally everywhere primitively or which satisfy at least for some fixed a that for each p there is $x \in \Lambda_p$ with $Q(x) = t$ and $a^{-1}x \notin \Lambda_p$ (one also says that t is represented locally everywhere with imprimitivity bounded by a).*

In both cases one has an asymptotic formula

$$r(\Lambda, t) = r(\operatorname{gen} \Lambda, t) + \mathrm{O}(t^{\frac{m}{4}-\delta})$$

for any $\delta < \frac{5}{16}$ for odd m and $\delta < \frac{1}{2}$ for even m, where the main term $r(\operatorname{gen} \Lambda, t)$ grows at least like $t^{\frac{m}{2}-1-\epsilon}$ for all $\epsilon > 0$ for t satisfying the conditions given.

Remark. (a) The condition on bounded imprimitivity for the local representations is automatically satisfied for all primes p for which Λ_p is isotropic (represents zero nontrivially), hence in particular for all p not dividing the determinant of Λ.

(b) The exponents in the error terms are better than the original ones; the bound for even m is the Ramanujan-Petersson bound (proven by Deligne), the bound for odd m is the bound from [4, (1.3)].

(c) In the Hardy Littlewood method the main term appears as the singular series.

This result already contains some essential features of the general situation (i.e., arbitrary n):

- Instead of an exact formula one has an asymptotic formula whose main term is determined by the local arithmetic of N and Λ.
- The asymptotic formula is unconditional for $m = 5 = 2 \cdot n + 3$ (with $n = 1$) and needs an additional primitivity assumption for $m = 4 = 2 \cdot n + 2 = n + 3$.
- Results for representation of sufficiently large integers follow directly from the asymptotic formula and can be made explicit.

The result has been generalized to the number field case and (as far as possible) to $m = 3$, for details see the survey [44] and notice that the bound in the error term for odd m has meanwhile been improved in [5]. All these results can be generalized to representations with congruence conditions and to statements about the equidistribution of lattice points on (higher dimensional) ellipsoid surfaces, see [12].

Remark. There are several results about representation of numbers by an integral quadratic form that don't fit well into this survey but should at least be mentioned:

- The 15-theorem of Conway and Schneeberger [2] states that a classically integral quadratic form represents all natural numbers if it represents all natural numbers up to 15. A modification of this result is the 290-conjecture, stating that an integral valued quadratic form represents all natural numbers if it represents all natural numbers up to 290; a proof of this conjecture has been announced by Bhargava and Hanke in 2008. A generalization of both these results to representation of quadratic forms has been proven by B. M. Kim, M.-H. Kim and Oh in [23].
- In several recent articles the representation of numbers by a quadratic form with restricted variables is investigated, e.g. [3] treats the number of representations of t as a sum of four squares whose largest prime factor is bounded ("smooth squares").

6 Representation of Forms ($n > 1$), Analytic Method

We continue to assume Λ to be positive definite (and F totally real), so that the representation number $r(\Lambda, N)$ is finite.

All results obtained for $r(\Lambda, N)$ with $n = \mathrm{rk}(N) > 1$ obtained so far by analytic methods are for the case $R = \mathbb{Z}$, $F = \mathbb{Q}$ and use the fact that the theta series of degree n of Λ is a Siegel modular form with respect to a suitable congruence subgroup of the modular group $\mathrm{Sp}_n(\mathbb{Z})$.

To fix some notation let

$$\mathrm{Sp}_n(\mathbb{R}) = \{g \in \mathrm{GL}_{2n}(\mathbb{R}) \mid {}^t g \begin{pmatrix} 0_n & -1_n \\ 1_n & 0_n \end{pmatrix} g = \begin{pmatrix} 0_n & -1_n \\ 1_n & 0_n \end{pmatrix}\}$$

be the real symplectic group of rank (degree, genus) n and $\mathcal{H}_n$ the Siegel upper half space of degree (genus) n, with $g = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{Sp}_n(\mathbb{R})$ acting by

$$Z \longmapsto g\langle Z\rangle := gZ := (AZ + B)(CZ + D)^{-1}.$$

A Siegel modular form of weight k for the congruence subgroup $\Gamma \subseteq \mathrm{Sp}_n(\mathbb{Z})$ is a holomorphic function $F : \mathcal{H}_n \longrightarrow \mathbb{C}$ satisfying $F(g\langle Z\rangle) = \det(CZ + D)^k F(Z)$ for all $g = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma$. If $\chi : \Gamma \longrightarrow \mathbb{C}^\times$ is a character we will also use Siegel modular forms with character (nebentype) χ, where one has $F(g\langle Z\rangle) = \chi(g)\det(CZ + D)^k F(Z)$ for $g \in \Gamma$. (If $n = 1$, one has to add a holomorphy condition at the cusps.) The theta series of degree (genus) n of the positive definite lattice Λ is given as

$$\vartheta^{(n)}(\Lambda, Z) = \sum_{\mathbf{x}=(x_1,\dots,x_n)\in\Lambda^n} \exp(2\pi i \mathrm{tr}(Q(\mathbf{x})Z))$$

with $Q(\mathbf{x}) = (B(x_i, x_j)) \in M_n^{\mathrm{sym}}(\frac{1}{2}\mathbb{Z})$, if S is a Gram matrix of Λ we can also write

$$\begin{aligned}\vartheta^{(n)}(\Lambda, Z) = \vartheta^{(n)}(S, Z) &= \sum_{X\in M_{m,n}(\mathbb{Z})} \exp(\pi i\,\mathrm{tr}({}^tXSXZ)) \\ &= \sum_T r(S,T)\exp(2\pi i\,\mathrm{tr}(TZ))\end{aligned}$$

where T runs over half integral positive semidefinite symmetric matrices with integral diagonal.

Proposition 7. *Let Λ as above have even rank and Gram matrix S, let M be an integer such that MS^{-1} is integral with even diagonal and write*

$$\chi\begin{pmatrix} A & B \\ C & D\end{pmatrix} = \left(\frac{(-1)^{\frac{m}{2}}\det S}{\det D}\right) \quad \textit{(generalized Jacobi symbol)}$$

for

$$\begin{pmatrix} A & B \\ C & D\end{pmatrix} \in \Gamma_0^{(n)}(M) = \{g = \begin{pmatrix} A & B \\ C & D\end{pmatrix} \in \mathrm{Sp}_n(\mathbb{Z}) \mid C \equiv 0 \bmod M\}.$$

Then $\vartheta^{(n)}(\Lambda, \cdot)$ is a Siegel modular form of weight $k = \frac{m}{2}$ with character χ for the group $\Gamma_0^{(n)}(M)$.

In particular, if $\det S = 1$ (Λ and S are then called even unimodular), $\vartheta^{(n)}(\Lambda, \cdot)$ is a Siegel modular form for the full modular group $\mathrm{Sp}_n(\mathbb{Z})$.

Proof. A proof can e.g. be found in [1], where also a similar formula is given for the case of odd rank m. □

For a Siegel modular form F of degree n for some congruence subgroup Γ the ϕ-operator is defined by

$$(F|\phi)(Z) = (\phi F)(Z) = \lim_{t\to\infty} F\begin{pmatrix} Z & 0 \\ 0 & it\end{pmatrix} \quad (Z \in \mathcal{H}_{n-1}),$$

the function $F|\phi$ is then a Siegel modular form of degree $n-1$ for the group

$$\{\gamma = \begin{pmatrix} A & B \\ C & D\end{pmatrix} \mid \begin{pmatrix} \begin{pmatrix} A & 0 \\ 0 & 1\end{pmatrix} & \begin{pmatrix} B & 0 \\ 0 & 0\end{pmatrix} \\ \begin{pmatrix} C & 0 \\ 0 & 0\end{pmatrix} & \begin{pmatrix} D & 0 \\ 0 & 1\end{pmatrix}\end{pmatrix} \in \Gamma\}.$$

F is a cusp form if $F|\gamma|\phi = 0$ for all $\gamma \in \mathrm{Sp}_n(\mathbb{Z})$, it is said to vanish in all zero-dimensional cusps if

$$(F|\gamma) \mid \Phi^{n-1} = 0 \quad \text{for all } \gamma \in \mathrm{Sp}_n(\mathbb{Z})$$

(i.e., if the constant term in the Fourier expansion of $F|\gamma$ vanishes for all $\gamma \in \mathrm{Sp}_n(\mathbb{Z})$). A well-known fact is:

Proposition 8. *If Λ, Λ' are lattices in the same genus, the function $\vartheta^{(n)}(\Lambda, \cdot) - \vartheta^{(n)}(\Lambda', \cdot)$ vanishes in all zero dimensional cusps. If we define*

$$\vartheta^{(n)}(\mathrm{gen}(\Lambda), Z) = \sum_T r(\mathrm{gen}\,\Lambda, N_T) \exp(\pi i \,\mathrm{tr}(TZ))$$

with N_T denoting a lattice with Gram matrix T, then also

$$\vartheta^{(n)}(\Lambda, \cdot) - \vartheta^{(n)}(\mathrm{gen}(\Lambda), \cdot)$$

vanishes in all zero dimensional cusps.

Proof. In the case $n = 1$ this has been noticed in [46, p. 376], the general case is an immediate consequence. The reason is that $\mathrm{Sp}_n(\mathbb{Z})$ is generated by $\begin{pmatrix} 0_n & -1_n \\ 1_n & 0_n \end{pmatrix}$ and matrices not changing the Fourier expansion of $\vartheta^{(n)}(\Lambda, \cdot)$ and that $F| \begin{pmatrix} 0_n & -1_n \\ 1_n & 0_n \end{pmatrix}$ can be expressed with the help of the Poisson summation formula as a sum of terms whose constant term in the Fourier expansion depends only on the congruence properties of S. □

The analytic approach to Problem 1' can now be formulated as follows.

Write $r(\Lambda, N) = r(\mathrm{gen}\,\Lambda, N) + (r(\Lambda, N) - r(\mathrm{gen}\,\Lambda, N))$ and try to estimate the main term $r(\mathrm{gen}\,\Lambda, N)$ from below and the error term $r(\Lambda, N) - r(\mathrm{gen}\,\Lambda, N)$ from above, using the fact that the latter expression is a Fourier coefficient of a Siegel modular form which vanishes in all zero dimensional cusps. In the case $n = 1$ we have to estimate the Fourier coefficients of a cusp form, which allows to use Deligne's theorem (i.e., the truth of the Ramanujan-Petersson conjecture) if m is even. For $n > 1$, the difference $\vartheta(\Lambda, \cdot) - \vartheta(\mathrm{gen}\,\Lambda, \cdot)$ will in general not be a cusp form; this makes the estimation of the error term from above considerably more difficult.

The first result for our problem in the case $n > 1$ is due to Raghavan:

Theorem 9 ([39]). *Let N run through positive definite integral lattices of rank n with $2n + 3 \leq m$ and $(\det N) \to \infty$ satisfying one of the equivalent conditions*

(a) $\min(N) \cdot \min(N^{\#}) < c_1$ *for some fixed* $c_1 > 0$
(b) $\min(N^{\#}) \geq c_1(\det N)^{-\frac{1}{n}}$ *for some fixed* $c_2 > 0$
(c) $\min(N) \geq c_3(\det N)^{\frac{1}{n}}$ *for some fixed* $c_3 > 0$

Then one has

$$r(\Lambda, N) = r(\operatorname{gen}\Lambda, N) + \mathrm{O}((\det N)^{\frac{m-n-1}{2}} \cdot (\min(N))^{(n+1-\frac{m}{2})/2}).$$

Proof. The idea of the proof is to compute the Fourier coefficient at T of $g(Z) := \vartheta^{(n)}(\Lambda, Z) - \vartheta^{(n)}(\operatorname{gen}\Lambda, Z)$ as

$$\int_{\mathfrak{C}} g(Z)\exp(-2\pi i \operatorname{tr}(TZ))dX,$$

where the variable $Z = X + iT^{-1}$ runs over a cube $\mathfrak{C}$ of side length 1 with one corner in T^{-1}, using a generalized Farey dissection of this cube which has been introduced by Siegel in [49].

Raghavan proves in fact more generally an estimate for the Fourier coefficient of $F - \varphi$, where F is a Siegel modular form of weight $k > n+1$ and φ the associated Eisenstein series; the analytic version of Siegel's theorem states that $\vartheta^{(n)}(\operatorname{gen}(\Lambda), \cdot)$ is the Eisenstein series associated to $\vartheta^{(n)}(\Lambda, \cdot)$.

Raghavan shows that the formula given above is indeed an asymptotic formula for the number of representations of N by Λ in the case $n = 2$; this is achieved by estimating the local densities (and hence the main term $r(\operatorname{gen}(\Lambda), N)$) from below; we will see a more general version of that below. □

Remark. Minkowski's reduction theory of positive definite quadratic forms implies that the minimum of N grows at most like a constant multiple of $(\det(N))^{\frac{1}{n}}$, so the condition of the theorem roughly says that the minimum of N grows as fast as it can.

Since Λ represents (by the case $n = 1$) only all sufficiently large numbers that it represents locally everywhere, there are in general some small numbers which are represented by Λ locally everywhere but not globally. It is then easy to construct a sequence of lattices N of growing determinant and one of these exceptional numbers as minimum which are represented locally everywhere but not globally (since their minimum is not represented globally). The most common example of this type is to take the Leech lattice for Λ and the number 1 as minimum of N. It is therefore clear that an asymptotic formula for $r(\Lambda, N)$ can not rely on the growth of $\det(N)$ alone.

Raghavan's result was further improved by Kitaoka in [24, 25].

Theorem 10 ([24]).

(a) *If $m \geq 2n+3$ holds, the product $\prod_p \alpha_p(\Lambda, N)$ is bounded from below and above by constants depending only on Λ for all N which are represented locally everywhere by Λ.*

In particular, one has

$$r(\operatorname{gen}\Lambda, N) > c_4(\det N)^{\frac{m-n-1}{2}}$$

for all N which are represented locally everywhere by Λ , with c_4 depending only on Λ.

(b) *The formula in Raghavan's theorem is an asymptotic formula for N which are represented locally everywhere by Λ and for which $\min(N) \geq c_3(\det N)^{\frac{1}{n}}$ with some constant $c_3 > 0$ independent of N.*
If N runs through lattices of rank n with $\min(N) > c_3(\det N)^{\frac{1}{n}}$, any N of sufficiently large minimum which is represented locally everywhere by Λ is represented globally by Λ.

(c) *If $n = 2$, $m = 7$, there is a constant c_5 depending only on Λ such that the condition $\min(N) \geq c_3(\det N)^{\frac{1}{2}}$ above can be replaced by $\min(N) \geq c_5$.*
In particular, any N which is represented locally everywhere by Λ and has minimum $\geq c_5$ is represented globally by Λ.

The method of proof here is essentially the same as for Raghavan's theorem, but considerably refined.

Another interesting result of Kitaoka is proved in [26].

Theorem 11 ([27]). *Let $n = 2$, $m = 6$. Let N_0 have Gram matrix T_0 and be represented by Λ locally everywhere. Then for $t \to \infty$ with $\gcd(t, \det(\Lambda)) = 1$ and such that tT_0 is represented by Λ_p for all p dividing $\det(\Lambda)$ one has*

$$r(\Lambda, tT_0) = r(\mathrm{gen}(\Lambda), tT_0) + O(t^{\frac{5}{2}+\delta}) \quad \textit{for all } \delta > 0,$$

where the main term $r(\mathrm{gen}(\Lambda), tT_0)$ grows at least like $t^{3-\epsilon}$ for all $\epsilon > 0$.

In particular, for t large enough and satisfying the conditions above the matrix tT_0 is represented by Λ.

For the case that Λ is even unimodular (i.e., the Gram matrix S of Λ has even diagonal and determinant 1), the estimation of the error term has also been investigated by Kitaoka using instead of the circle method as above the decomposition of $\vartheta^{(n)}(\Lambda, \cdot) - \vartheta^{(n)}(\mathrm{gen}\,\Lambda, \cdot)$ into a cusp form and Klingen-Eisenstein series associated to cusp forms of degree $r < n$. The result is a similar asymptotic formula as above for the range $m \geq 4n + 4$, namely

$$r(\Lambda, N) = r(\mathrm{gen}\,\Lambda, N) + \mathrm{O}((\min(N))^{1-\frac{m}{4}})(\det N)^{\frac{m-n-1}{2}}.$$

In particular, in exchange for the restriction on Λ and the stronger condition $m \geq 4n + 4$ one gets rid of the condition $\min(N) > c_3(\det N)^{\frac{1}{n}}$.

In all of the above results one can deduce global representability from local representability only for lattices N of large minimum, a condition which excludes many cases of interest in which the determinant of N grows but the minimum remains small.

The examples which show its necessity are lattices N of large determinant which have small minimum or more generally a sublattice N' of small determinant, so that it can happen that N' is not represented globally by Λ. On the other hand, a lattice N' of rank $n' < n$ and small determinant of which one already knows that it is represented globally by Λ may have extensions to lattices N of rank n which are represented locally everywhere. A result towards the global representation of such N by Λ has been obtained in [8].

Theorem 12 ([8]). *Let T_1 be positive definite symmetric Minkowski reduced of rank n_1 and Λ be even unimodular. Then for Minkowski reduced symmetric $n \times n$-matrices $T = \begin{pmatrix} T_1 & T_2 \\ {}^tT_2 & T_4 \end{pmatrix}$ with $m > 4n$ and sufficiently large $\min(T_4)$ the primitive representation number $r^*(\Lambda, T)$ satisfies*

$$r^*(\Lambda,T) = c_6 r^*(\Lambda,T_1)\cdot(\det T_4)^{\frac{m-n-1}{2}} + O((\det T_4)^{\frac{m-n-1}{2}})\min(T_4)^{-\frac{m}{4}+v(n_1)}$$

for some constant $c_6 \neq 0$, with $v(n_1) < \frac{m}{4}$.

The proof uses again the decomposition of $\vartheta^{(n)}(\Lambda) - \vartheta^{(n)}(\operatorname{gen}\Lambda)$ into a sum of Klingen-Eisenstein series associated to cusp forms of degrees $\leq n$. Notice that for even unimodular Λ and $m \geq 2n+3$ the condition of local representability is satisfied automatically.

In the case $n_1 = 1$ Böcherer has shown in [6] that for a square free integer $t_1 = T_1$ this problem can also be treated using the theory of Jacobi forms; a generalization of that result to general t_1 and to not necessarily unimodular Λ will be the subject of the PhD thesis of T. Paul in Saarbrücken.

7 Representation of Forms, Arithmetic Method

In order to present the arithmetic method we need some terminology. We denote by $O_V(F)$ the group of isometries of V with respect to Q (the orthogonal group of the quadratic space (V,Q)), by $O_V(\mathbb{A})$ its adelization, and by $SO_V(F)$ resp. $SO_V(\mathbb{A})$ their subgroups of elements of determinant 1. For a lattice Λ on V we denote its automorphism group (or unit group) $\{\sigma \in O_V(F) \mid \sigma(\Lambda) = \Lambda\}$ by $O_\Lambda(R)$ and similarly for the local or adelic analogues. $\mathrm{Spin}_V(\mathbb{A})$ is the adelic spin group and $O'_V(\mathbb{A})$ its image in $O_V(\mathbb{A})$, i.e., the subgroup of adelic transformations of determinant and spinor norm 1.

The orbit of a fixed lattice Λ under $O_V(\mathbb{A})$ consists then of all lattices on V in the genus of Λ, the lattices on V in the spinor genus of Λ comprise the orbit under $O_V(F)O'_V(\mathbb{A})$.

The proof of the theorem of Sect. 4 on representation by indefinite lattices rests on the strong approximation theorem for the spin group with respect to an archimedean place of F at which Q is indefinite.

In the case of a definite lattice one can, following Eichler [15], consider it as "arithmetically indefinite" if there is a non-archimedean place w of F for which Λ_w is isotropic (i.e., represents zero nontrivially). The strong approximation theorem gives then

Lemma 13. *Let w be a non-archimedean place of F for which Λ_w is isotropic.*

(a) *Each class in the spinor genus of Λ has a representative Λ' such that $\Lambda'_v = \Lambda_v$ for all places $v \neq w$.*
(b) *If the genus of Λ consists of only one spinor genus there is an integer s such that Λ represents every R-lattice N for which N_v is represented by $\mathfrak{p}_w^s \Lambda_v$ for all finite places v of F (where $\mathfrak{p}_w$ is the ideal of R corresponding to w).*
(c) *If $m \geq n+3$ and N is represented (primitively) locally everywhere by Λ there is a lattice Λ' in the spinor genus of Λ with $\Lambda'_v = \Lambda_v$ for all places $v \neq w$ and Λ'_w in the $\operatorname{Spin}(F_w)$-orbit of Λ_w, such that N is represented (primitively) by Λ.*

The lemma alone is not sufficient to deduce global representability of N by Λ from representability locally everywhere since in the definite situation the spinor genus consists in general of more than one class. We will see in the next section that it provides the starting point for the ergodic method of Ellenberg and Venkatesh. It is also basic for the purely arithmetic method of Hsia, Kitaoka and Kneser.

For N which is represented by Λ locally everywhere they construct in [20], using the local arithmetic of lattices, a finite set of sublattices $K(J)$ of rank n and $L(J)$ of rank $m-n \geq n+3$ of Λ which are orthogonal to each other and such that for each finite place v of F the lattice N_v is represented either by $K(J)_v$ or by $\mathfrak{p}_v^s L(J)_v$. With the help of (b) of the Lemma and some additional rather tricky approximation arguments they can then deduce that N is represented by one of the $K(J)+L(J)$ and hence by Λ if the minimum of N is large enough. The final result is

Theorem 14 ([20]). *There is a constant $c_7 = c_7(\Lambda)$ such that for $m \geq 2n+3$ every lattice N which is represented locally everywhere by Λ and has minimum $\geq c_7$ is represented by Λ.*

The constant in the theorem can in principle be made effective; such an effective version (with a rather large constant) has been given by Chan and Icaza in [10] for $m \geq 3n+3$ and for $n=2, m=7$. It has been shown by Jöchner and Kitaoka in [22] that the proof of the theorem can be modified to give the same result for representations with additional congruence and primitivity conditions and by Hsia and Prieto-Cox in [21] that it can also be generalized to hermitian forms.

Kitaoka has further noticed (see [8, p. 95]) that a version of the result on extensions of representations in Theorem 12 can also be obtained by the arithmetic method, the result given in [8] has been further improved by Chan, B. M. Kim, M.-H. Kim, and Oh in [11] to give

Theorem 15 ([11]). *Let $F = \mathbb{Q}$, let K be a lattice of rank k on the space of Λ, let $\sigma : K \longrightarrow \Lambda$ be a representation. Then there is a constant $c_8 > 0$ such that one has:*

If $N \supseteq K$ is a lattice of rank n with $m \geq k + 2(n-k) + 3$ on the space of Λ such that for all primes p the local representation $\sigma_p : K_p \longrightarrow \Lambda_p$ can be extended to a representation $\rho_p : N_p \longrightarrow \Lambda_p$ and such that the minimum of the orthogonal projection $\pi(N)$ on the orthogonal complement of $\mathbb{Q}K$ in $\mathbb{Q}\Lambda$ is larger than c_8, the representation σ can be extended to a representation $\rho : N \longrightarrow \Lambda$. One can in addition specify congruence conditions modulo an integer prime to $2\det(K)\det(\Lambda)$.

8 Representation of Forms, Ergodic Method

The result of (c) of Lemma 13 can be rephrased group theoretically:

There exists an isometric embedding of N into $\Lambda' = u\Lambda$ with

$$u \in O_V(F)(\prod_{v \neq w} O_\Lambda(R_v))\mathrm{Spin}_V(F_w).$$

Representability of N by Λ is equivalent to being able to choose $u \in O_V(F)O_\Lambda(\mathbb{A})$ instead with $O_\Lambda(\mathbb{A}) = \prod_v O_\Lambda(R_v)$.

If we consider N as a sublattice of Λ' we can clearly modify u by a suitable element of $O_{W_1}(F_w)$, where $W_1 = (FN)^\perp$.

Ellenberg and Venkatesh show in [16] that the necessary modification of u is indeed possible for N of sufficiently large minimum if one has $m \geq n+3$, the lattice N has square free determinant, and N satisfies some additional conditions; their proof uses ergodic theory, in particular results of Ratner and Margulis/Tomanov (see [37, 40]). In view of the fact that before their work it was generally considered to be possible that $m = 2n + 2$ is the natural barrier for the validity of a representability result this represented a dramatic breakthrough. That their conditions on the lattice N (but not on its dimension) can be further relaxed has been shown in [45], where also the arithmetic parts of their proof were reformulated in a way closer to previous work in the arithmetic theory of quadratic forms.

The final result is:

Theorem 16. *Let $(V, Q), \Lambda$ be as before, fix a finite place w of F and $j \in \mathbb{N}, a \in R$.*

Then there exists a constant $c_9 := c_9(\Lambda, j, w, a)$ such that Λ represents all R-lattices N of rank $n \leq m - 3$ satisfying

(a) *N is represented by Λ locally everywhere with imprimitivity bounded by a and with isotropic orthogonal complement at the place w.*
(b) $\mathrm{ord}_w(\det(N_w)) \leq j$
(c) *The minimum of N is $\geq c_9$.*

The representation may be taken to be of imprimitivity bounded by a.

The isotropy condition is satisfied automatically if $n \leq m - 5$ or if w is such that the determinants of the local lattices Λ_w and N_w are units in R_w.

It is not difficult to adapt the method in order to obtain a version for extensions of representations:

Corollary 17. *Let $(V, Q), \Lambda$ be as before, fix a finite place w of F and $j \in \mathbb{N}$, $a \in R$.*

Let $K \subseteq \Lambda$ be a fixed R-lattice of rank k, $\sigma : K \longrightarrow \Lambda$ a representation of K by Λ and assume that K_w is unimodular.

Then there exists a constant $c_{10} := c_{10}(\Lambda, R, j, w, a)$ such that one has: If $N \supseteq K$ is an R-lattice of rank $n \leq m - 3$ and

(a) *For each place v of F there is a representation $\tau_v : N_v \longrightarrow \Lambda_v$ with $\tau_v|_{K_v} = \sigma_v$ with imprimitivity bounded by a and with isotropic orthogonal complement in Λ at the place w*
(b) *For the w-adic order* $\mathrm{ord}_w(\det(N_w))$ *of the determinant of a Gram matrix of N_w one has* $\mathrm{ord}_w(\det(N_w)) \leq j$
(c) *The minimum of $N \cap (FK)^\perp$ is $\geq c_{10}$,*

then there exists a representation $\tau : N \longrightarrow \Lambda$ with $\tau|_K = \sigma$.

The representation may be taken to be of imprimitivity bounded by a.

The isotropy condition is satisfied automatically if $n \leq m - 5$ or if w is such that the local lattices Λ_w and N_w are unimodular.

Ellenberg and Venkatesh prove the theorem in [16] under the stronger restriction that the determinant of N is square free; the version of it given here and the corollary are proven in [45].

For the reader's convenience we add a matrix version of the main result for the case $F = \mathbb{Q}$:

Theorem 18. *Let $S \in M_m^{\mathrm{sym}}(\mathbb{Z})$ be a positive definite integral symmetric $m \times m$-matrix, fix a prime q and positive integers j, a.*

Then there is a constant c_{11} such that a positive definite matrix $T \in M_n^{\mathrm{sym}}(\mathbb{Z})$ with $n \leq m - 3$ is represented by S (i.e., $T =^t XSX$ with $X \in M_{mn}(\mathbb{Z})$) provided it satisfies:

(a) *For each prime p there exists a matrix $X_p \in M_{mn}(\mathbb{Z}_p)$ with ${}^tX_pSX_p = T$ such that the elementary divisors of X_p divide a and such that the equations ${}^tX_qS\mathbf{y} = \mathbf{0}$ and ${}^t\mathbf{y}S\mathbf{y} = 0$ have a nontrivial common solution $\mathbf{y} \in \mathbb{Z}_q^m$*
(b) $q^j \nmid \det(T)$
(c) $\min\{{}^t\mathbf{y}T\mathbf{y} \mid \mathbf{0} \neq \mathbf{y} \in \mathbb{Z}^n\} > c_{11}$

The matrix X may be chosen to have elementary divisors dividing a.

As remarked earlier the primitivity (or bounded imprimitivity) condition is satisfied automatically in the range $m \geq 2n + 3$ covered by the analytic and arithmetic results. Kitaoka has proved in [27, 28, 30, 31] some lemmas which imply that one can drop or weaken the primitivity conditions in some lower dimensional cases; the original purpose of those lemmas was to obtain improved estimates for the main term in the analytic method which could be used once the analytic estimate for the error term in the asymptotic formula could be improved sufficiently much. This leads to the following corollaries, proven in [45]:

Corollary 19. *Let $F = \mathbb{Q}$, let $(V, Q), \Lambda$ be as before and fix a prime q and $j \in \mathbb{N}$.*

(a) *Let $n \geq 6$ and $m = \dim(V) \geq 2n$. Then there exists a constant $c_{12} := c_{12}(\Lambda, j, q)$ such that Λ represents all $\mathbb{Z}$ - lattices N of rank n which are represented by Λ locally everywhere, have minimum $\geq c_{12}$ and satisfy* $\mathrm{ord}_q(\det(N)) \leq j$.

(b) *Let* $n \geq 3$ *and* $m = \dim(V) \geq 2n+1$. *Then there exists a constant* $c_{13} := c_{13}(\Lambda, j, q)$ *such that* Λ *represents all* $\mathbb{Z}$ *- lattices* N *of rank* n *which are represented by* Λ *locally everywhere, have minimum* $\geq c_{13}$*, satisfy* $\mathrm{ord}_q(\det(N)) \leq j$ *and which are in the case* $n = 3$ *such that the orthogonal complement of* N_q *in* Λ_q *is isotropic.*

(c) *Let* $n = 2$ *and* $m = \dim(V) \geq 6$. *Then there exists a constant* $c_{14} := c_{14}(\Lambda, j, q)$ *such that* Λ *represents all* $\mathbb{Z}$ *- lattices* N *of rank* n *which are represented by* Λ *locally everywhere, have minimum* $\geq c_{14}$*, satisfy* $\mathrm{ord}_q(\det(N)) \leq j$ *and which are such that the orthogonal complement of* N_q *in* Λ_q *is isotropic.*

(d) *Let a positive definite* $\mathbb{Z}$*-lattice* N_0 *of rank* $n_0 \leq m-3$ *with Gram matrix* T_0 *be given. Let* S *be a finite set of primes with* $q \in S$ *such that one has*

(i) Λ_p *and* $(N_0)_p$ *are unimodular for all primes* $p \notin S$ *and for* $p = q$.

(ii) Each isometry class in the genus of Λ *has a representative* Λ' *on* V *such that* $\Lambda'_p = \Lambda_p$ *for all primes* $p \notin S$.

Then there exists a constant $c_{15} := c_{15}(\Lambda, T_0, S)$ *such that for all sufficiently large integers* $t \in \mathbb{Z}$ *which are not divisible by a prime in* S*, the* $\mathbb{Z}$*-lattice* N *with Gram matrix* tT_0 *is represented by* Λ *if it is represented by all completions* Λ_p.

9 Comparison of Results

Concerning dimension bounds the theorem of Ellenberg and Venkatesh is clearly superior to the results obtained by other methods, and it should not be difficult to show that it is best possible in this respect. The method makes it necessary to impose a bound on the power to which some fixed prime is allowed to divide the determinant of the lattice N; this is not necessary for the arithmetic and the analytic results in the dimension range where they are valid. At least at present the ergodic method gives neither an effective bound on the "sufficiently large" minimum of the lattice N nor an asymptotic formula for the number of representations. This may of course change with further refinements of the results from ergodic theory which make the proof possible.

Results of Kitaoka (see [25, 27, 29]) on estimates of local densities show that at least the main term $r(\mathrm{gen}\Lambda, N)$ is still growing like $(\det\ N)^{\frac{m-n-1}{2}}$ in the range $n+3 < m \leq 2n+2$ if one puts suitable restrictions on N, e.g., if one supposes a Gram matrix of N to have square free determinant.

On the other hand, even for a Siegel cusp form of weight k the best known estimtes for the Fourier coefficient $a(F,T)$ at T bound it by a term of the type $(\det\ T)^r$ where r is not much smaller than $\frac{k}{2}$, see [7] for some results in that direction. The famous conjecture of Resnikoff and Saldaña [41] (for which meanwhile counterexamples are known, see [36]) predicts an estimate

$$|a(F,T)| = O((\det(T))^{\frac{k}{2}-\frac{n+1}{4}+\epsilon}),$$

hence (with $m = 2k$) an exponent $\frac{m-n-1}{4} + \epsilon$ at $\det(T)$, which, like the exponent in the main term, depends only on the difference $m - n$ but not on m itself.

An asymptotic formula for $r(\Lambda, N)$ valid in a range $m \geq n + n_0$ for some fixed n_0 would have a main term growing like $(\det(T))^{\frac{n_0-1}{2}}$ (with some restrictions on T), in particular the exponent would be independent of the weight of the theta series. Its validity would therefore in particular require that the Fourier coefficients of the modular form $\vartheta^{(n)}(\Lambda) - \vartheta^{(n)}(\mathrm{gen}(\Lambda))$, which in general is not cuspidal, satisfy an estimate similar to that of the Resnikoff-Saldaña conjecture for Fourier coefficients of cusp forms.

References

1. A. N. Andrianov, V. G. Zhuravlev: Modular forms and Hecke operators. Translations of Mathematical Monographs, **145**. American Mathematical Society, Providence, RI, 1995.
2. M. Bhargava: On the Conway-Schneeberger fifteen theorem. Quadratic forms and their applications (Dublin, 1999), 27–37, Contemp. Math., **272**, Amer. Math. Soc., Providence, RI, 2000.
3. V. Blomer, V, J. Brüdern, R. Dietmann: Sums of smooth squares. Compos. Math. **145** (2009), 1401–1441.
4. V. Blomer, G. Harcos: Hybrid bounds for twisted L-functions. J. Reine Angew. Math. **621** (2008), 53–79
5. V. Blomer, G. Harcos: Twisted L-functions over number fields and Hilbert's eleventh problem, Geom. Funct. Anal. **20** (2010), 1–52
6. S. Böcherer: On the Fourier-Jacobi-coefficients of Eisenstein series of Klingen type, Preprint
7. S. Böcherer, W. Kohnen: Estimates for Fourier coefficients of Siegel cusp forms. Math. Ann. **297** (1993), 499–517
8. S. Böcherer, S. Raghavan: On Fourier coefficients of Siegel modular forms. J. Reine Angew. Math. **384** (1988), 80–101.
9. J.W.S. Cassels: Rational quadratic forms. London Mathematical Society Monographs, 13. Academic Press, Inc., London-New York, 1978.
10. W.K. Chan, M.I. Icaza: Effective results on representations of quadratic forms. Algebraic and arithmetic theory of quadratic forms, 73–83, Contemp. Math., **344**, Amer. Math. Soc., Providence, RI, 2004
11. W.K. Chan, B.M. Kim, M.-H. Kim, B.-K. Oh: Extensions of representations of integral quadratic forms, Ramanujan J. **17** (2008), 145–153
12. W. Duke, R. Schulze-Pillot: Representation of integers by positive ternary quadratic forms and equidistribution of lattice points on ellipsoids. Invent. Math. **99** (1990), 49–57.
13. A.G. Earnest, J.S. Hsia: Spinor norms of local integral rotations. II. Pacific J. Math. **61** (1975), 71–86.
14. A.G. Earnest, J.S. Hsia, D.C. Hung: Primitive representations by spinor genera of ternary quadratic forms. J. London Math. Soc. (2) **50** (1994), 222–230
15. M. Eichler: Die Ähnlichkeitsklassen indefiniter Gitter. Math. Z. **55** (1952), 216–252.
16. J. Ellenberg, A. Venkatesh: Local-global principles for representations of quadratic forms. Invent. Math. **171** (2008), 257–279.
17. H. Hasse: Über die Darstellbarkeit von Zahlen durch quadratische Formen im Körper der rationalen Zahlen. J. für Math. **152**, 129–148 (1923)
18. J.S. Hsia: Spinor norms of local integral rotations. I. Pacific J. Math. **57** (1975), 199–206
19. J. S. Hsia: Representations by spinor genera. Pacific J. Math. **63** (1976), 147–152.

20. J. S. Hsia, Y. Kitaoka, M. Kneser: Representations of positive definite quadratic forms. J. Reine Angew. Math. **301** (1978), 132–141.
21. J.S. Hsia, J.P. Prieto-Cox: Representations of positive definite Hermitian forms with approximation and primitive properties. J. Number Theory **47** (1994), 175–189.
22. M. Jöchner, Y. Kitaoka: Representations of positive definite quadratic forms with congruence and primitive conditions. J. Number Theory **48** (1994), 88–101.
23. B.M. Kim, M.-H. Kim, B.-K. Oh: A finiteness theorem for representability of quadratic forms by forms. J. Reine Angew. Math. **581** (2005), 23–30
24. Y. Kitaoka: Modular forms of degree n and representation by quadratic forms. II. Nagoya Math. J. **87** (1982), 127–146.
25. Y. Kitaoka: Lectures on Siegel modular forms and representation by quadratic forms. Tata Institute of Fundamental Research Lectures on Mathematics and Physics, **77**. Published for the Tata Institute of Fundamental Research, Bombay; by Springer-Verlag, Berlin, 1986.
26. Y. Kitaoka: Modular forms of degree n and representation by quadratic forms. V. Nagoya Math. J. **111** (1988), 173–179.
27. Y. Kitaoka: Local densities of quadratic forms. Investigations in number theory, 433–460, Adv. Stud. Pure Math., **13**, Academic Press, Boston, MA, 1988.
28. Y. Kitaoka: Some remarks on representations of positive definite quadratic forms. Nagoya Math. J. **115** (1989), 23–41.
29. Y. Kitaoka: Representations of positive definite quadratic forms and analytic number theory. Translation of Sugaku **43** (1991), no. 2, 115–127
30. Y. Kitaoka: The minimum and the primitive representation of positive definite quadratic forms. Nagoya Math. J. **133** (1994), 127–153.
31. Y. Kitaoka: The minimum and the primitive representation of positive definite quadratic forms. II. Nagoya Math. J. **141** (1996), 1–27.
32. H.D. Kloosterman: On the representation of numbers in the form $ax^2 + by^2 + cz^2 + dt^2$., Acta Math. **49** (1927), 407–464.
33. H.D. Kloosterman: Asymptotische Formeln für die Fourierkoeffizienten ganzer Modulformen. (German) Abh. Math. Sem. Hamburg **5** (1927), 337–352.
34. M. Kneser: Klassenzahlen indefiniter quadratischer Formen in drei oder mehr Veränderlichen. Arch. Math. **7** (1956), 323–332.
35. M. Kneser: Darstellungsmaße indefiniter quadratischer Formen. Math. Z. **77** (1961), 188–194.
36. W. Kohnen: On the fourier coefficients of certain special Siegel cusp forms. Math. Z. **248** (2004), 345–350
37. G.A. Margulis, G.M. Tomanov: Invariant measures for actions of unipotent groups over local fields on homogeneous spaces. Invent. Math. **116** (1994), 347–392.
38. O.T. O'Meara: Introduction to quadratic forms, Springer-Verlag 1973.
39. S. Raghavan: Modular forms of degree n and representation by quadratic forms. Ann. of Math. (2) **70** (1959) 446–477
40. M. Ratner: Raghunathan's conjectures for Cartesian products of real and p-adic Lie groups. Duke Math. J. **77** (1995), 275–382
41. H. L. Resnikoff and R. L. Saldaña: Some properties of Fourier coefficients of Eisenstein series of degree two. J. reine angew. Math. **265** (1974) 90–109.
42. R. Schulze-Pillot: Darstellung durch Spinorgeschlechter ternärer quadratischer Formen, J. Number Theory **12** (1980), 529–540
43. R. Schulze-Pillot: Darstellungsmaße von Spinorgeschlechtern ternärer quadratischer Formen, J. Reine Angew. Math. **352** (1984), 114–132
44. R. Schulze-Pillot: Representation by integral quadratic forms—a survey. Algebraic and arithmetic theory of quadratic forms, 303–321, Contemp. Math., **344**, Amer. Math. Soc., Providence, RI, 2004
45. R. Schulze-Pillot: Local conditions for global representations of quadratic forms, Acta Arith. **138** (2009), 289–299
46. C.L. Siegel: Über die analytische Theorie der quadratischen Formen. Ann. of Math. (2) **36** (1935), 527–606

47. C.L. Siegel: Über die analytische Theorie der quadratischen Formen II, Ann. of Math. (2) **37** (1936), 230–263
48. C.L. Siegel: Über die analytische Theorie der quadratischen Formen III, Ann. of Math. (2) **38** (1937), 212–291
49. C. L. Siegel: On the theory of indefinite quadratic forms, Ann. of Math. **45** (1944), 577–622
50. W. Tartakowsky: Die Gesamtheit der Zahlen, die durch eine positive quadratische Form $F(x_1, x_2, \ldots, x_s)$ $(s \geqq 4)$ darstellbar sind. I, II. Bull. Ac. Sc. Leningrad 7 **2** (1929); 111–122, 165–196 (1929)
51. A. Weil: Sur la théorie des formes quadratiques. in: Colloq. Théorie des Groupes Algébriques (Bruxelles, 1962) pp. 9–22
52. Fei Xu: Representation masses of spinor genera. Duke Math. J. **110** (2001), 279–307
53. Fei Xu: On representations of spinor genera II. Math. Ann. **332** (2005), 37–53.

Identifying the Matrix Ring: Algorithms for Quaternion Algebras and Quadratic Forms

John Voight*

Abstract We discuss the relationship between quaternion algebras and quadratic forms with a focus on computational aspects. Our basic motivating problem is to determine if a given algebra of rank 4 over a commutative ring R embeds in the 2×2-matrix ring $\mathrm{M}_2(R)$ and, if so, to compute such an embedding. We discuss many variants of this problem, including algorithmic recognition of quaternion algebras among algebras of rank 4, computation of the Hilbert symbol, and computation of maximal orders.

Key words and Phrases Quadratic forms • Quaternion algebras • Maximal orders • Algorithms • Matrix ring • Number theory

Mathematics Subject Classification (2010): Primary 11R52; Secondary 11E12

Since the discovery of the division ring of quaternions over the real numbers by Hamilton, and continuing with work of Albert and many others, a deep link has been forged between quadratic forms in three and four variables over a field F and quaternion algebras over F. Starting with a *quaternion algebra* over F, a central simple F-algebra of dimension 4, one obtains a quadratic form via the reduced norm (restricted to the trace zero subspace); the split quaternion algebra over F, the 2×2-matrix ring $\mathrm{M}_2(F)$, corresponds to an isotropic quadratic form, one that represents zero nontrivially. (Conversely, one recovers the quaternion algebra via the Clifford algebra of the quadratic form.) In this article, we give an exposition of this

*Research partially supported by NSA Grant H98230-09-1-0037 and NSF Grant DMS-0901971.

J. Voight (✉)
Department of Mathematics and Statistics, University of Vermont, 16 Colchester Ave,
Burlington, VT 05401, USA
e-mail: jvoight@gmail.com

K. Alladi et al. (eds.), *Quadratic and Higher Degree Forms*, Developments
in Mathematics 31, DOI 10.1007/978-1-4614-7488-3_10,

link relating quaternion algebras and quadratic forms from an explicit, algorithmic perspective and in a wider context.

Let R be a noetherian, commutative domain. We say that R is *computable* if there exists an encoding of R into bits with algorithms to perform ring operations in R and to test if an element of R is zero. The following basic algorithmic problem, along with its many variants, forms the core of this article. (See Sect. 1 for further definitions and algorithmic specifications.)

Problem (IsMatrixRing). *Given a computable domain R and an R-algebra $\mathcal{O}$ of rank 4, determine if $\mathcal{O}$ embeds in $\mathrm{M}_2(R)$ and, if so, compute an explicit embedding $\mathcal{O} \hookrightarrow \mathrm{M}_2(R)$ of R-algebras.*

The Problem (IsMatrixRing) captures in an important way the link between quadratic forms and quaternion algebras. In the simplest case where $R = F$ is a field—when such an embedding is necessarily an isomorphism—this problem corresponds to asking if a ternary quadratic form over F represents zero nontrivially, and for this reason it arises in a wide variety of situations. When F is a local field, this problem corresponds to the computation of the Hilbert symbol. In the case where R is a local ring, it corresponds to the computation of an (explicit) integral splitting of a quaternion order and thereby appears as a foundational step in many algorithms in arithmetic geometry (as in work of Kirschmer and the author [18]). Finally, when R is a Dedekind domain, roughly speaking, the problem of approximating (IsMatrixRing) naturally gives rise to the problem of computing a maximal order containing $\mathcal{O}$. In these and other ways, therefore, the Problem (IsMatrixRing) will serve as kind of unifying and motivating question.

In Sect. 1, we introduce the basic terminology we will use throughout concerning computable rings and quaternion algebras. In Sect. 2, we consider algebras equipped with a standard involution and we exhibit an algorithm to test if an F-algebra B has a standard involution. In Sect. 3, we relate algebras with a standard involution to quadratic forms via the reduced norm; we introduce the theory of quadratic forms over local PIDs, providing an algorithm to compute a *normalization* of such a form. As a consequence, we exhibit an algorithm to test if an F-algebra B is a quaternion algebra and, if so, to compute *standard generators* for B. With these reductions, we turn in Sect. 4 to Problem (IsMatrixRing) for quaternion algebras and prove that this problem is deterministic polynomial-time equivalent to the problem of determining if a conic defined over F has an F-rational point (and, if so, to exhibit one).

In Sect. 5, we consider Problem (IsMatrixRing) in the case where F is a local field, which corresponds to the computation of the Hilbert symbol; in Sect. 6 we treat the more delicate case of a local dyadic field, and putting these together prove that there is a deterministic polynomial-time algorithm to compute the Hilbert symbol (Theorem 6.1). We thereby exhibit an algorithm to compute the generalized Jacobi symbol for computable Euclidean domains. In Sect. 7, we turn to the case of a Dedekind domain R and relate Problem (IsMatrixRing) to the problem of computing a maximal R-order; we prove that the problem of computing a maximal order for a quaternion algebra B over a number field F is probabilistic polynomial-time equivalent to the problem of factoring integers. Finally, in Sect. 8, we consider the Problem (IsMatrixRing) over $\mathbb{Q}$, and show that recognizing the matrix ring is deterministic polynomial-time equivalent to the problem of quadratic residuosity.

Many of the results in this paper fit into the more general setting of semisimple algebras; however, we believe that the special link to quadratic forms, along with the wide application of quaternion algebras (analogous to that of quadratic field extensions), justifies the specialized treatment they are afforded here.

The author would like to thank his Ph.D. advisor Hendrik Lenstra for his many helpful comments, the Magma group at the University of Sydney for their support while writing this paper, and David Kohel for his valuable input. We are indebted to Carl Pomerance for the citation [2] and would like to thank Asher Auel, Jonathan Hanke, Kate Thompson, and the referee for helpful corrections and suggestions. Some of the results herein occur in the author's Ph.D. thesis [39].

1 Rings and Algebras

We begin by introducing some notation and background that will be used throughout. Let R be a commutative, noetherian domain (with 1), and let F be the field of fractions of R.

Let $\mathcal{O}$ be an *R-algebra*, an associative ring with 1 equipped with an embedding $R \hookrightarrow \mathcal{O}$ of rings (taking $1 \in R$ to $1 \in \mathcal{O}$) whose image lies in the center of $\mathcal{O}$; we identify R with its image under this embedding. We will assume without further mention that $\mathcal{O}$ is a finitely generated, projective (equivalently, locally free) R-module of rank $n \in \mathbb{Z}_{\geq 1}$.

Computable rings and algebras. We will follow the conventions of Lenstra [22] for rings and algorithms, with the notable exception that we do not require all rings to be commutative.

A domain R is *computable* if R comes equipped with a way of encoding elements of R in bits (i.e. the elements of R are recursively enumerable, allowing repetitions) along with deterministic algorithms to perform ring operations in R (addition, subtraction, and multiplication) and to test if $x = 0 \in R$; a ring is *polynomial-time computable* if these algorithms run in polynomial time (in the bit size of the input). A field is *computable* if it is a computable ring and furthermore there exists an algorithm to divide by a nonzero element. For precise definitions and a thorough survey of the subject of computable rings we refer to Stoltenberg-Hansen and Tucker [34] and the references contained therein.

Example 1.1. A domain R which is the localization of a ring which is finitely generated over its prime ring is computable by the theory of Gröbner bases [13]. For example, any finitely generated algebra over $\mathbb{Z}$ or $\mathbb{Q}$ (without zerodivisors, since we restrict to domains) is computable, and in particular the coordinate ring of any integral affine variety over a finitely generated field is computable.

Example 1.2. If R is a computable domain, then F is a computable field if elements are represented in bits as pairs of elements of R in the usual way.

Remark 1.3. Inexact fields (e.g. local fields, such as $\mathbb{Q}_p$ or $\mathbb{R}$) are not computable, since they are uncountable! However, see the discussion in Sect. 5 for the use of a computable subring which works well in our situation.

Example 1.4. A number field F is computable, specified by the data of the minimal polynomial of a primitive element (itself described by the sequence of its coefficients, given as rational numbers); elements of F are described by their standard representation in the basis of powers of the primitive element [6, Sect. 4.2.2]. For a detailed exposition of algorithms for computing with a number field F, see Cohen [6, 7] and Pohst and Zassenhaus [27].

Remark 1.5. Global function fields, i.e. finite extensions of $k(T)$ with k a finite field, can be treated in a parallel fashion to number fields. Unfortunately, at the present time the literature is much less complete in providing a suite of algorithms for computing with integral structures in such fields—particularly in the situation where one works in a relative extension of such fields—despite the fact that some of these algorithms have already been implemented in MAGMA [3] by Hess [14]. Therefore, in this article we will often consider just the case of number fields and content ourselves to notice that the algorithms we provide will generalize with appropriate modifications to the global function field setting.

Throughout this article, when discussing algorithms, we will assume that the domain R and its field of fractions F are computable.

Let B be a F-algebra with $\dim_F B = n$ and basis $e_1, e_2, \dots, e_n$ (as an F-vector space), and suppose $e_1 = 1$. A *multiplication table* for B is a system of n^3 elements $(c_{ijk})_{i,j,k=1,\dots,n}$ of F, called *structure constants*, such that multiplication in B is given by

$$e_i e_j = \sum_{k=1}^{n} c_{ijk} e_k$$

for $i, j \in \{1, \dots, n\}$.

An F-algebra B is represented in bits by a multiplication table and elements of F are represented in the basis e_i. Note that basis elements in B can be multiplied directly by the multiplication table but multiplication of arbitrary elements in B requires $O(n^3)$ arithmetic operations (additions and multiplications) in F; in either case, note the output is of polynomial size in the input for fixed B.

Remark 1.6. We have assumed that B is associative as an F-algebra; however, this property can be verified by simply checking the associative law on a basis.

Remark 1.7. We require that the element 1 be included as a generator of B, since by our definition an F-algebra is equipped with an embedding $F \hookrightarrow B$. This is not a serious restriction, for the equations which uniquely define the element 1 in B are linear equations and so $1 \in B$ can be (uniquely) recovered by linear algebra over F. (And an algebra without 1 embeds inside an algebra with 1.)

An R-algebra $\mathcal{O}$ is represented in bits by the F-algebra $B = \mathcal{O} \otimes_R F$ and a set of R-module generators $x_1, \dots, x_m \in B$ with $x_1 = 1$. A morphism between R-algebras is represented by the underlying R-linear map, specified by a matrix in the given sets of generators for the source and target.

Quaternion algebras. We refer to Vignéras [38] and Reiner [28] for background relevant to this section.

An F-algebra B is *central* if the center of B is equal to F, and B is *simple* if the only two-sided ideals of B are (0) and B (or equivalently that any F-algebra homomorphism with domain B is either the zero map or injective).

Remark 1.8. One can compute the center of B by solving the n linear equations $xe_i = e_i x$ for $x = x_1 e_1 + \cdots + x_n e_n$ and thereby, for example, verify that B is central.

Definition 1.9. A *quaternion algebra* B over F is a central simple F-algebra with $\dim_F B = 4$.

An F-algebra B is a quaternion algebra if and only if there exist $i, j \in B$ which generate B as an F-algebra such that

$$i^2 = a, \quad j^2 = b, \quad ji = -ij \tag{1.10}$$

with $a, b \in F^\times$ if $\operatorname{char} F \neq 2$, and

$$i^2 + i = a, \quad j^2 = b, \quad ji = (i+1)j \tag{1.11}$$

with $a \in F$ and $b \in F^\times$ if $\operatorname{char} F = 2$. We give an algorithmic proof of this equivalence in Sect. 3. We accordingly denote an algebra (1.10), (1.11) by $B = \left(\dfrac{a, b}{F}\right)$, say that B is in *standard form*, and call the elements i, j *standard generators*. Note that B has basis $1, i, j, ij$ as an F-vector space, so indeed $\dim_F B = 4$.

Example 1.12. The ring $\mathrm{M}_2(F)$ of 2×2-matrices with coefficients in F is a quaternion algebra over F. Indeed, we have $\left(\dfrac{1, 1}{F}\right) \cong \mathrm{M}_2(F)$ with $j \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and

$$i \mapsto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{or} \quad i \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

according as $\operatorname{char} F \neq 2$ or $\operatorname{char} F = 2$.

Every quaternion algebra over a separably (or algebraically) closed field F is isomorphic to $\mathrm{M}_2(F)$.

Example 1.13. The $\mathbb{R}$-algebra $\mathbb{H} = \left(\dfrac{-1, -1}{\mathbb{R}}\right)$, generated by i, j satisfying $i^2 = j^2 = (ij)^2 = -1$ is the usual division ring of quaternions over $\mathbb{R}$. Every quaternion algebra over $\mathbb{R}$ is isomorphic to either $\mathrm{M}_2(\mathbb{R})$ or $\mathbb{H}$, according to the theorem of Frobenius.

Let B be an F-algebra. An *R-order* in B is a subring $\mathcal{O} \subset B$ that is finitely generated as an R-module and such that $\mathcal{O}F = B$. We see that an R-algebra $\mathcal{O}$ is an R-order in $B = \mathcal{O} \otimes_R F$, and we will use this equivalence throughout, sometimes thinking of $\mathcal{O}$ as an R-algebra on its own terms and at other times thinking of $\mathcal{O}$ as arising as an order inside an algebra over a field.

A *quaternion order* over R is an R-order in a quaternion algebra B over F. Equivalently, an R-algebra $\mathcal{O}$ is a quaternion order if $B = \mathcal{O} \otimes_R F$ is a quaternion algebra over F.

Example 1.14. $\mathrm{M}_2(R)$ is a quaternion order in $\mathrm{M}_2(F)$.

If $a, b \in R \setminus \{0\}$, then $\mathcal{O} = R \oplus Ri \oplus Rj \oplus Rij$ is a quaternion order in $B = \left(\dfrac{a, b}{F}\right)$. So for example $\mathbb{Z} \oplus \mathbb{Z}i \oplus \mathbb{Z}j \oplus \mathbb{Z}ij$ is a $\mathbb{Z}$-order in the rational Hamiltonians $B = \left(\dfrac{-1, -1}{\mathbb{Q}}\right)$.

Further examples of quaternion orders will be defined in the next section (see Lemma 2.11).

Modules over Dedekind domains. Let R be a Dedekind domain, an integrally closed (noetherian) domain in which every nonzero prime ideal is maximal. Every field is a Dedekind domain (vacuously), as is the integral closure of $\mathbb{Z}$ or $\mathbb{F}_p[T]$ in a finite (separable) extension of $\mathbb{Q}$ or $\mathbb{F}_p(T)$, respectively. The localization of a Dedekind domain at a multiplicative subset is again a Dedekind domain. If R is the ring of integers of a number field, then we call R a *number ring*.

Over a Dedekind domain R, every projective R-module M can be represented as the direct sum of projective R-modules of rank 1, which is to say that there exist projective (equivalently, locally principal) R-modules $\mathfrak{a}_1, \dots, \mathfrak{a}_n \subset F$ (also known as *fractional ideals* of R) and elements $x_1, \dots, x_n \in M$ such that

$$M = \mathfrak{a}_1 x_1 \oplus \cdots \oplus \mathfrak{a}_n x_n;$$

we say then that the elements x_i are a *pseudobasis* for M with *coefficient ideals* $\mathfrak{a}_i$. More generally, if $M = \mathfrak{a}_1 x_1 + \cdots + \mathfrak{a}_m x_m$ (the sum not necessarily direct), then we say the elements x_i are a *pseudogenerating set* for M (with *coefficient ideals* $\mathfrak{a}_i$).

In fact, the above characterization can be made computable as follows.

Proposition 1.15. *Let R be a number ring. Then there exists an algorithm which, given a projective R-module M specified by a pseudogenerating set, returns a pseudobasis for M.*

The algorithm in Proposition 1.15 is a generalization of the Hermite normal form (HNF) for matrices over $\mathbb{Z}$; see Cohen [7, Chap. 1]. Therefore, from now on we represent a quaternion order $\mathcal{O}$ over a number ring R by a pseudobasis; in such a situation, we may and do assume that $\mathfrak{a}_1 = R$ and $x_1 = 1$ (by employing the HNF).

Remark 1.16. Recalling Remark 1.5, in particular there seems to be no comprehensive reference for results akin to Proposition 1.15 in the global function field case.

2 Standard Involutions and Degree

Quaternion algebras, or more generally algebras which have a standard involution, possess a quadratic form called the reduced norm. In this section, we discuss this association and we give an algorithm which verifies that an algebra has a standard involution. As a reference, see Jacobson [17, Sect. 1.6], Knus [19], and work of the author [40].

In this section, let R be an integrally closed (noetherian) domain with field of fractions F. Let $\mathcal{O}$ be an R-algebra and let $B = \mathcal{O} \otimes_R F$.

Degree. We first generalize the notion of degree from field extensions to R-algebras.

Definition 2.1. The *degree* of $x \in \mathcal{O}$ over R, denoted $\deg_R(x)$, is the smallest positive integer n such that x satisfies a monic polynomial of degree n with coefficients in R. The *degree* of $\mathcal{O}$ over R, denoted $\deg_R(\mathcal{O})$, is the smallest positive integer n such that every element of $\mathcal{O}$ has degree at most n.

Every $x \in \mathcal{O}$ satisfies the characteristic polynomial of (left) multiplication by x on a set of generators for $\mathcal{O}$ as an R-module, and consequently $\deg_R(\mathcal{O}) < \infty$ (under our continuing hypothesis that $\mathcal{O}$ is projective of finite rank).

Lemma 2.2. *We have* $\deg_R(\mathcal{O}) = \deg_F(B)$.

Proof. Since $\mathcal{O}$ is finitely generated as an R-module and R is noetherian, the R-submodule $R[x] \subset \mathcal{O}$ is finitely generated, so x is integral over R. Since R is integrally closed, the minimal polynomial of $x \in \mathcal{O}$ over F has coefficients in R by Gauss's lemma, so $\deg_R(x) = \deg_F(x)$ and thus $\deg_R(\mathcal{O}) \leq \deg_F(B)$. On the other hand, if $y \in B$ then there exists $0 \neq d \in R$ such that $x = yd \in \mathcal{O}$ so $\deg_F(x) = \deg_F(y) = \deg_R(y)$ so $\deg_F(B) \leq \deg_R(\mathcal{O})$. □

From the lemma, we need only consider the degree of an algebra over a field.

Example 2.3. B has degree 1 if and only if $B = F$.

If K is a separable field extension of F with $\dim_F K = n$, then K has degree n as a F-algebra (in the above sense) by the primitive element theorem.

If $\dim_F B = n$, then B has degree at most n but even if B is commutative one may still have $\deg_F(B) < \dim_F B$: for example, $B = F[x,y,z]/(x,y,z)^2$ has rank 4 over the field F but has degree 2.

Standard involutions. We will see in a moment that quaternion orders and algebras are algebras of degree 2; this will be a consequence of the fact that they possess a standard involution. Indeed, the link between algebras with an involution and quadratic forms the heart of much important work [20].

Definition 2.4. An *anti-automorphism* of $\mathcal{O}$ is an R-linear map $\overline{} : \mathcal{O} \to \mathcal{O}$ with $\overline{1} = 1$ and $\overline{xy} = \overline{y}\,\overline{x}$ for all $x \in \mathcal{O}$. An *involution* is an anti-automorphism such that $\overline{\overline{x}} = x$ for all $x \in \mathcal{O}$. An involution is *standard* if $x\overline{x} \in R$ for all $x \in \mathcal{O}$.

Note that if $x\overline{x} \in R$ for all $x \in \mathcal{O}$, then $(x+1)(\overline{x}+1) = x\overline{x} + (x+\overline{x}) + 1 \in R$ and hence $x + \overline{x} \in R$ for all $x \in \mathcal{O}$ as well. Note that $\overline{x}x = x\overline{x}$ for all $x \in \mathcal{O}$ since $x(x+\overline{x}) = (x+\overline{x})x$ (and R is central in $\mathcal{O}$).

Example 2.5. If $\mathcal{O} = \mathrm{M}_n(R)$, then the transpose map is an anti-automorphism which is standard if and only if $n = 1$; the adjoint map is a standard involution for $n \leq 2$ but is not R-linear for $n \geq 3$.

Suppose now that $\mathcal{O}$ has a standard involution $\overline{}$. Then we define the *reduced trace* and *reduced norm*, respectively, to be the maps

$$\begin{aligned} \operatorname{trd} : \mathcal{O} &\to R & \qquad \operatorname{nrd} : \mathcal{O} &\to R \\ x &\mapsto x + \overline{x} & x &\mapsto x\overline{x} = \overline{x}x \end{aligned}$$

We have

$$x^2 - \operatorname{trd}(x)x + \operatorname{nrd}(x) = x^2 - (x+\overline{x})x + x\overline{x} = 0 \tag{2.6}$$

for all $x \in \mathcal{O}$. It follows that if $\mathcal{O}$ has a standard involution then either $\mathcal{O} = R$ (so the standard involution is the identity and $\mathcal{O} = R$ has degree 1) or $\mathcal{O}$ has degree 2.

Example 2.7. Let $B = \left(\dfrac{a,b}{F}\right)$ be a quaternion algebra over F. Then B has a standard involution, defined as follows. For $x = t + ui + vj + wk$, we have

$$\overline{x} = t - ui - vj - wk$$

so $\operatorname{trd}(x) = 2t$ and $\operatorname{nrd}(x) = t^2 - au^2 - bv^2 + abw^2$ if $\operatorname{char} F \neq 2$ and

$$\overline{x} = t + (u+1)i + vj + wk$$

so $\operatorname{trd}(x) = 2u$ and $\operatorname{nrd}(x) = t^2 + tu + au^2 + bv^2 + bvw + abw^2$ if $\operatorname{char} F = 2$.

Lemma 2.8. *$\mathcal{O}$ has a standard involution if and only if $B = \mathcal{O} \otimes_R F$ has a standard involution.*

Proof. If $\mathcal{O}$ has a standard involution, we obtain one on B by extending F-linearly. Conversely, suppose B has a standard involution and let $x \in \mathcal{O}$. Then as in the proof of Lemma 2.2, x is integral over R so its minimal polynomial over F has coefficients in R. If $x \in R$, then $\overline{x} = x$ and there is nothing to prove. If $x \notin R$, this minimal polynomial must be given by (2.6), so $\operatorname{trd}(x) = x + \overline{x} \in R$ and thus $\overline{x} = \operatorname{trd}(x) - x \in \mathcal{O}$ has $x\overline{x} = \operatorname{nrd}(x) \in R$ as well. □

An R-algebra S is *quadratic* if S has rank 2 as an R-module.

Lemma 2.9. *Let S be a quadratic R-algebra. Then S is commutative and has a unique standard involution.*

Proof. By Lemma 2.8, it suffices to prove the lemma for $K = S \otimes_R F$. But then for any $x \in K \setminus F$ we have $K = F \oplus Fx$ so K is commutative. Moreover, we have $x^2 - tx + n = 0$ for some unique $t, n \in F$ and so the (necessarily unique) standard involution is given by $x \mapsto t - x$, extending by F-linearity. (See also Scharlau [33, Sect. 8.11] for a proof of this lemma.) □

Corollary 2.10. *If $\mathcal{O}$ has a standard involution, then this involution is unique.*

This corollary follows immediately from Lemma 2.9 by restricting to quadratic subalgebras K of B.

Quaternion orders. Having identified the standard involution on a quadratic algebra, we now generalize the construction of quaternion algebras (1.10), (1.11) to quaternion orders. Let S be a quadratic R-algebra, and suppose S is *separable*, so the minimal polynomial of every $x \in S$ has distinct roots over the algebraic closure $\overline{F}$ of F. Let $J \subset S$ be an invertible S-ideal (equivalently, a locally principal S-module) and let $b \in R \setminus \{0\}$. We denote by $\left(\dfrac{S, J, b}{R}\right)$ the R-algebra $S \oplus Jj$ subject to the relations $j^2 = b$ and $ji = \bar{i}j$ for all $i \in S$, where $\overline{}$ denotes the unique standard involution on S obtained from Lemma 2.9. We say that such an algebra is in *standard form.*

Lemma 2.11. *The R-algebra $\mathcal{O} = \left(\dfrac{S, J, b}{R}\right)$ is a quaternion order.*

Proof. We consider $B = \mathcal{O} \otimes_R F$. Let $K = S \otimes_R F$ and let $i \in K \setminus F$. Since K is separable, if $\operatorname{char} F \neq 2$ by completing the square we may assume $i^2 = a$ with $a \in F^\times$; if $\operatorname{char} F = 2$, we may assume $i^2 + i = a$ with $a \in F$. Now since J is projective we have $J \otimes_R F = J \otimes_S K \cong K$ so $B \cong K \oplus Kj$ as an F-algebra. Finally, since $ji = \bar{i}j = (\operatorname{trd}(i) - i)j$ and $\operatorname{trd}(i) = 0, 1$ according as $\operatorname{char} F \neq 2$ or not, we have identified B as isomorphic to the quaternion algebra $\left(\dfrac{a, b}{F}\right)$. □

Algorithmically identifying a standard involution. We conclude this section with an algorithm to test if an F-algebra B (of dimension n) has a standard involution.

First, we note that if B has a standard involution $\overline{} : B \to B$, then this involution and hence also the reduced trace and norm can be computed efficiently. Indeed, let $\{e_i\}_i$ be a basis for B; then $\operatorname{trd}(e_i) \in F$ is simply the coefficient of e_i in e_i^2, and so $\overline{e_i} = \operatorname{trd}(e_i) - e_i$ for each i can be precomputed for B; one recovers the involution on B (and hence also the trace) for an arbitrary element of B by F-linearity. Therefore the involution and the reduced trace can be computed using $O(n)$ arithmetic operations in F (with output linear in the input for fixed B) and the reduced norm using $O(n^2)$ operations in F (with output quadratic in the input).

Algorithm 2.12. Let B be an F-algebra given by a multiplication table in the basis $e_1, \dots, e_n$ with $e_1 = 1$. This algorithm returns TRUE if and only if B has a standard involution.

1. For $i = 2, \dots, n$, let $t_i \in F$ be the coefficient of e_i in e_i^2, and let $n_i = e_i^2 - t_i e_i$. If some $n_i \notin F$, return FALSE.
2. For $i = 2, \dots, n$ and $j = i+1, \dots, n$, let $n_{ij} = (e_i + e_j)^2 - (t_i + t_j)(e_i + e_j)$. If some $n_{ij} \notin F$, return FALSE. Otherwise, return TRUE.

Proof of correctness. Let $F[x] = F[x_1, \dots, x_n]$ be the polynomial ring over F in n variables, and let $B_{F[x]} = B \otimes_F F[x]$. Let $\xi = x_1 + x_2 e_2 + \dots + x_n e_n \in B_{F[x]}$, and define

$$t_\xi = \sum_{i=1}^n t_i x_i$$

and

$$n_\xi = \sum_{i=1}^n n_i x_i^2 + \sum_{1 \le i < j \le n} (n_{ij} - n_i - n_j) x_i x_j.$$

Let

$$\xi^2 - t_\xi \xi + n_\xi = \sum_{i=1}^n c_i(x_1, \dots, x_n) e_i$$

with $c_i(x) \in F[x]$. Each $c_i(x)$ is a homogeneous polynomial of degree 2. The algorithm then verifies that $c_i(x) = 0$ for $x \in \{e_i\}_i \cup \{e_i + e_j\}_{i,j}$, and this implies that each $c_i(x)$ vanishes identically. Therefore, the specialization of the map $\xi \mapsto \overline{\xi} = t_\xi - \xi$ is the unique standard involution on B. □

Remark 2.13. Algorithm 2.12 requires $O(n)$ arithmetic operations in F, since e_i^2 can be computed directly from the multiplication table and hence $(e_i + e_j)^2 = e_i^2 + e_i e_j + e_j e_i + e_j^2$ can be computed using $O(4n) = O(n)$ operations.

3 Algebras with a Standard Involution and Quadratic Forms

In this section, we describe a relationship between R-algebras with a standard involution and quadratic forms over R. The main result of this section is an algorithm which verifies that an R-algebra $\mathcal{O}$ over a local PID is a quaternion order and, if so, exhibits standard generators for $\mathcal{O}$. Specializing, we will thereby recognize quaternion algebras over a field F. We then extend this to recognizing quaternion orders over a number ring R. Over fields, a reference for this section is Lam [21], and for more about algebras equipped with a quadratic norm form, we refer the reader to Knus [19].

Quadratic forms over rings. We begin by defining quadratic forms over a (noetherian) domain R.

Definition 3.1. A *quadratic form* over R is a map $Q : M \to R$, where M is a finitely generated projective R-module, such that:

(i) $Q(ax) = a^2 Q(x)$ for all $a \in R$ and $x \in M$; and
(ii) The map $T : M \times M \to R$ defined by

$$T(x,y) = Q(x+y) - Q(x) - Q(y)$$

is R-bilinear.

A symmetric bilinear form $T : M \times M \to R$ is *even* if $T(x,x) \in 2R$ for all $x \in M$. If T arises from a quadratic form, then T is even, and conversely if T is even and 2 is a nonzerodivisor in R then one recovers the quadratic form as $Q(x) = T(x,x)/2$.

Let $Q : M \to R$ be a quadratic form and suppose that M is free over R with basis $e_1, \dots, e_n$. The *Gram matrix* of Q with respect to the basis $e_1, \dots, e_n$ is the matrix $A = (T(e_i, e_j))_{i,j=1,\dots,n} \in \mathrm{M}_n(R)$. The matrix A has the property that $x^t A y = T(x,y)$, where we identify $x = x_1 e_1 + \dots + x_n e_n$ with the column vector $(x_1, \dots, x_n)^t$, and similarly for y. In particular we have $x^t A x = 2Q(x)$.

Let $Q : M \to R$ be a quadratic form. We say $x, y \in M$ are *orthogonal* (with respect to Q) if $T(x,y) = 0$.

Example 3.2. Let $\mathcal{O}$ be an R-algebra with a standard involution $\overline{}$. Then the reduced norm $\mathrm{nrd} : \mathcal{O} \to R$ (defined by $x \mapsto x\overline{x}$ for $x \in \mathcal{O}$) is a quadratic form on $\mathcal{O}$ with associated bilinear form

$$T(x,y) = x\overline{y} + y\overline{x} = \mathrm{trd}(x\overline{y}) = \mathrm{trd}(x)y + \mathrm{trd}(y)x - (xy+yx) = \mathrm{trd}(\overline{x}y) \quad (3.3)$$

for $x, y \in \mathcal{O}$. In particular $T(1,x) = T(x,1) = \mathrm{trd}(x)$. Note that $x, y \in \mathcal{O}$ are orthogonal if and only if $x\overline{y} = -y\overline{x}$, and if further $\mathrm{trd}(x) = \mathrm{trd}(y) = 0$ then $\overline{x} = -x$ and $\overline{y} = -y$ so x, y are orthogonal if and only if $xy = -yx$.

Example 3.4. Let $\mathcal{O}_0 = \{x \in \mathcal{O} : \mathrm{trd}(x) = 0\}$ be the R-submodule of elements of reduced trace zero. Then $\mathcal{O}/\mathcal{O}_0$ is torsion-free, since if $rx \in \mathcal{O}_0$ then $\mathrm{trd}(rx) = r\,\mathrm{trd}(x) = 0$ so $\mathrm{trd}(x) = 0$ so $x \in \mathcal{O}_0$. Thus if R is a Dedekind domain or $2 \in R^\times$, then $\mathcal{O}_0$ is a projective R-submodule of $\mathcal{O}$ and $\mathcal{O} \supset R \oplus \mathcal{O}_0$. We therefore obtain a quadratic form $\mathrm{nrd}_0 = \mathrm{nrd}\,|_{\mathcal{O}_0} : \mathcal{O}_0 \to R$.

If $Q : M \to R$ and $Q' : M' \to R$ are quadratic forms, we define the form $Q \perp Q'$ on $M \oplus M'$ by requiring that $(T \perp T')(x + x') = T(x) + T(x')$ and $(Q \perp Q')(x + x') = Q(x) + Q(x')$. (Note that $T(x,x) = 2Q(x)$ for all $x \in M$ so if $2 \neq 0 \in R$ then the second condition follows from the first.)

Let $Q : M \to R$ be a quadratic form and suppose that M is free (of finite rank). In this case, a basis $e_1, \dots, e_n$ for M gives an isomorphism $M \cong R^n$ in which Q can be written

$$Q(x) = Q(x_1e_1 + \cdots + x_ne_n) = \sum_i Q(e_i)x_i^2 + \sum_{i<j} T(e_i, e_j)x_ix_j$$

with $x = (x_1, \ldots, x_n) \in R^n$.

For $a \in R$, the quadratic form $Q(x) = ax^2$ on R is denoted $\langle a\rangle$; similarly, for $a_1, \ldots, a_n \in R$, we abbreviate $\langle a_1\rangle \perp \cdots \perp \langle a_n\rangle = \langle a_1, \ldots, a_n\rangle$. For $a, b, c \in R$, the quadratic form $Q(x, y) = ax^2 + bxy + cy^2$ on R^2 is denoted $[a, b, c]$.

Example 3.5. Let $B = \left(\dfrac{a, b}{F}\right)$ be a quaternion algebra over F. Then as in Example 2.7, in the basis $1, i, j, ij$ we have $\operatorname{nrd} \cong \langle 1, -a, -b, ab\rangle \cong \langle 1, -a\rangle \perp -b\langle 1, -a\rangle$ if $\operatorname{char} F \neq 2$ and $\operatorname{nrd} \cong [1, 1, a] \perp b[1, 1, a]$ if $\operatorname{char} F = 2$.

Similarly, for $\operatorname{nrd}_0 : B_0 \to F$ we have $\operatorname{nrd}_0 \cong \langle -a, -b, ab\rangle \cong \langle -a\rangle \perp -b\langle 1, -a\rangle$ if $\operatorname{char} F \neq 2$ and $\operatorname{nrd}_0 \cong \langle 1\rangle \perp b[1, 1, a]$ if $\operatorname{char} F = 2$.

Quadratic forms over DVRs. Now let R be a local PID. Then R has valuation $\operatorname{ord}_v : R \to \mathbb{Z}_{\geq 0} \cup \{\infty\}$ and uniformizer π. If $R = F$ is a field, then $\pi = 1$ and the valuation is trivial, i.e. $\operatorname{ord}_v(x) = 0$ for $x \in F^\times$ (and $\operatorname{ord}_v(0) = \infty$).

Let $Q : M \to R$ be a quadratic form over R. Then since R is a PID, M is free; let n be the rank of M over R. We will now seek to find a basis for R^n in which a quadratic form Q has a particularly simple form: we will seek to diagonalize Q as far as possible. In cases where $2 \in R^\times$, we can accomplish a full diagonalization; otherwise, we can at least break up the form as much as possible, as follows.

A quadratic form Q over R is *atomic* if either:

(i) $Q \cong \langle a\rangle$ for some $a \in R^\times$, or
(ii) $2 \notin R^\times$ and $Q \cong [a, b, c]$ with $a, b, c \in R$ satisfying

$$\operatorname{ord}_v(b) < \operatorname{ord}_v(2a) \leq \operatorname{ord}_v(2c) \text{ and } \operatorname{ord}_v(a)\operatorname{ord}_v(b) = 0.$$

In case (ii), we necessarily have $\operatorname{ord}_v(2) > 0$ and $\operatorname{ord}_v(b^2 - 4ac) = 2\operatorname{ord}_v(b)$.

Example 3.6. If $2 \in R^\times$, then a quadratic form Q is atomic if and only if $Q(x) = ax^2$ for $a \in R^\times$.

Example 3.7. If $R = F$ is a field with $\operatorname{char} F = 2$, then $[a, b, c]$ is atomic if and only if $b \in F^\times$; scaling y by a/b realizes this form as isomorphic to $a[1, 1, ca/b^2]$ with $a \in F^\times$. Therefore, over fields, recording the middle coefficient is unnecessary, and indeed other texts use $[a, b]$ to denote the quadratic form $ax^2 + xy + by^2$.

For example, take $R = \mathbb{Z}_2[\sqrt{2}]$ with normalized valuation $\operatorname{ord}_v(\sqrt{2}) = 1$ and let $Q(x, y) = x^2 + \sqrt{2}xy$. Then according to our definition, Q is atomic, since $\operatorname{ord}_v(b) = 1 < \operatorname{ord}_v(2a) = 2 \leq \operatorname{ord}_v(2c) = \infty$ and $\operatorname{ord}_v(a) = 0$. But this form is not globally divisible by any element of positive valuation, and a calculation shows that any isomorphic (equivalent) form has middle coefficient of positive valuation.

Example 3.8. Suppose $R = \mathbb{Z}_2$ is the ring of 2-adic integers, so that $\operatorname{ord}_v(x) = \operatorname{ord}_2(x)$ is the largest power of 2 dividing $x \in \mathbb{Z}_2$. Recall that $\mathbb{Z}_2^\times/\mathbb{Z}_2^{\times 2}$ is represented by the elements $\pm 1, \pm 5$, therefore a quadratic form Q over $\mathbb{Z}_2$ is atomic of type (i) above if and only if $Q(x) \cong \pm x^2$ or $Q(x) \cong \pm 5x^2$. For forms of type (ii), the conditions $\operatorname{ord}_v(b) < \operatorname{ord}_v(2a) = \operatorname{ord}_v(a) + 1$ and $\operatorname{ord}_v(a) \operatorname{ord}_v(b) = 0$ imply in fact $\operatorname{ord}_v(b) = 0$, and so a quadratic form Q over $\mathbb{Z}_2$ is atomic of type (ii) if and only if $Q(x, y) \cong ax^2 + xy + cy^2$ with $\operatorname{ord}_2(a) \leq \operatorname{ord}_2(c)$. Replacing x by ux and y by $u^{-1}y$ for $u \in \mathbb{Z}_2^\times$ we may assume $a = \pm 2^t$ or $a = \pm 5 \cdot 2^t$ with $t \geq 0$, and then the atomic representative $[a, 1, c]$ of the isomorphism class of Q is unique.

A quadratic form Q is *decomposable* if Q can be written as the orthogonal sum of two quadratic forms ($Q \cong Q_1 \perp Q_2$) and is *indecomposable* otherwise.

It follows by induction on the rank of M that Q is the orthogonal sum of indecomposable forms. We will soon give an algorithmic proof of this fact and write each indecomposable form as a scalar multiple of an atomic form. We begin with the following lemma.

Lemma 3.9. *An atomic form Q is indecomposable.*

Proof. If Q is atomic of type (i) then the space underlying Q has rank 1, so this is clear. So suppose $Q = [a, b, c]$ is atomic of type (ii) and suppose Q is decomposable. It follows that if $x, y \in M$ then $T(x, y) \in 2R$. Thus we cannot have $\operatorname{ord}_v(b) = 0$, so $\operatorname{ord}_v(a) = 0$, and further $\operatorname{ord}_v(b) \geq \operatorname{ord}_v(2) = \operatorname{ord}_v(2a)$; this contradicts the fact that Q is atomic. □

Proposition 3.10. *Let R be a local PID and let $Q : M \to R$ be a quadratic form. Then there exists a basis of M such that the form Q can be written*

$$Q \cong \pi^{e_1} Q_1 \perp \cdots \perp \pi^{e_n} Q_n$$

where the forms Q_i are atomic and $0 \leq e_1 \leq \cdots \leq e_n \leq \infty$.

In the above proposition, we interpret $\pi^\infty = 0$. A form as presented in Proposition 3.10 is called *normalized*, and the integer e_i is called the *valuation* of $\pi^{e_i} Q_i$. The tuple of valuations e_i for Q is unique.

Example 3.11. By Example 3.5, if B is a quaternion algebra over a field F then the quadratic form nrd is normalized in the basis $1, i, j, ij$, with a similar statement for nrd_0.

We give an algorithmic proof of Proposition 3.10. (Over fields, see Lam [21, Sect. 1.2], and see Scharlau [33, Sect. 9.4] for fields of characteristic 2.)

Algorithm 3.12. Let R be a computable ring which is a local PID with (computable) valuation $\operatorname{ord}_v : R \to \mathbb{Z}_{\geq 0} \cup \{\infty\}$.

Let $Q : M \to R$ be a quadratic form over R and let $e_1, \ldots, e_n$ be a basis for M. This algorithm returns a basis of M in which Q is normalized.

1. If $T(e_i, e_j) = 0$ for all i, j, return $f_i := e_i$. Otherwise, let (i, j) with $1 \leq i \leq j \leq n$ be such that $\operatorname{ord}_v T(e_i, e_j)$ is minimal, taking $i = j$ if possible and if not taking i minimal.

2. If $i = j$, let $f_1 := e_i$ and proceed to Step 3. If $i \neq j$ and $2 \in R^\times$, let $f_1 := e_i + e_j$ and proceed to Step 3. Otherwise, proceed to Step 4.
3. Let $e_i := e_1$. For $k = 2, \dots, n$ let

$$f_k := e_k - \frac{T(f_1, e_k)}{T(f_1, f_1)} f_1.$$

Let $m = 2$ and proceed to Step 5.
4. (We have $2 \notin R^\times$ and $i \neq j$.) Let

$$f_1 := \frac{\pi^{\operatorname{ord}_v T(e_i, e_j)}}{T(e_i, e_j)} e_i,$$

$f_2 := e_j$, $e_i := e_1$ and $e_j := e_2$. Let $d := T(f_1, f_1)T(f_2, f_2) - T(f_1, f_2)^2$. For $k = 3, \dots, n$, let

$$\begin{aligned} t_k &:= T(f_1, f_2)T(f_2, e_k) - T(f_2, f_2)T(f_1, e_k) \\ u_k &:= T(f_1, f_2)T(f_1, e_k) - T(f_1, f_1)T(f_2, e_k) \end{aligned}$$

and let

$$f_k := e_k + \frac{t_k}{d} f_1 + \frac{u_k}{d} f_2.$$

Let $m = 3$.
5. Recursively call the algorithm with $M = Rf_m \oplus \cdots \oplus Rf_n$, and return $f_1, \dots, f_{m-1}$ concatenated with the output basis.

Given such a basis, one recovers the normalized quadratic form by factoring out in each atomic form the minimal valuation achieved. (One can also keep track of this valuation along the way in the above algorithm, if desired.)

Remark 3.13. Note that if $2 \in R^\times$, then this algorithm computes a diagonalization of the form Q, ordering the coefficients by their valuation.

Proof of correctness. In Step 3, we verify that $\operatorname{ord}_v T(f_1, f_1) \leq \operatorname{ord}_v T(f_1, e_k)$. Indeed, we have

$$T(f_1, f_1) = T(e_i, e_i) + 2T(e_i, e_j) + T(e_j, e_j)$$

and so $\operatorname{ord}_v T(f_1, f_1) = \operatorname{ord}_v T(e_i, e_j)$ by the ultrametric inequality and the hypotheses that $\operatorname{ord}_v T(e_i, e_j) < \operatorname{ord}_v T(e_i, e_i), T(e_j, e_j)$ and $\operatorname{ord}_v(2) = 0$. So Steps 2 and 3 give correct output.

We have left to check Step 4. This is proven by letting $f_k = e_k + t_k f_1 + u_k f_2$ and solving the linear equations $T(f_1, f_k) = T(f_2, f_k) = 0$ for t_k, u_k. The result then follows from a direct calculation, coupled with the fact that $\operatorname{ord}_v(d) = 2 \operatorname{ord}_v T(f_1, f_2) \leq \operatorname{ord}_v(t_k)$ (and similarly with u_k). This case only arises if (and only if)

$$\operatorname{ord}_v T(f_1, f_2) < \operatorname{ord}_v T(f_1, f_1) = \operatorname{ord}_v(2Q(f_1)) \leq \operatorname{ord}_v(2Q(f_2))$$

so the corresponding block is indeed atomic. □

Example 3.14. Consider the binary quadratic form $[a, b, c]$ over $\mathbb{Z}_2$. Then $T(e_1, e_1) = 2a$, $T(e_1, e_2) = b$, and $T(e_2, e_2) = 2c$. We follow the course of Algorithm 3.12. If $\operatorname{ord}_v(2a)$ is minimal, then in Steps 2 and 3 we diagonalize (complete the square): we have $f_1 = e_1$ and $f_2 = e_2 - (b/2a)e_1$ and so we obtain the (isomorphic) form $\langle a, c + b^2/4a\rangle$. If $\operatorname{ord}_v(2c)$ is minimal, then we similarly obtain $\langle c, a + b^2/4c\rangle$. Finally, if $\operatorname{ord}_2(b)$ is minimal, then we enter Step 4. Since (i, j) was taken with i minimal, for illustration we may suppose $i = 1$ and $j = 2$. Then we have $t = \operatorname{ord}_v(b) < \operatorname{ord}_v(2a) \leq \operatorname{ord}_v(2c)$. Writing $a = 2^t a'$, $b' = 2^t b'$ and $c' = 2^t c'$, in Step 4, we simply have $f_1 = (1/b')e_1$ and $f_2 = e_2$ and we obtain the form $2^t[a'/(b')^2, 1, c']$ and $[a'/(b')^2, 1, c']$ is indeed atomic.

Example 3.15. Consider the form $q(x, y, z) = xy + xz$ over $\mathbb{Z}_2$. We enter Step 4 with $f_1 = e_1$ and $f_2 = e_2$. We compute that $d = -T(f_1, f_2) = -1$, and $t_3 = 0$ and $u_3 = 1$. Thus $f_3 = e_3 - f_2 = e_3 - e_2$, and we obtain the form $[0, 1, 0] \perp \langle 0\rangle$.

We note that Algorithm 3.12 requires $O(n^2)$ arithmetic operations in R. This algorithm can be modified suitably to operate on the Gram matrix $(T(e_i, e_j))_{i,j}$ of the quadratic form Q, which as explained above recovers the quadratic form when $2 \neq 0 \in R$.

For a quadratic form $Q : M \to R$, we define

$$\operatorname{rad}(Q) = \{x \in M : T(x, y) = 0 \text{ for all } y \in M\};$$

we say Q is *nonsingular* if $\operatorname{rad}(Q) = \{0\}$.

Example 3.16. We have $\operatorname{rad}(Q \perp Q') = \operatorname{rad}(Q) \oplus \operatorname{rad}(Q')$, and if Q is atomic then $\operatorname{rad}(Q) = \{0\}$. In particular, one can read off $\operatorname{rad}(Q)$ directly from a normalized form by the corresponding valuations.

Identifying quaternion algebras. Using the above normalization of a quadratic form in the case where $R = F$ is a field, we can directly identify quaternion algebras amongst algebras with a standard involution.

Proposition 3.17. *Let B be an F-algebra with a standard involution. If $\dim_F B = 4$, then B is a quaternion algebra if and only if* nrd *is nonsingular.*

Proof. If B is a quaternion algebra, then nrd is nonsingular by Example 3.5.

Conversely, B has a basis $1, i, j, k$ which is a normalized basis for Q. First suppose $\operatorname{char} F \neq 2$. By orthogonality we have $\operatorname{trd}(i) = 0$ so $i^2 = -\operatorname{nrd}(i) = a \neq 0$ by nonsingularity and similarly $j^2 = b \neq 0$, and $ji + ij = 0$ from (3.3) so $(ij)^2 = -ab$. Thus $B \supset \left(\dfrac{a, b}{F}\right)$ hence this map is an isomorphism. The case $\operatorname{char} F = 2$ follows similarly: now instead we have $i^2 + i = a$ and $ji = \bar{i}j = (i + 1)j$. □

Proposition 3.17 yields the following algorithm.

Algorithm 3.18. Let B be an F-algebra with $\dim_F B = 4$ (specified by a multiplication table). This algorithm returns TRUE if and only if B is a quaternion algebra, and if so returns an isomorphism $B \cong \left(\dfrac{a,b}{F}\right)$.

1. Verify that B has a standard involution by calling Algorithm 2.12. If not, return FALSE.
2. Compute a normalized basis $1, i, j, k$ for the quadratic form $\operatorname{nrd} : B \to F$ by calling Algorithm 3.12.
3. Test if nrd is nonsingular as in Example 3.16. If so, return TRUE and the quaternion algebra $\left(\dfrac{a,b}{F}\right)$ given by the standard generators i, j.

Remark 3.19. Given a quaternion algebra over $\mathbb{Q}$, Rónyai [29, Theorem 2.1] gives an algorithm to compute a standard representation, but this algorithm tests a polynomial of degree 2 over $\mathbb{Q}$ for irreducibility; the above algorithm requires no such test.

Remark 3.20. If in Step 3 one finds that nrd is not nonsingular, then one has the further refinement of Algorithm 3.18 as follows.

We denote by $\operatorname{rad}(B)$ the *Jacobson radical* of B, the largest two-sided *nil ideal* of B, i.e. the largest two-sided ideal in which every element is nilpotent. An algebra B for which $\operatorname{rad}(B) = \{0\}$ is called *semisimple*. We claim that $\operatorname{rad}(B) = \operatorname{rad}(\operatorname{nrd})$. Indeed, let $e \in B$ be nilpotent, so that $e^2 = 0$. For any $x \in B$, we have by (3.3) that

$$xe + ex = \operatorname{trd}(x)e + \operatorname{trd}(xe).$$

It follows that e generates a nil ideal if and only if $T(x, e) = 0$ for all $x \in B$, which holds if and only if $x \in \operatorname{rad}(\operatorname{nrd})$. Thus $\operatorname{rad}(B) = \operatorname{rad}(\operatorname{nrd})$. One can then easily modify the algorithm to output $\operatorname{rad}(B) = \operatorname{rad}(\operatorname{nrd})$.

Remark 3.21. Another algorithm which tests if B is a quaternion algebra (but does not give a standard representation) under the assumption $\operatorname{char} F = 0$ runs as follows. (See Lam [21, Chap. 4] for the standard facts we use.) By the Wedderburn-Artin theorem and a dimension count, the algebra B over F is a quaternion algebra if and only if B is central and semisimple. We verify that B is central as in Remark 1.8. To verify semisimplicity, if $\operatorname{char} F = 0$, Dickson [10, Sect. 66] showed that B with $\dim_F B = n$ is semisimple if and only if the matrix $(\operatorname{Tr}(e_i e_j))_{i,j=1,\dots,n}$ has full rank n, where Tr is the (left) algebra trace.

In view of Algorithm 3.18, we assume from now on that a quaternion algebra B over a field F is given as input by a standard representation.

Over a general domain R, the above algorithms do not generalize directly, as we cannot hope to normalize a quadratic form in such a simple way for over rings that are no longer local PIDs. Indeed, the category of quadratic forms over a general domain R can be quite complicated—already forms over the integers $\mathbb{Z}$ are of significant interest. However, over Dedekind domains, we can still recognize quaternion orders, and one instead understands these orders as in Sect. 1 via their localizations, a subject which will consume the later sections of this article.

Identifying quaternion orders. Let F be a number field and let $\mathbb{Z}_F$ be its ring of integers. In this section, we give an algorithm which allows us in many cases to put quaternion orders in a standard form as in the discussion of Lemma 2.11.

Algorithm 3.22. Let $\mathcal{O} \subset B$ be a quaternion order over $\mathbb{Z}_F$. Let $\iota : K \to B$ be an embedding of F-algebras with K a field such that $[K : F] = 2$ and $\iota(K) \cap \mathcal{O} = \mathbb{Z}_K$ is maximal. This algorithm returns a fractional ideal $\mathfrak{b}$ of K, an element $j \in \mathcal{O}$ such that $\mathcal{O} = \iota(\mathbb{Z}_K) \oplus \iota(\mathfrak{b})j \cong \left(\dfrac{\mathbb{Z}_K, \mathfrak{b}, b}{\mathbb{Z}_F}\right)$.

1. Identify K with $\iota(K)$. Let $K = F \oplus Fi$ with $i \in B$. Compute $j \in B$ orthogonal to $1, i$.
2. Let $x_1, \dots, x_m$ be a generating set for $\mathcal{O}$ as a $\mathbb{Z}_F$-module. Write $x_k = a_k + b_k j$ with $a_k, b_k \in K$ for $k = 1, \dots, m$.
3. Compute a pseudo-basis $\mathbb{Z}_K \oplus \mathfrak{b}j$ for the $\mathbb{Z}_K$-module generated by (a_k, b_k) for $k = 1, \dots, m$ using a HNF.
4. Let a, b be generators for $\mathfrak{b}$ as an $\mathbb{Z}_F$-module. If $\operatorname{trd}(j) \neq 0$, then let $c := \operatorname{trd}(bj)a - \operatorname{trd}(aj)b$, let $j := cj$ and $\mathfrak{b} := (1/c)\mathfrak{b}$. Return $\mathfrak{b}$ and the element j.

Proof of correctness. In Step 4, we check directly that $\operatorname{trd}(j) = \operatorname{trd}(ij) = 0$, as desired. □

Remark 3.23. One can extend Algorithm 3.22 when $\iota(K) \cap \mathcal{O} = S$ is no longer maximal by an appropriate modification of the HNF algorithm over S.

4 Identifying the Matrix Ring

In this section, we continue the pursuit of our motivating question and address the computational complexity of identifying the matrix ring over a field. Throughout this section, let F be a computable field. We represent a quaternion algebra B over F by a standard form $B = \left(\dfrac{a, b}{F}\right)$.

Problem (IsMatrixRing). *Given a quaternion algebra B over F, determine if $B \cong \mathrm{M}_2(F)$.*

We may also ask for a solution to the more difficult problem of constructing an explicit isomorphism.

Problem (ExhibitMatrixRing). *Given a quaternion algebra B over F, determine if $B \cong \mathrm{M}_2(F)$ and, if so, output such an isomorphism.*

Zerodivisors. Let B be a quaternion algebra. The following structural lemma allows us to address the above problems.

Lemma 4.1. *The following are equivalent:*

(i) $B \cong \mathrm{M}_2(F)$;
(ii) *B is not a division ring;*
(iii) *There exists a nonzero $e \in B$ such that $e^2 = 0$; and*
(iv) *B has a proper, nonzero left (or right) ideal I.*

If $B \cong \mathrm{M}_2(F)$, we say that B is *split*. More generally, if $K \supset F$ is a field containing F, then we say K is a *splitting field* for B if $B_K = B \otimes_F K$ is split.

We give a proof of Lemma 4.1 in an algorithmically effective way in this section. The implication (i) ⇒ (ii) is clear. The implication (ii) ⇒ (iii) is obtained as follows.

Algorithm 4.2. Let $x \in B$ be a zerodivisor. This algorithm returns a nonzero element $e \in B$ such that $e^2 = 0$.

```
1. If $\operatorname{trd}(x) = 0$, return $x$.
2. Compute $0 \neq y \in B$ orthogonal to $1, x$ with respect to the quadratic form $\operatorname{nrd}$.
   If $xy = 0$, return $y$; otherwise, return $xy$.
```

Proof of correctness. The element $x \neq 0$ is a zerodivisor if and only if $\operatorname{nrd}(x) = x\overline{x} = 0$. Since y is orthogonal to 1 we have $\operatorname{trd}(y) = 0$ so $\overline{y} = -y$; similarly, since y is orthogonal to x we have $\operatorname{trd}(xy) = -\operatorname{trd}(x\overline{y}) = 0$. If $xy = 0$ then y is a zerodivisor. If $xy \neq 0$ then $\operatorname{nrd}(xy) = \operatorname{nrd}(x)\operatorname{nrd}(y) = 0$, as desired. □

The implication (iii) ⇒ (iv) follows, since e generates a proper left (or right) ideal. Below, in the proof of correctness of the following algorithm, we will show that if $I = Be$ then $\dim_F I = 2$; the final implication (iv) ⇒ (i) then follows since left multiplication gives a nonzero F-algebra map $B \to \operatorname{End}_F(I) \cong \mathrm{M}_2(F)$ which is injective since B is simple and therefore an isomorphism as $\dim_F B = 4 = \dim_F \mathrm{M}_2(F)$.

Algorithm 4.3. Let $e \in B$ satisfy $e^2 = 0$. This algorithm returns a standard representation $B \cong \left(\dfrac{1,1}{F}\right) \cong \mathrm{M}_2(F)$.

```
1. Find $k \in \{i, j, ij\}$ such that $\operatorname{trd}(ek) = s \neq 0$. Let $t = \operatorname{trd}(k)$ and $n = \operatorname{nrd}(k)$,
   and let $e' = (1/s)e$.
2. Let $j' = k + (-tk + n + 1)e'$ and let
     $i' = \begin{cases} e'k - (k+t)e', & \text{if } \operatorname{char} F \neq 2; \\ k + ((t+1)k + n + 1)e', & \text{if } \operatorname{char} F = 2. \end{cases}$
   Return $i', j'$.
```

Proof of correctness. In Step 1, if $\operatorname{trd}(ek) = 0$ for all such k then $e \in \operatorname{rad}(\operatorname{nrd})$, contradicting Lemma 3.17. We have $\operatorname{trd}(e'k) = \operatorname{trd}(ke') = 1$ so $\operatorname{trd}(\overline{e'k}) = -1$.

Consider $I = Fe' + Fke'$. Note $\operatorname{trd}(ke') \neq 0$ implies that e', ke' are linearly independent. Let A be the subalgebra of B generated by e' and k. We have $e'k + ke' = te' + 1$ from (3.3) and $k^2 = tk - n$, and thus we compute that left multiplication yields a map

$$A \to \operatorname{End}_F(I) \cong \mathrm{M}_2(F)$$
$$e', k \mapsto \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -n \\ 1 & t \end{pmatrix}.$$

A direct calculation then reveals that $j' \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $i' \mapsto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ if $\operatorname{char} F \neq 2$ and $i' \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ if $\operatorname{char} F = 2$, as in Example 1.12.

It follows all at once that $A = B$, that $I = Be'$, and that the map $B \to \mathrm{M}_2(F)$ is an isomorphism. □

Remark 4.4. An algorithm like the above which requires linear algebra in F is claimed but not exhibited explicitly by Rónyai [29]; see also further of Rónyai [32, Sect. 5.1].

Conics. We have already seen in Lemma 4.1 that $B \cong \mathrm{M}_2(F)$ if and only if there exists $0 \neq e \in B$ such that $e^2 = 0$. To this end, as in the previous section let

$$B_0 = \{e \in B : \operatorname{trd}(e) = 0\}.$$

We have $\dim_F B_0 = 3$, and given a standard representation for $B = \left(\dfrac{a, b}{F}\right)$, we have a basis for B_0 given by i, j, ij if $\operatorname{char} F \neq 2$ and $1, j, ij$ if $\operatorname{char} F = 2$, as in Example 3.5.

We may identify the set $\mathbb{P}(B_0) = B_0^\times / F^\times$ with the points of the projective plane $\mathbb{P}^2(F)$ over F. Then the equation $\operatorname{nrd}_0(x, y, z) = 0$ yields a *conic* $C \subset \mathbb{P}^2_F$ defined over F, a nonsingular projective plane curve of degree 2.

Lemma 4.5. *The following are equivalent:*

(i) $B \cong \mathrm{M}_2(F)$*;*
(v) *The quadratic form* $Q = \operatorname{nrd}|_{B_0}$ *associated to* B *represents zero over* F*; and*
(vi) *The conic* C *associated to* B *has an* F*-rational point.*

Therefore we are led to the following problems.

Problem 4.6 (HASPOINT). *Given a conic* C *defined over a field* F*, determine if* C *has an* F*-rational point.*

Problem 4.7 (EXHIBITPOINT). *Given a conic* C *defined over a field* F*, determine if* C *has an* F*-rational point and, if so, output such a point.*

These problems could be equivalently formulated as follows: given a nonsingular ternary quadratic form $Q : V \to F$, determine if F is *isotropic* (represents zero nontrivially) and, if so, find $0 \neq x \in V$ such that $Q(x) = 0$. We find the geometric language here to be more suggestive, but really these are equivalent ways to describe the same situation.

By Algorithm 3.12, given a conic C over F, there is a (deterministic, polynomial-time) algorithm which computes a change of coordinates in which C is given by the equation

$$ax^2 + by^2 + cz^2 = 0$$

if $\operatorname{char} F \neq 2$, with $a, b, c \in F^\times$, and

$$ax^2 + axy + aby^2 + cz^2 = 0$$

if $\operatorname{char} F = 2$, with $a, c \in F^\times$ and $b \in F$ by Example 3.7. In the first case, multiplying through by $abc \neq 0$ we obtain $bc(ax)^2 + ac(by)^2 + (abc^2)z^2 = 0$ which arises as the form associated to $\left(\dfrac{-bc, -ac}{F}\right)$; in the second case, we multiply through by $c \neq 0$ to obtain $(ac)x^2 + (ac)xy + b(ac)y^2 + (cz)^2 = 0$ which is associated to $\left(\dfrac{b, ac}{F}\right)$. Together with Algorithm 4.3, therefore, we arrive at the following lemma.

Proposition 4.8. *The association $B \mapsto C = \operatorname{nrd}_0$ gives a bijection between quaternion algebras over F up to isomorphism and conics over F up to isomorphism.*

Problems (IsMatrixRing), (ExhibitMatrixRing) *are (deterministic polynomial-time) equivalent to Problems* (HasPoint), (ExhibitPoint), *respectively.*

Proof. We need only identify isomorphisms: we need to show that two quaternion algebras $B \cong B'$ are isomorphic if and only if the induced conics $C \cong C'$ are isomorphic.

We treat only the case $\operatorname{char} F \neq 2$; the case $\operatorname{char} F = 2$ follows similarly. If $\phi : B \to B'$ is an isomorphism of quaternion algebras, then $\phi(1) = 1$ so $\phi(B_0) = B_0'$, and the reduced norm is determined by the standard involution which is unique, so $\operatorname{nrd}_B = \operatorname{nrd}_{B'} \circ \phi$.

Conversely, suppose $\psi : C \to C'$ is an isomorphism. Choose a quadratic form Q so that C is given by $Q = 0$ in $\mathbb{P}^2_F$, normalized and scaled so that $Q \cong \operatorname{nrd}_0$ for some $B \cong \left(\dfrac{a, b}{F}\right)$. Choose similarly Q' for C'. Then ψ is given by an element of $\mathrm{PGL}_3(F)$ and there exists a lift of ψ to $\mathrm{GL}_3(F)$ such that $Q = Q' \circ \psi$. The F-linear map $\psi : B_0 \to B_0'$ extends naturally (defining $\phi(1) = 1$) to an F-linear map which we also denote $\psi : B \to B'$, and we must show that ψ is an F-algebra isomorphism.

Suppose $B = \left(\dfrac{a, b}{F}\right)$. Then we have $\operatorname{nrd}(\psi(i)) = \operatorname{nrd}(i) = -a$ and $\operatorname{nrd}(\psi(i)) = \psi(i)\overline{\psi(i)} = -\psi(i)^2$ so $\psi(i)^2 = a$. Similarly we have $\psi(j)^2 = b$.

We have $ji = -ij$ since i, j are orthogonal, but then $\psi(i), \psi(j)$ are orthogonal so $\psi(j)\psi(i) = -\psi(i)\psi(j)$. Finally, we have that both $\psi(ij)$ and $\psi(i)\psi(j)$ are orthogonal to $1, \psi(i), \psi(j)$, and $\psi(ij)^2 = -ab = (\psi(i)\psi(j))^2$, so $\psi(ij) = \pm\psi(i)\psi(j)$. If the negative sign occurs, we replace ψ by the linear map defined on the basis $1, i, j, ij$ unmodified on $1, i, j$ but negated on ij; this map is now an F-algebra homomorphism. Together, these imply that $B' \cong \left(\dfrac{a,b}{F}\right)$ as well. □

We conclude this section by considering a simple case of the above problems. First, let $F = \mathbb{F}_q$ be a finite field with q elements. Indeed, Problem (HASPOINT) is trivial: since every conic over a finite field has a point (an elementary argument), one can simply always output TRUE!

For Problem (EXHIBITPOINT), we will make use of the following related problem.

Problem 4.9 (SQUAREROOT). *Given $a \in F^{\times 2}$, output $b \in F^{\times}$ such that $b^2 = a$.*

We have two cases. First, if q is even, then one can solve Problem (SQUAREROOT) in deterministic polynomial time (by repeated squaring, since $q - 1 = \#\mathbb{F}_{2^r}^{\times}$ is odd); for a conic in the form given in Example 3.5, given up to scaling by $x^2 + by^2 + byz + abz^2$ with $a, b \in \mathbb{F}_q$ and $b \neq 0$, this is already sufficient to solve Problem (EXHIBITPOINT). If q is odd, then there exists a deterministic polynomial-time algorithm to solve (EXHIBITPOINT) over $\mathbb{F}_q$ by work of van de Woestijne [37]. There also exists a probabilistic polynomial-time algorithm, which intersects the conic with a random line and then calls (SQUAREROOT), and there is a probabilistic polynomial-time algorithm to solve (SQUAREROOT) but no deterministic such algorithm (without further assumption of a generalized Riemann hypothesis). The latter algorithm is extremely efficient in practice.

Remark 4.10. It would also be interesting to study the corresponding problem where $\mathrm{M}_2(F)$ is replaced by another quaternion algebra B': in other words, to test if two quaternion algebras B, B' over F are isomorphic and, if so, to compute an explicit isomorphism. Since the reduced norm is determined by the standard involution on a quaternion algebra, and this involution is unique, it follows that if $B \cong B'$ then $\mathrm{nrd}_B \cong \mathrm{nrd}_{B'}$; in fact, this is an equivalence even when restricted to the trace zero subspace [21]. Therefore one is led to consider the problem of determining if two quadratic forms are isometric and, if so, to compute an explicit isometry.

Remark 4.11. More generally, one can establish a functorial bijection between twisted similarity classes of ternary quadratic forms over a commutative ring R and quaternion rings over R via the Clifford algebra; see work of the author [41]. It would be interesting to investigate the algorithmic implications of this correspondence.

5 Splitting Fields and the Hilbert Symbol

In this section, we exhibit algorithms for solving the Problem (IsMatrixRing) over a local field with residue characteristic not 2: in this setting, our problem is otherwise known as computing the Hilbert symbol.

Hilbert symbol. Let F be a field with $\operatorname{char} F \neq 2$, and let $a, b \in F^\times$. The *Hilbert symbol* is defined to be

$$(a,b)_F = \begin{cases} 1, & \text{if } \left(\dfrac{a,b}{F}\right) \cong \mathrm{M}_2(F); \\ -1, & \text{otherwise.} \end{cases}$$

We begin by recalling a well-known criterion [38, Corollaire 2.4].

Lemma 5.1. *A quaternion algebra* $\left(\dfrac{a,b}{F}\right)$ *is split if and only if* $b \in N_{K/F}(K^\times)$*, where* $K = F[i]$.

Here, we write $K = F[i] = F \oplus Fi$ to be the quadratic F-algebra generated by i.

Proof. If $\mathrm{N}_{K/F}(u+vi) = \operatorname{nrd}(u+vi) = b$ with $x, y \in F$, then $x = u + vi + j$ has $\operatorname{nrd}(x) = \operatorname{nrd}(u+vi+j) = \operatorname{nrd}(u+vi) + \operatorname{nrd}(j) = b - b = 0$, so B is not a division ring, so $B \cong \mathrm{M}_2(F)$ by Lemma 4.1. Conversely, if $B \xrightarrow{\sim} \mathrm{M}_2(F)$, then after conjugating by an element of $\mathrm{GL}_2(F)$ we may assume $i \mapsto \begin{pmatrix} 0 & a \\ 1 & 0 \end{pmatrix}$ (rational canonical form). The condition that $ji = -ij$ implies that $j \mapsto \begin{pmatrix} u & -av \\ v & -u \end{pmatrix}$ and $j^2 = u^2 - av^2 = b = \mathrm{N}_{K/F}(u+vi)$. □

Lemma 5.2. *We have* $(a,b)_F = (b,a)_F$ *and* $(a,b)_F = (-ab,b)_F$*. If* $u, v \in F^\times$ *then* $(a,b)_F = (au^2, bv^2)_F$*.*

Proof. Interchanging i, j gives an isomorphism $\left(\dfrac{a,b}{F}\right) \cong \left(\dfrac{b,a}{F}\right)$; replacing i, j by ui, vj gives an isomorphism $\left(\dfrac{a,b}{F}\right) \cong \left(\dfrac{u^2a, v^2b}{F}\right)$. By considering the algebra generated by ij, j we see that $\left(\dfrac{a,b}{F}\right) \cong \left(\dfrac{a,-ab}{F}\right)$. □

Local Hilbert symbol. For the rest of this section, let F be a number field. For a place v of F, let F_v denote the completion of F at v and let R_v be its valuation ring. Let π_v be a uniformizer for F_v and let k_v be the residue field of F_v.

If $a, b \in F_v^\times$, we abbreviate $(a,b)_v = (a,b)_{F_v}$. We now proceed to discuss the computability of $(a,b)_v$, and thereby Problem (IsMatrixRing) for local fields F_v with $\operatorname{char} k_v \neq 2$.

Remark 5.3. With Lemma 5.1 in mind, we recall the following facts about local norms. There is a unique unramified quadratic extension K_v of F_v, obtained from the corresonding unique such extension of residue fields. Then $\mathrm{N}_{K_v/F_v}(K_v^\times) = R_v^\times \times \pi_v^{2\mathbb{Z}}$ by Hensel's lemma, since the norm map in an extension of finite fields is surjective. For further details, see Neukirch [25, Corollary V.1.2] or Fröhlich [12, Proposition 7.3].

We begin by recalling the following fundamental result concerning division quaternion algebras over a local field [38, Théorèmes II.1.1, II.1.3].

Lemma 5.4. *Let v be a noncomplex place of F. Then there is a unique quaternion algebra B_v over F_v which is a division ring, up to F_v-algebra isomorphism.*

Note that there is no division quaternion algebra over $\mathbb{C}$ since $\mathbb{C}$ is algebraically closed. The unique division algebra over $\mathbb{R}$ is the classical ring of Hamiltonians $\mathbb{H} = \left(\dfrac{-1,-1}{\mathbb{R}}\right)$. If v is nonarchimedean, then the unique division ring over F_v is given by $B_v \cong \left(\dfrac{K_v, \pi_v}{F_v}\right)$, where K_v is the (unique) unramified quadratic extension of F_v.

Let B be a quaternion algebra over F. We say B is *unramified* (or *split*) at v if $B \otimes_F F_v \cong \mathrm{M}_2(F_v)$, i.e. F_v is a splitting field for B; otherwise (if B_v is a division ring) we say B is *ramified* at v.

A place v of F is *odd* if either v is real or v is nonarchimedean and $\#k_v$ is odd; v is *even* if v is nonarchimedean and $\#k_v$ is even. (A complex place is neither odd nor even.) For an odd place v and $a \in F_v^\times$, we define the *square symbol*

$$\left\{\frac{a}{v}\right\} = \begin{cases} 1, & \text{if } a \in F_v^{\times 2}; \\ -1, & \text{if } a \notin F_v^{\times 2} \text{ and } \mathrm{ord}_v(a) \text{ is even}; \\ 0, & \text{if } a \notin F_v^{\times 2} \text{ and } \mathrm{ord}_v(a) \text{ is odd}. \end{cases}$$

Here we set the convention that v is a real place then $\pi_v = -1$ is a uniformizer for $F_v \cong \mathbb{R}$ and that $a = (-1)^{\mathrm{ord}_v(a)}|a|$; in other words, $\left\{\dfrac{a}{v}\right\} = 1$ or 0 according as $a > 0$ or $a < 0$.

Suppose v is nonarchimedean. If $\mathrm{ord}_v(a) = 0$, then $\left\{\dfrac{a}{v}\right\} = \left(\dfrac{a}{v}\right)$ is the usual Legendre symbol (see (5.7) below); in fact, $\left\{\dfrac{a}{v}\right\} = 0$ if and only if $\mathrm{ord}_v(a)$ is odd. Note that the square symbol is not multiplicative, for example $\left\{\dfrac{\pi_v^2}{v}\right\} = 1 \neq 0 = \left\{\dfrac{\pi_v}{v}\right\}^2$; it is multiplicative when restricted to the subgroup of elements with even valuation, however.

Finally, we note that $\left\{\dfrac{a}{v}\right\} = -1$ if and only if $F_v(\sqrt{a})$ is an unramified field extension of F_v and $\left\{\dfrac{a}{v}\right\} = 0$ if and only if $F_v(\sqrt{a})$ is ramified; when v is real, we follow the convention that $\mathbb{C}$ is considered to be ramified over $\mathbb{R}$.

Proposition 5.5. *Let v be an odd place of F and let $a, b \in F_v^\times$. Then $(a,b)_v = 1$ if and only if*

$$\left\{\frac{a}{v}\right\} = 1 \ \textit{or} \ \left\{\frac{b}{v}\right\} = 1 \ \textit{or} \ \left\{\frac{-ab}{v}\right\} = 1 \ \textit{or} \ \left\{\frac{a}{v}\right\} = \left\{\frac{b}{v}\right\} = \left\{\frac{-ab}{v}\right\} = -1.$$

Proof. First, suppose v is archimedean. Then $(a,b)_v = 1$ if and only if $v(a) > 0$ or $v(b) > 0$ if and only if $\left\{\dfrac{a}{v}\right\} = 1$ or $\left\{\dfrac{b}{v}\right\} = 1$. So we suppose v is nonarchimedean.

Let $B_v = \left(\dfrac{a,b}{F_v}\right)$, and let $K_v = F_v[i]$, where we recall $i^2 = a$. Since $(a,b)_v = (b,a)_v = (a,-ab)_v$, the statement is symmetric in interchanging a, b and replacing b by $-ab$. If one of $\left\{\dfrac{a}{v}\right\} = 1$ or $\left\{\dfrac{b}{v}\right\} = 1$ or $\left\{\dfrac{-ab}{v}\right\} = 1$, then we may suppose $\left\{\dfrac{a}{v}\right\} = 1$; consequently, K_v is not a field, so B_v is not a division ring and by Lemma 4.1 we have $(a,b)_v = 1$. We cannot have $\left\{\dfrac{a}{v}\right\} = \left\{\dfrac{b}{v}\right\} = \left\{\dfrac{-ab}{v}\right\} = 0$. Thus we have only to consider the case $\left\{\dfrac{a}{v}\right\} = -1$.

If $\left\{\dfrac{b}{v}\right\} = -1$, then since K_v is the unique unramified quadratic extension of F_v and $\operatorname{ord}_v(b)$ is even, we have $b \in \mathrm{N}_{K_v/F_v}(K_v^\times)$ by Remark 5.3, so by Lemma 5.1 we have that B_v is split so $(a,b)_v = 1$. Otherwise, $\left\{\dfrac{b}{v}\right\} = 0$. But now $F_v[i] = K_v$ is the unramified quadratic extension of F_v so $b \notin \mathrm{N}_{K_v/F_v}(K_v^\times)$ and thus B_v is a division ring by Lemma 5.1, so $(a,b)_v = -1$. □

Corollary 5.6. *Let $a, b \in R_v \setminus \{0\}$ and suppose $a \in R_v^\times$. Then $(a,b)_v = \left(\dfrac{a}{v}\right)^{\operatorname{ord}_v b}$.*

Representing local fields. When discussing computability for local fields, we immediately encounter the following issue: a local field F_v is uncountable, so it is not computable.

One has at least two choices for overcoming this obstacle. One possibility is to use *exact local field arithmetic*, where one includes with the specification of an element its precision. One then requires the output of algorithms to be a continuous function of the input and to be correct with whatever output precision is given.

This way of working with $\mathbb{R}$ (or $\mathbb{C}$) also goes by the name *exact real* (or *complex*) *arithmetic*. This model has several advantages. In practice, for many applications this works extremely well: if more precision is required in the output, one simply gives more precision in the input. Consequently this model is also very efficient. Although this method does not realize a local field F as a computable field, all of the algorithms we discuss in this article work well in this model for F_v.

A second method is simply to work in a computable subfield F of the local field F_v. Indeed, any subfield F which is countably generated over its prime field is computable. In this article, we will take this approach; it is more appropriate for the theoretical discussion below (even as it will be less efficient in practice).

With this discussion in mind, we represent a local field as follows. First, let F be a number field. Let v be a place of F. If v is archimedean, then it is specified by some ordering of the roots of f in $\mathbb{C}$. If v is nonarchimedean, then v is specified by a prime ideal in the ring of integers in F. We can thereby compute a uniformizer $\pi_v \in F$ for the place v by the Chinese remainder theorem.

We then represent the local field as $F_v^{\mathrm{alg}} = \overline{F} \cap F_v$, an algebraic closure of F in F_v. Given a (monic) polynomial g with coefficients in F, there exists a deterministic algorithm which returns the roots of g in F_v (as elements of F_v^{alg}). In the nonarchimedean case, Hensel's lemma provides the essential ingredient to show that one can (efficiently) compute with F_v^{alg}. With this choice, by computing in the subfield generated by any element $x \in F_v^{\mathrm{alg}}$ we can compute the discrete valuation $\mathrm{ord}_v : F \to \mathbb{Z} \cup \{\infty\}$ as well as the reduction map $R_v \to k_v$ modulo π_v. When v is real, we recall that $\mathrm{ord}_v(a) = 0, 1$ according as $a > 0$ or $a < 0$, and so the computability of ord_v follows from well-known algorithms for exact real root finding.

The above discussion applies equally well to the case of global function fields; see Remark 1.5. For more on computably algebraically closed fields, we refer again to Stoltenberg-Hansen and Tucker [34].

Computing the local Hilbert symbol. To conclude, we discuss the computability of the Hilbert symbol for odd places using Proposition 5.5. We use Proposition 5.5 and the correspondence above to relate Problem (HASPOINT) to the problem of computing the square symbol.

Suppose F_v is archimedean. The Hilbert symbol for $F_v \cong \mathbb{C}$ is trivial. If v is real, then $\left\{\dfrac{a}{v}\right\} = 1, 0$ according as $a > 0$ or $a < 0$, so by the correspondence above this solves (HASPOINT) for these fields. It follows that Problem (EXHIBITPOINT) is equivalent to Problem (SQUAREROOT), and there is a deterministic algorithm to solve this problem in the computable subfield $F_v^{\mathrm{alg}} = \overline{F} \cap \mathbb{R}$ by hypothesis.

Next, suppose F_v is nonarchimedean and that v is odd. Then we can evaluate $\left\{\dfrac{a}{v}\right\}$ by simply computing $\mathrm{ord}_v(a) = e$; if e is odd then $\left\{\dfrac{a}{v}\right\} = 0$, whereas if e is even then $\left\{\dfrac{a}{v}\right\} = \left(\dfrac{a_0}{v}\right)$ where $a_0 = a\pi_v^{-e} \in R_v$ and $\left(\dfrac{a_0}{v}\right) = \left(\dfrac{a_0}{\mathfrak{p}}\right)$ is the usual Legendre symbol, defined by

$$\left(\frac{a_0}{\mathfrak{p}}\right) = \begin{cases} 0, & \text{if } a_0 \equiv 0 \pmod{\mathfrak{p}}; \\ 1, & \text{if } a_0 \not\equiv 0 \pmod{\mathfrak{p}} \text{ and } a_0 \text{ is a square modulo } \mathfrak{p}; \\ -1, & \text{otherwise.} \end{cases} \qquad (5.7)$$

The Legendre symbol can be computed in deterministic polynomial time by Euler's formula

$$\left(\frac{a_0}{\mathfrak{p}}\right) \equiv a_0^{(q-1)/2} \pmod{\mathfrak{p}}$$

using repeated squaring, where $q = \#k_v$.

To solve Problem (HASPOINT), by Proposition 5.5 we have two cases. In the first case, where one value of the square symbol is equal to 1, we reduce to Problem (SQUAREROOT) over F_v^{alg} which we can solve by the above. Otherwise, if all three symbols in Proposition 5.5 are -1, then also by Hensel's lemma, Problem (EXHIBITPOINT) over F_v^{alg} is reducible to Problem (EXHIBITPOINT) over k_v, which was discussed at the end of the previous section.

If we restrict our input to a global field F, then a runtime analysis of the above method yields the following.

Proposition 5.8. *Let F be a number field and let v be an odd place of F. Then there exists a deterministic polynomial-time algorithm to evaluate the Hilbert symbol $(a, b)_v$ for $a, b \in F^\times$.*

Remark 5.9. By *Hilbert reciprocity*, we have

$$\prod_v (a, b)_v = 1 \qquad (5.10)$$

whenever F is a global field and $a, b \in F^\times$. Consequently, if one can compute all but one local Hilbert symbol $(a, b)_v$, then the final symbol can be recovered from the above relation. In particular, this means for a number field F, if there exists a unique prime above 2 (e.g. when $F = \mathbb{Q}$) then one can evaluate $(a, b)_2$ in this way.

6 The Even Local Hilbert Symbol

In this section, we discuss the computation of the local Hilbert symbol for an even place of a number field F. The main result of this section is the following theorem.

Theorem 6.1. *Let F be a number field and let v be a place of F. Then there exists a deterministic polynomial-time algorithm to evaluate the Hilbert symbol $(a, b)_v$ for $a, b \in F^\times$.*

If v is complex, this theorem is trivial; if v is an odd place of F then Theorem 6.1 follows from Proposition 5.8. So suppose that v is an even place of F, i.e. $\#k_v$ is even. Let $\mathbb{Z}_F$ be the ring of integers of F and let $\mathfrak{p}$ be the prime of $\mathbb{Z}_F$ corresponding to v.

We first give an algorithm which gives a solution to an integral norm form via a Hensel-like lift.

Algorithm 6.2. Let $\mathfrak{p}$ an even prime with ramification index $e = \operatorname{ord}_{\mathfrak{p}} 2$, and let $a, b \in F$ be such that $\operatorname{ord}_{\mathfrak{p}}(a) = 0$ and $\operatorname{ord}_{\mathfrak{p}}(b) = 1$. This algorithm outputs a solution to the congruence

$$1 - ay^2 - bz^2 \equiv 0 \pmod{\mathfrak{p}^{2e}}$$

with $y, z \in \mathbb{Z}_F/\mathfrak{p}^{2e}$ and $y \in (\mathbb{Z}_F/\mathfrak{p})^\times$.

1. Let $f \in \mathbb{Z}_{\geq 1}$ be the residue class degree of $\mathfrak{p}$ (so that $\#(\mathbb{Z}_F/\mathfrak{p}) = 2^f$) and let $q = 2^f$. Let π be a uniformizer at $\mathfrak{p}$.
2. Initialize $(y, z) := (1/\sqrt{a}, 0)$.
3. Let $N := 1 - ay^2 - bz^2 \in \mathbb{Z}_F/4\mathbb{Z}_F$ and let $t := \operatorname{ord}_{\mathfrak{p}}(N)$. If $t \geq 2e$, return y, z. Otherwise, if t is even, let

   $$y := y + \sqrt{\frac{N}{a\pi^t}}\pi^{t/2}$$

   and if t is odd, let

   $$z := z + \sqrt{\frac{N}{b\pi^{t-1}}}\pi^{\lfloor t/2 \rfloor}.$$

   Return to Step 3.

In this algorithm, when we write $\sqrt{u}$ for $u \in (\mathbb{Z}_F/\mathfrak{p}^{2e})^\times$ we mean any choice of a lift of $\sqrt{u} \in (\mathbb{Z}_F/\mathfrak{p})^\times$ to $\mathbb{Z}_F/\mathfrak{p}^{2e}$.

Proof of correctness. The key calculation in Step 3 is as follows: if t is even, we make the substitution

$$1 - a(y + u\pi^{t/2})^2 - bz^2 = N - 2au\pi^{t/2}y - au^2\pi^t \equiv 0 \pmod{\mathfrak{p}^{t+1}}$$

and solve for u. Note that since $t < 2e$ we have $\operatorname{ord}_{\mathfrak{p}}(2\pi^{t/2}) = e + t/2 \geq t + 1$; solving we get $u^2 \equiv N/(a\pi^t) \pmod{\mathfrak{p}}$ as claimed. The case where t is odd is similar: we have

$$\begin{aligned}
1 - ay^2 - b(z + \sqrt{N/b\pi^{t-1}}\pi^{\lfloor t/2 \rfloor})^2 \\
&= N - 2bz\sqrt{N/b\pi^{t-1}}\pi^{\lfloor t/2 \rfloor} - b(N/b\pi^{t-1})\pi^{t-1} \\
&\equiv N - N \equiv 0 \pmod{\mathfrak{p}^{t+1}}
\end{aligned}$$

and the middle term above vanishes modulo $\mathfrak{p}^{t+1}$ since $t < 2e$ implies $e + 1 + \lfloor t/2 \rfloor = e + 1 + (t-1)/2 \geq t + 1$. □

Remark 6.3. Alternatively, we can compute a solution modulo 2 directly. The map

$$\begin{aligned}(\mathbb{Z}_F/\mathfrak{p}^e)^2 &\to \mathbb{Z}_F/2\mathbb{Z}_F \\ (y,z) &\mapsto 1 - ay^q - bz^q\end{aligned}$$

is $\mathbb{Z}_F/\mathfrak{p} \cong \mathbb{F}_q$-linear since $2 \equiv 0 \pmod{\mathfrak{p}^e}$. Let (y_0, z_0) be in the kernel of this map. Letting $(x,y,z) := (1, y_0^{q/2}, z_0^{q/2})$, we see $1 - ay^2 - bz^2 \equiv 0 \pmod 2$.

Remark 6.4. This is better than the algorithm provided in Simon's thesis [35] because we do not need to make a brute force search, which might not run in polynomial time.

We reduce to the above Hensel lift by the following algorithm.

Algorithm 6.5. Let $\mathfrak{p}$ an even prime with ramification index $e = \operatorname{ord}_{\mathfrak{p}} 2$ and let $a, b \in F^\times$ be such that $v(a) = 0$ and $v(b) \in \{0, 1\}$. This algorithm outputs $y, z, w \in \mathbb{Z}_F/\mathfrak{p}^{2e}$ such that

$$1 - ay^2 - bz^2 + abw^2 \equiv 0 \pmod{\mathfrak{p}^{2e}}$$

and $y \in (\mathbb{Z}_F/\mathfrak{p})^\times$. Let π be a uniformizer for $\mathfrak{p}$.

1. If $v(b) = 1$, return the output $(y, z, 0)$ of Algorithm 6.2 with input a, b.
2. Suppose $a \in (\mathbb{Z}_F/\mathfrak{p}^e\mathbb{Z}_F)^{\times 2}$ and $b \in (\mathbb{Z}_F/\mathfrak{p}^e\mathbb{Z}_F)^{\times 2}$. Let $(a_0)^2 a \equiv 1 \pmod{\mathfrak{p}^e}$ and $(b_0)^2 b \equiv 1 \pmod{\mathfrak{p}^e}$. Return
$$y := a_0,\ z := b_0,\ w := a_0 b_0.$$
3. Swap a, b if necessary so that $a \in (\mathbb{Z}_F/\mathfrak{p}^e\mathbb{Z}_F)^\times \setminus (\mathbb{Z}_F/\mathfrak{p}^e\mathbb{Z}_F)^{\times 2}$. Let t be the largest integer such that $a \in (\mathbb{Z}_F/\mathfrak{p}^t)^{\times 2}$ but $a \notin (\mathbb{Z}_F/\mathfrak{p}^e)^{\times 2}$. Then t is odd; write $a = a_0^2 + \pi^t a_t$ with $a_0, a_t \in \mathbb{Z}_F$. Let y, z be the output of Algorithm 6.2 with input $a' := a$, $b' := -\pi a_t / b$. Return
$$y' := \frac{1}{a_0},\ z' := \frac{\pi^{\lfloor t/2 \rfloor}}{a_0 z},\ w' := \frac{y\pi^{\lfloor t/2 \rfloor}}{a_0 z}$$
(reswapping if necessary).

Proof of correctness. In Step 2, writing $aa_0^2 = 1 + 2a'$ and $bb_0^2 = 1 + 2b'$ with $a', b' \in \mathbb{Z}_F$ we indeed have

$$\begin{aligned}&1 - a(a_0)^2 - b(b_0)^2 + ab(a_0 b_0)^2 \\ &\quad = 1 - (1 + 2a') - (1 + 2b') + (1 + 2a')(1 + 2b') \equiv 0 \pmod{\mathfrak{p}^{2e}}\end{aligned}$$

since $4 \in \mathfrak{p}^{2e}$.

Now we discuss Step 3. Write $a = a_0 + a_1\pi + \cdots + a_{e-1}\pi^{e-1}$ with $a_i \in \mathbb{Z}_F/\mathfrak{p}$. Then indeed $a \in (\mathbb{Z}_F/\mathfrak{p}^e)^{\times 2}$ if and only if and $a_i = 0$ for i odd by the freshperson's dream, so in particular $t < e$ is odd. Now suppose from Algorithm 6.2 we have

$$1 - ay^2 + (\pi a_t/b)z^2 \equiv 0 \pmod{\mathfrak{p}^{2e}}.$$

Note $\operatorname{ord}_{\mathfrak{p}}(z) \le \lfloor t/2 \rfloor = (t-1)/2$ since otherwise $a \in (\mathbb{Z}_F/\mathfrak{p}^{t+1})^{\times 2}$, a contradiction. Multiplying by $-b\pi^{t-1}/z^2 = -b(\pi^{\lfloor t/2 \rfloor}/z)^2$ gives

$$-b(\pi^{\lfloor t/2 \rfloor}/z)^2 + ab(y\pi^{\lfloor t/2 \rfloor}/z)^2 - \pi^t a_t \equiv 0 \pmod{\mathfrak{p}^{2e}}$$

so

$$a_0^2 - (a_0^2 + \pi^t a_t) - b(\pi^{\lfloor t/2 \rfloor}/z)^2 + ab(y\pi^{\lfloor t/2 \rfloor}/z)^2 \equiv 0 \pmod{\mathfrak{p}^{2e}}$$

so since $a = a_0^2 + \pi^t a_t$, dividing by a_0^2 we have the result. □

We say that $\pi^{-1} \in F$ is an *inverse uniformizer* for $\mathfrak{p}$ if $\operatorname{ord}_{\mathfrak{p}}(\pi^{-1}) = -1$ and $\operatorname{ord}_{\mathfrak{q}}(\pi^{-1}) \ge 0$ for all $\mathfrak{q} \ne \mathfrak{p}$.

We are now prepared to evaluate the even Hilbert symbol.

Algorithm 6.6. Let $B = \left(\dfrac{a,b}{F}\right)$ be a quaternion algebra with $a, b \in F^\times$, and let $\mathfrak{p}$ be an even prime of F. This algorithm returns the value of the Hilbert symbol $(a,b)_{\mathfrak{p}}$.

1. Scale a, b if necessary by an element of $\mathbb{Q}^{\times 2} \cap \mathbb{Z}$ so that $a, b \in \mathbb{Z}_F$.
2. Let π^{-1} be an inverse uniformizer for $\mathfrak{p}$. Let $a := (\pi^{-1})^{2\lfloor \operatorname{ord}_{\mathfrak{p}}(a)/2 \rfloor} a$ and $b := (\pi^{-1})^{2\lfloor \operatorname{ord}_{\mathfrak{p}}(b)/2 \rfloor} b$. If $\operatorname{ord}_{\mathfrak{p}} a = \operatorname{ord}_{\mathfrak{p}} b = 1$, let $a := (\pi^{-1})^2(-ab)$. Swap if necessary so that $\operatorname{ord}_{\mathfrak{p}} a = 0$.
3. Call Algorithm 6.5, and let $i' := (1 + yi + zj + wij)/2$. Let $f(T) = T^2 - T + \operatorname{nrd}(i')$ be the minimal polynomial of i'. If f has a root modulo $\mathfrak{p}$, return 1.
4. Let $j' := (zb)i - (ya)j$ and let $b' := (j')^2$. If $\operatorname{ord}_v b'$ is even, return 1, otherwise return -1.

Proof of correctness. If in Step 2 we have a root modulo $\mathfrak{p}$, then by Hensel's lemma, f has a root $t \in F_{\mathfrak{p}}$, hence $t - i'$ is a zero divisor and we return 1 correctly. Otherwise, by Lemma 5.4, we have $K_{\mathfrak{p}} = F_{\mathfrak{p}}[i']$ is the unramified field extension of $F_{\mathfrak{p}}$. We compute that $\operatorname{trd}(j') = \operatorname{trd}(i'j') = 0$, so $B_{\mathfrak{p}} \cong \left(\dfrac{K_{\mathfrak{p}}, b'}{F_{\mathfrak{p}}}\right)$ and $B_{\mathfrak{p}}$ is split if and only if $\operatorname{ord}_{\mathfrak{p}} b'$ is even. □

Note that the above algorithms run in deterministic polynomial time.

Example 6.7. Let $F = \mathbb{Q}(u)$ where $u = \sqrt[8]{500}$. Then $2\mathbb{Z}_F = (2, \sqrt[8]{500})^4 = \mathfrak{p}^4$, so $\mathbb{Z}_{F,\mathfrak{p}}$ is a ramified extension of $\mathbb{Z}_2$ of residue degree 2 and ramification degree $e = 4$. Using Algorithm 6.6, we compute $(a,b)_{\mathfrak{p}}$ where $b = u^2 + 40$ and $a = u^2 + u + 1$.

In Step 2, we compute the inverse uniformizer $\pi^{-1} = u^3/10$ satisfying the polynomial $T^8 - 5/4$. We compute $\mathrm{ord}_{\mathfrak{p}}(a) = 0$ and $\mathrm{ord}_{\mathfrak{p}}(b) = 2$. So we let $b := (\pi^{-1})^2 b = \frac{1}{5}(2u^6 + 25)$ with now $\mathrm{ord}_{\mathfrak{p}}(b) = 0$.

In Step 3, we call Algorithm 6.5. We use the uniformizer $\pi = u$. We compute that $b \equiv 1 \pmod{\mathfrak{p}^e}$ so $b \in (\mathbb{Z}_F/\mathfrak{p}^e\mathbb{Z}_F)^{\times 2}$ but $a \equiv 1 + \pi + \pi^2 \pmod{\mathfrak{p}^e}$. So we write $a = a_0 + \pi^t a_t$ with $a_0 = 1$ and $a_t = u + 1$.

We then call Algorithm 6.2 with input $a' := a$ and $b' := -\pi a_t/b$. We initialize $(y, z) = (1, 0)$. In Step 3 of this algorithm, we have $N := 1 - (1 + u + u^2) = -(u + u^2)$ with valuation $t := 1$. We let $z := \sqrt{N/b} = 1$ and return; now $N := 1 - ay^2 - bz^2$ has valuation $t := 9 > 2e$, so we exit the loop with output $y = z = 1$.

We then exit Algorithm 6.5 with $y' := 1/a_0 = 1$, $z' := \pi^{\lfloor t/2 \rfloor}/(a_0 z) = 1$, and $w' := y\pi^{\lfloor t/2 \rfloor}/(a_0 z) = 1$. We verify that $1 - a(y')^2 - b(z')^2 + ab(w')^2 = 1 - a - b + ab \equiv 0 \pmod{4}$.

Returning to Algorithm 6.6, we let $i' := (1+i+j+ij)/2$ and compute $\mathrm{nrd}(i') = 1/10(w^7 + 10w^2 + 10w + 500) \equiv 0 \pmod{\mathfrak{p}}$, so $f(T) = T^2 - T + \mathrm{nrd}(i')$ has a root modulo $\mathfrak{p}$, and we return $(a, b)_{\mathfrak{p}} = 1$.

Computing the Jacobi symbol. An interesting consequence of the above algorithm is that one can evaluate the Jacobi symbol in deterministic polynomial time in certain cases analogous to the way ("reduce and flip") that one computes this symbol using quadratic reciprocity in the case $F = \mathbb{Q}$. (See Lenstra [23] for an alternative approach which works in greater generality.)

We extend the definition of the Legendre symbol (5.7) to a symbol $\left(\dfrac{a}{\mathfrak{b}}\right)$ with $\mathfrak{b}$ odd by multiplicativity, and we define $\left(\dfrac{a}{b}\right) = \left(\dfrac{a}{b\mathbb{Z}_F}\right)$.

We write $v \mid 2\infty$ for the set of finite even places and real archimedean places of F.

Proposition 6.8. *Let $a, b \in \mathbb{Z}_F$ satisfy $a\mathbb{Z}_F + b\mathbb{Z}_F = \mathbb{Z}_F$, with b odd, and suppose $a = a_0 a_1$ with a_1 odd. Then*

$$\left(\frac{a}{b}\right)\left(\frac{b}{a_1}\right) = \prod_{v \mid 2\infty} (a, b)_v.$$

Proof. By Hilbert reciprocity (5.10), we have

$$\prod_v (a, b)_v = 1 = \prod_{v \mid 2\infty} (a, b)_v \prod_{\mathfrak{p} \nmid 2} (a, b)_{\mathfrak{p}}.$$

By Lemma 5.5, if $\mathfrak{p}$ is odd and $\mathrm{ord}_{\mathfrak{p}}(a) = \mathrm{ord}_{\mathfrak{p}}(b) = 0$ then $(a, b)_{\mathfrak{p}} = 1$. Therefore

$$\prod_{\mathfrak{p} \mid a_1 b} (a, b)_{\mathfrak{p}} = \prod_{v \mid 2\infty} (a, b)_v.$$

For $\mathfrak{p}$ odd, if $\operatorname{ord}_{\mathfrak{p}} a_1 > 0$ then $\operatorname{ord}_{\mathfrak{p}} b = 0$ by assumption and thus

$$(a,b)_{\mathfrak{p}} = \left(\frac{b}{\mathfrak{p}}\right)^{\operatorname{ord}_{\mathfrak{p}} a} = \left(\frac{b}{\mathfrak{p}}\right)^{\operatorname{ord}_{\mathfrak{p}} a_1}.$$

Similarly if $\operatorname{ord}_{\mathfrak{p}} b > 0$ then $(a,b)_{\mathfrak{p}} = \left(\frac{a}{\mathfrak{p}}\right)^{\operatorname{ord}_{\mathfrak{p}} b}$, hence

$$\prod_{\mathfrak{p} \mid a_1 b} (a,b)_{\mathfrak{p}} = \left(\frac{a}{b}\right)\left(\frac{b}{a_1}\right).$$

The result follows. □

A *Euclidean function* on F is a map $N : \mathbb{Z}_F \setminus \{0\} \to \mathbb{Z}_{\geq 0}$ such that for all $a,b \in \mathbb{Z}_F$ we have $N(ab) = N(a)N(b)$ and there exists $q,r \in \mathbb{Z}_F$ such that $a = qb + r$ with either $r = 0$ or $N(r) < N(b)$. A Euclidean function is *computable* if given a,b, the elements q,r as above are computable.

Algorithm 6.9. Let F be a number field with a computable Euclidean function N and let $a,b \in \mathbb{Z}_F \setminus \{0\}$. This algorithm returns the Jacobi symbol $\left(\frac{a}{b}\right)$.

```
1. Initialize $z = 1$.
2. If $b\mathbb{Z}_F = \mathbb{Z}_F$, return $z$. Otherwise, compute $q,r \in \mathbb{Z}_F$ such that $a = qb + r$. If
   $r = 0$, return 0. Let $a := r$. Write $a = a_0 a_1$ with $a_1 \in \mathbb{Z}_F$ odd.
3. Multiply $z$ by $\prod_{v \mid 2,\infty} (a,b)_v$, computed using Algorithm 6.6. Return to Step 2,
   with $(a,b) = (b, a_1)$.
```

Proof of correctness. The division algorithm associated to N implies that $\mathbb{Z}_F$ has unique factorization, so we can indeed write $a = a_0 a_1$ with a_1 odd. The algorithm terminates because in Step 4 we have $N(a_1) \leq N(a) = N(r) < N(b)$. □

Remark 6.10. For any fixed F, one can precompute a table of the values $(a,b)_{\mathfrak{p}}$ for a,b in appropriate residue classes modulo an even prime $\mathfrak{p}$; this is what is usually done for $F = \mathbb{Q}$, for example.

Relationship to conics. In view of the results in Sect. 4, we now relate the above algorithms to the geometric problem of rational points on conics.

Theorem 6.11 (Hasse-Minkowski). *A quaternion algebra B has $B \cong \mathrm{M}_2(F)$ if and only if B is unramified at all places.*

Equivalently, a conic C has $C(F) \neq \emptyset$ if and only if $C(F_v) \neq \emptyset$ for all places v of F. For a proof of the Hasse-Minkowski Theorem, see Lam [21], O'Meara [26], or Vignéras [38, Sect. III.3.1]

Proposition 6.12. *Problem* (IsMatrixRing) *is deterministic polynomial-time reducible to the problem of factoring ideals in $\mathbb{Z}_F$.*

Proof. Given a quaternion algebra $B = \left(\dfrac{a,b}{F}\right)$, we have $B_v \cong \mathrm{M}_2(F_v)$ for all $v \nmid 2ab\infty$, and by factoring by the above algorithms for each $v \mid 2ab\infty$ we check if $B_v \cong \mathrm{M}_2(F_v)$ by computing the Hilbert symbol $(a,b)_v$ in deterministic polynomial time. □

7 Maximal Orders

In this section, we consider some integral versions (for orders) of the above algorithms relating quadratic forms and quaternion algebras. Our main result relates identifying the matrix ring to computing a maximal order. Throughout this section, let F be a number field, let $\mathbb{Z}_F$ be its ring of integers, and let $\mathcal{O}$ be a ($\mathbb{Z}_F$-)order in a quaternion algebra B over F. For further reading, see Reiner [28] or Vignéras [38].

Computing maximal orders, generally. There exists a deterministic algorithm to compute the ring of integers $\mathbb{Z}_F$ (see Cohen [6, Sect. 6.1], [7, Algorithm 2.4.9]): in fact, computing $\mathbb{Z}_F$ is deterministic polynomial-time equivalent to the problem of finding the largest square divisor of a positive integer [5, 22]; no polynomial-time algorithm is known for this problem (though see work of Buchmann and Lenstra [4] for a way of "approximating" $\mathbb{Z}_F$).

Example 7.1. If $F = \mathbb{Q}(\sqrt{D})$, then $R = \mathbb{Z} \oplus \mathbb{Z}(d + \sqrt{d})/2$ where $D = df^2$ and f^2 is the largest square divisor of D subject to the requirement that $d \equiv 0, 1 \pmod{4}$.

We consider in this section the noncommutative analogues of this problem. We have the following general result due to Ivanyos and Rónyai [16, Theorem 5.3], which was rediscovered by Nebe and Steel [24]; see also Friedrichs [11].

Theorem 7.2. *There exists an explicit algorithm which, given a semisimple F-algebra B, computes a maximal order $\mathcal{O} \subset B$. This algorithm runs in deterministic polynomial time given oracles for the problems of factoring integers and factoring polynomials over finite fields.*

At present, it is not known if there exist deterministic polynomial-time algorithms to solve either of these latter two problems. Indeed, we have already noted that computing a maximal order in F is as hard as computing the largest squarefree divisor of a positive integer; therefore, it is reasonable to expect that the problem for a noncommutative algebra B is no less complicated. (See a more precise characterization of this complexity at the end of this section.)

We do not discuss the algorithm exhibited in Theorem 7.2; rather, we consider the special case of quaternion algebras, and by manipulations with quadratic forms we obtain a simpler algorithm.

Discriminants. We begin by analyzing the following problem.

Problem 7.3 (IsMaximal). *Given an order $\mathcal{O} \subset B$, determine if $\mathcal{O}$ is a maximal order.*

This problem has a very simple solution as follows. The *discriminant* $\mathfrak{D}(B)$ of B is the ideal equal to the product of all primes of $\mathbb{Z}_F$ where B is ramified:

$$\mathfrak{D}(B) = \prod_{\mathfrak{p} \text{ ramified}} \mathfrak{p}.$$

On the other hand, the *discriminant* $\operatorname{disc}(\mathcal{O})$ of an order $\mathcal{O} \subset B$ is the ideal generated by the set

$$\{\det(\operatorname{trd}(x_i x_j))_{i,j=1,\dots,4} : x_1, \dots, x_4 \in \mathcal{O}\}.$$

The discriminant $\operatorname{disc}(\mathcal{O})$ is the square of an ideal in $\mathbb{Z}_F$, and the *reduced discriminant* $\mathfrak{d}(\mathcal{O})$ of $\mathcal{O}$ is the ideal satisfying $\mathfrak{d}(\mathcal{O})^2 = \operatorname{disc}(\mathcal{O})$.

Given a pseudobasis $(\mathfrak{a}_i, x_i)$ for $\mathcal{O}$ we have

$$\operatorname{disc}(\mathcal{O}) = (\mathfrak{a}_1 \cdots \mathfrak{a}_4)^2 \det(\operatorname{trd}(x_i x_j))_{i,j=1,\dots,4}.$$

Remark 7.4. Although we will not use this in the sequel, the reduced discriminant can in fact be computed more simply: if $\mathcal{O} = \mathbb{Z}_F \oplus \mathfrak{a}i \oplus \mathfrak{b}j \oplus \mathfrak{c}k$ then

$$\mathfrak{d}(\mathcal{O}) = \mathfrak{a}\mathfrak{b}\mathfrak{c} \operatorname{trd}((ij - ji)\overline{k}).$$

Lemma 7.5. *An order $\mathcal{O} \subset B$ is maximal if and only if $\mathfrak{d}(\mathcal{O}) = \mathfrak{D}(B)$.*

Proof. We give only a sketch of the proof. For a prime $\mathfrak{p}$ of $\mathbb{Z}_F$, let $\mathbb{Z}_{F,\mathfrak{p}}$ be the completion of $\mathbb{Z}_F$ at $\mathfrak{p}$ and $F_\mathfrak{p}$ the completion of F at $\mathfrak{p}$; write $\mathcal{O}_\mathfrak{p} = \mathcal{O} \otimes_{\mathbb{Z}_F} \mathbb{Z}_{F,\mathfrak{p}}$ and similarly $B_\mathfrak{p} = B \otimes_F F_\mathfrak{p}$.

We have $\mathfrak{d}(\mathcal{O}) = \mathfrak{D}(B)$ if and only if $\mathfrak{d}(\mathcal{O})_\mathfrak{p} = \mathfrak{d}(\mathcal{O}_\mathfrak{p}) = \mathfrak{D}(B_\mathfrak{p}) = \mathfrak{D}(B)_\mathfrak{p}$ for all primes $\mathfrak{p}$, and the order $\mathcal{O}$ is maximal if and only if $\mathcal{O}_\mathfrak{p}$ is maximal for every prime $\mathfrak{p}$ of $\mathbb{Z}_F$ (see [28, 11.2]). So it suffices to note that if $\mathfrak{p}$ is unramified then any maximal order of $B_\mathfrak{p}$ has discriminant $\mathbb{Z}_{F,\mathfrak{p}}$ and if $\mathfrak{p}$ is ramified then the unique maximal order of $B_\mathfrak{p}$ has reduced discriminant $\mathfrak{p}\mathbb{Z}_{F,\mathfrak{p}}$ [28, Theorem 14.9]. □

Putting these together with the computation of the local Hilbert symbol, we have shown that one can solve Problem (IsMaximal) in deterministic polynomial time given an oracle to factor integers and polynomials over finite fields, since this allows the factorization of the discriminant $\mathfrak{D}(B)$ [6, Proposition 6.2.8, Algorithm 6.2.9]; note that this need only be done once for a quaternion algebra B.

Computing maximal orders. We now turn to the problem of computing a maximal order in a quaternion algebra.

Problem 7.6 (AlgebraMaxOrder). *Given a quaternion algebra B over F, compute a maximal order $\mathcal{O} \subset B$.*

A more general problem is as follows.

Problem 7.7 (MAXORDER). *Given an order $\Lambda \subset B$ in a quaternion algebra B over F, compute a maximal order $\mathcal{O} \supset \Lambda$.*

One immediately reduces from the former to the latter by exhibiting any order in B, as follows. (First, we compute $\mathbb{Z}_F$ as above; this can be considered a precomputation step if F is fixed.) If $B = \left(\dfrac{a,b}{F}\right)$, we may scale a, b by a nonzero square integer so that $a, b \in \mathbb{Z}_F$, and then

$$\Lambda = \mathbb{Z}_F \oplus \mathbb{Z}_F i \oplus \mathbb{Z}_F j \oplus \mathbb{Z}_F ij \tag{7.8}$$

is an order, where i, j are the standard generators for B.

An order $\mathcal{O}$ is $\mathfrak{p}$-*maximal* for a prime $\mathfrak{p}$ if $\mathcal{O}_\mathfrak{p} = \mathcal{O} \otimes_{\mathbb{Z}_F} \mathbb{Z}_{F,\mathfrak{p}}$ is maximal (as an $\mathbb{Z}_{F,\mathfrak{p}}$-order). Note that if $\operatorname{ord}_\mathfrak{p}(\mathfrak{d}(\mathcal{O}_\mathfrak{p})) = 0$ then necessarily $\mathcal{O}$ is $\mathfrak{p}$-maximal. To solve Problem (MAXORDER), we recursively compute a $\mathfrak{p}$-maximal order for every prime $\mathfrak{p} \mid \mathfrak{d}(\mathcal{O})$, proceeding in two steps.

We say an order $\mathcal{O}$ is $\mathfrak{p}$-*saturated* if $\operatorname{nrd}|_{\mathcal{O}_\mathfrak{p}}$ has a normalized basis $1, i, j, k$ (see Proposition 3.10) such that each atomic block has valuation at most 1; we then say that $1, i, j, k$ is a $\mathfrak{p}$-*saturated* basis for $\mathcal{O}$.

We compute a $\mathfrak{p}$-saturated order in the following straightforward way. Recall that $\pi^{-1} \in F$ is an *inverse uniformizer* for $\mathfrak{p}$ if $\operatorname{ord}_\mathfrak{p}(\pi^{-1}) = -1$ and $\operatorname{ord}_\mathfrak{q}(\pi^{-1}) \geq 0$ for all $\mathfrak{q} \neq \mathfrak{p}$.

Algorithm 7.9. Let

$$\Lambda = \mathbb{Z}_F \oplus \mathfrak{a}i \oplus \mathfrak{b}j \oplus \mathfrak{c}k \subset B$$

be an order and let $\mathfrak{p}$ be prime. This algorithm computes a $\mathfrak{p}$-saturated order $\mathcal{O} \supset \Lambda$ and a $\mathfrak{p}$-saturated basis for $\mathcal{O}$.

1. Choose $d \in \mathfrak{a}$ such that $\operatorname{ord}_\mathfrak{p}(d) = \operatorname{ord}_\mathfrak{p}(\mathfrak{a})$ and let $i := di$; compute similarly with j, k. Let $\mathcal{O} := \Lambda$.
2. Run Algorithm 3.12 over the localization of $\mathbb{Z}_F$ at $\mathfrak{p}$ with input the quadratic form $\operatorname{nrd}|_\mathcal{O}$ and the basis $1, i, j, k$; let $1, i^*, j^*, k^*$ be the output. Let $c \in \mathbb{Z}_F$ be such that $\operatorname{ord}_\mathfrak{p} c = 0$ and such that $ci^* \in \mathcal{O}$, and let $i := ci^*$; compute similarly with j, k.
3. Let π^{-1} be an inverse uniformizer for $\mathfrak{p}$. For each atomic form Q in $\operatorname{nrd}_\mathcal{O}$, let e be the valuation of Q, and multiply each basis element in Q by $(\pi^{-1})^{\lfloor e/2 \rfloor}$. Return $\mathcal{O} := \Lambda + (\mathbb{Z}_F i \oplus \mathbb{Z}_F j \oplus \mathbb{Z}_F k)$ and the basis $1, i, j, k$.

Proof of correctness. In Step 3, we are asserting that the output of Algorithm 3.12 leaves 1 as the first basis element. Indeed, we note that $\operatorname{ord}_\mathfrak{p} \operatorname{trd}(j) \leq \operatorname{ord}_\mathfrak{p} \operatorname{trd}(i(ij))$ since $\operatorname{trd}(i(ij)) = \operatorname{trd}(i)^2 - \operatorname{trd}(j)\operatorname{nrd}(i)$ and similarly $\operatorname{ord}_\mathfrak{p} \operatorname{trd}(i) \leq \operatorname{ord}_\mathfrak{p} \operatorname{trd}((ij)j)$.

Let $1, i, j, k$ be the basis computed in Step 3. By definition, this basis is $\mathfrak{p}$-saturated; we need to show that $\mathcal{O}$ is indeed an order. But $\mathcal{O}$ is an order if and only if $\mathcal{O}_\mathfrak{q}$ is an order for all primes $\mathfrak{q}$, and we have $\mathcal{O}_\mathfrak{q} = \Lambda_\mathfrak{q}$ for all primes $\mathfrak{q} \neq \mathfrak{p}$.

For any $x, y \in B$ we have $xy + yx = \operatorname{trd}(y)x + \operatorname{trd}(x)y - T(x, y)$, so if $\mathcal{O}$ is an order then $\mathcal{O} + \mathbb{Z}_F x$ is multiplicatively closed if and only if $T(x, y) \in \mathbb{Z}_F$ for all $y \in \mathcal{O}$. We have $T(x, y) = 0$ if x, y are orthogonal, and if x, y are a basis for an atomic block Q then by definition the valuation of $T(x, y)$ is at least the valuation of Q and so we can multiply each by $(\pi^{-1})^{\lfloor e/2 \rfloor}$, preserving integrality. □

After $\mathfrak{p}$-saturating, one can compute a maximal order as follows.

Algorithm 7.10. Let Λ be an order and let $\mathfrak{p}$ be prime. This algorithm computes a $\mathfrak{p}$-maximal order $\mathcal{O} \supset \Lambda$.

1. Compute a $\mathfrak{p}$-saturated order $\mathcal{O} \supset \Lambda$ and let $1, i, j, k$ be a $\mathfrak{p}$-saturated basis for $\mathcal{O}$. Let π^{-1} be an inverse uniformizer for $\mathfrak{p}$.
2. Suppose $\mathfrak{p}$ is odd. Swap i for j or k if necessary so that $a := i^2$ has $\operatorname{ord}_\mathfrak{p}(a) = 0$. Let $b := j^2$. If $\operatorname{ord}_\mathfrak{p} b = 0$, return $\mathcal{O}$. Otherwise, if $\operatorname{ord}_\mathfrak{p} b = 1$ and $(a/\mathfrak{p}) = 1$, solve
$$x^2 \equiv a \pmod{\mathfrak{p}}$$
for $x \in \mathbb{Z}_F/\mathfrak{p}$. Adjoin the element $\pi^{-1}(x - i)j$ to $\mathcal{O}$, and return $\mathcal{O}$.
3. Otherwise, $\mathfrak{p}$ is even. Let $t := \operatorname{trd}(i)$, let $a := -\operatorname{nrd}(i)$, and let $b := j^2$.
   (a) Suppose $\operatorname{ord}_\mathfrak{p} t = 0$. If $\operatorname{ord}_\mathfrak{p} b = 0$, return $\mathcal{O}$. If $\operatorname{ord}_\mathfrak{p} b = 1$ and $T^2 - tT + a = 0$ has a root x modulo $\mathfrak{p}$, and return $\mathcal{O} + \mathbb{Z}_F \pi^{-1}(x - i)j$.
   (b) Suppose $\operatorname{ord}_\mathfrak{p} \operatorname{trd}(i) > 0$. Let y, z, w be the output of Algorithm 6.5 with input a, b. Let
$$i' := (\pi^{-1})^e(1 + yi + zj + wij).$$
   Adjoin i' to $\mathcal{O}$, and return to Step 1.

Proof of correctness. At every step in the algorithm, for each prime $\mathfrak{q} \neq \mathfrak{p}$ the order $\mathcal{O}_\mathfrak{q}$ does not change, so we need only verify that $\mathcal{O}_\mathfrak{p}$ is indeed a maximal order.

In Step 2, we have that b is a uniformizer for $\mathfrak{p}$, that $\mathfrak{d}(\mathcal{O}_\mathfrak{p}) = 4ab\mathbb{Z}_{F,\mathfrak{p}}$. If $\operatorname{ord}_\mathfrak{p}(b) = 0$ then $\operatorname{ord}_\mathfrak{p} \mathfrak{d}(\mathcal{O}_\mathfrak{p}) = 0$ so $\mathcal{O}$ is indeed maximal. Otherwise, we have $\mathfrak{d}(\mathcal{O}_\mathfrak{p}) = \mathfrak{p}$ and $B_\mathfrak{p} \cong \left(\dfrac{K_\mathfrak{p}, b}{F_\mathfrak{p}}\right)$ where $K_\mathfrak{p} = F_\mathfrak{p}[i]$. We conclude that $B_\mathfrak{p}$ is a division ring (and hence $\mathcal{O}_\mathfrak{p}$ is maximal) if and only if $(a/\mathfrak{p}) = -1$. If $(a/\mathfrak{p}) = 1$ and $j' = \pi^{-1}(x - i)j$, then $1, i, j', ij'$ form the $\mathbb{Z}_{F,\mathfrak{p}}$-basis for a maximal order, since $(j')^2 = (\pi^{-1})^2(x^2 - a)b \in \mathbb{Z}_{F,\mathfrak{p}}$ and $j'i = -ij'$.

In Step 3, first note that ij is also orthogonal to $1, i$: we have i orthogonal to j so $\operatorname{trd}(ij) = 0$ so ij is orthogonal to 1, and similarly $\operatorname{trd}(ij\bar{i}) = \operatorname{trd}(\operatorname{nrd}(i)j) = 0$. In particular, we have $B_\mathfrak{p} = \left(\dfrac{K_\mathfrak{p}, b}{F_\mathfrak{p}}\right)$ where $K_\mathfrak{p} = F_\mathfrak{p}[i]$. By a comparison of discriminants, using the fact that the basis is normalized, we see that $1, i, j, ij$ is a $\mathfrak{p}$-saturated basis for $\mathcal{O}$ as well, so without loss of generality we may take $k = ij$.

Suppose first that $\operatorname{ord}_{\mathfrak{p}} \operatorname{trd}(i) = 0$, so we are in Step 3a. If $\operatorname{ord}_{\mathfrak{p}} b = 0$, then $\operatorname{ord}_{\mathfrak{p}} \mathfrak{d}(\mathcal{O}_{\mathfrak{p}}) = 0$ so $\mathcal{O}_{\mathfrak{p}}$ is maximal. If $\operatorname{ord}_{\mathfrak{p}} b > 0$, then since the basis is $\mathfrak{p}$-saturated we have $\operatorname{ord}_{\mathfrak{p}} b = 1$. Thus as in the case for $\mathfrak{p}$ odd, we have $B_{\mathfrak{p}}$ is a division ring if and only if $K_{\mathfrak{p}}$ is not a field, and as above the adjoining the element $\pi^{-1}(x - i)j$ yields a maximal order.

So suppose we are in Step 3b, so $\operatorname{ord}_{\mathfrak{p}} \operatorname{trd}(i) > 0$. Since $1, i, j, k$ is normalized, we have $\operatorname{ord}_{\mathfrak{p}} \operatorname{trd}(i) = \operatorname{ord}_{\mathfrak{p}} T(1, i) \leq \operatorname{ord}_{\mathfrak{p}} T(j, k)$. Adjoining i' to $\mathcal{O}$ gives a $\mathbb{Z}_{F,\mathfrak{p}}$-module with basis $1, i', j, i'j$ since $y \in (\mathbb{Z}_F/\mathfrak{p})^\times$; adjoining j' gives a module with basis $1, i', j', i'j'$ for the same reason. We verify that $\mathcal{O}_{\mathfrak{p}}$ after these steps is indeed an order: we have $\operatorname{trd}(i') = 2(\pi^{-1})^e \in \mathbb{Z}_{F,\mathfrak{p}}$ and $\operatorname{nrd}(i') = (\pi^{-1})^{2e}(1 - ay^2 - bz^2 + abw^2) \in \mathbb{Z}_{F,\mathfrak{p}}$ by construction, so at least $\mathbb{Z}_{F,\mathfrak{p}}[i] = \mathbb{Z}_{F,\mathfrak{p}} \oplus \mathbb{Z}_{F,\mathfrak{p}} i$ is a ring. Similarly we have $(j')^2 = b' \in \mathbb{Z}_{F,\mathfrak{p}}$. Finally, we have $\operatorname{trd}(i'i) = 2(\pi^{-1})^e ya$ and $\operatorname{trd}(i'j) = 2(\pi^{-1})^e zb$, so it follows that $\operatorname{trd}(i'j') = 0$, and hence $j'i' = -\overline{i'}j' = -i'j' - \operatorname{trd}(i')j'$, so indeed we have an order. □

Remark 7.11. One must really treat the even and odd prime cases separately. Consider, for example, $F = \mathbb{Q}$, and the quaternion algebra $B = \left(\dfrac{-3, 5}{\mathbb{Q}}\right)$. Then we have the maximal orders $\mathbb{Z}[(1 + i)/2] \subset \mathbb{Q}(i) \cong \mathbb{Q}(\sqrt{-3})$ and $\mathbb{Z}[(1 + j)/2] \subset \mathbb{Q}(j) \cong \mathbb{Q}(\sqrt{5})$, but we find that

$$\left(\frac{1+j}{2}\right)\left(\frac{1+i}{2}\right) = \left(\frac{1-i}{2}\right)\left(\frac{1+j}{2}\right) + \frac{ij}{2},$$

which is not integral (since $ij/2$ has norm $15/4$).

Remark 7.12. In the proof of correctness for Algorithm 7.10, in each case where $\mathfrak{p}$ is ramified in B we have in fact written $B_{\mathfrak{p}} \cong \left(\dfrac{K_{\mathfrak{p}}, \pi}{F_{\mathfrak{p}}}\right)$ where $K_{\mathfrak{p}}$ is the unramified extension of $F_{\mathfrak{p}}$. The reader will note the similarity between this algorithm and the algorithm to compute the Hilbert symbol: the former extends the latter by taking a witness for the fact that the algebra is split, namely a zerodivisor modulo $\mathfrak{p}$, and uses this to compute a larger order (giving rise therefore to the matrix ring).

Combining these two algorithms, we have the following immediate corollary.

Corollary 7.13. *There exists an algorithm to solve* (ExhibitMatrixRing) *for orders over* $\mathbb{Z}_{F,\mathfrak{p}}$.

(We recall the discussion in Sect. 4 for the representation of local fields and rings.) In other words, if $\mathcal{O} \subset B$ is an order in a quaternion algebra B over a number field F and $\mathfrak{p}$ is prime of $\mathbb{Z}_F$ which is unramified in B, then there exists an algorithm to compute an explicit embedding $\mathcal{O} \hookrightarrow \mathrm{M}_2(\mathcal{O}_{\mathfrak{p}})$.

Putting these two algorithms together, we have proved the following theorem.

Theorem 7.14. *Problem* (MaxOrder) *is deterministic polynomial-time reducible to the problem of factoring ideals in* $\mathbb{Z}_F$.

Proof. Given any order Λ, we factor its discriminant $\mathfrak{d}(\Lambda)$, and for each prime $\mathfrak{p} \mid \mathfrak{d}(\Lambda)$, we compute a $\mathfrak{p}$-saturated order containing Λ from Algorithm 7.9 and a $\mathfrak{p}$-maximal order $\mathcal{O}$ containing it using Algorithm 7.10. □

Complexity analysis. Given Theorem 7.14, we prove the following result which characterizes the abstract complexity class of this problem, following a hint of Ronyai [30, Sect. 6].

Theorem 7.15. *Problem* (ALGEBRAMAXORDER) *for any fixed number field F is probabilistic polynomial-time equivalent to the problem of factoring integers.*

To prove the theorem, we will use two lemmas. The first lemma is a standard fact.

Lemma 7.16. *The problem of factoring integral ideals $\mathfrak{a}$ of an arbitrary number field is probabilistic polynomial-time equivalent to the problem of factoring integers.*

Proof. Suppose $\mathfrak{a}$ is an integral ideal of F. After factoring the absolute discriminant d_F of F, we can in deterministic polynomial time compute the ring of integers $\mathbb{Z}_F$ of F as above. Now let $\mathfrak{a}$ be an ideal with norm $\mathrm{N}(\mathfrak{a}) = a$. After we factor a, for each prime $p \mid a$, we decompose $p\mathbb{Z}_F = \prod_i \mathfrak{p}_i^{e_i}$ into primes by a probabilistic polynomial-time algorithm due to Buchmann and Lenstra (see Cohen [6, Algorithm 6.2.9]): this algorithm uses a probabilistic algorithm to factor polynomials over a finite field, such as the Cantor-Zassenhaus algorithm; see von zur Gathen and Gerhard [13, Theorem 14.14] or Cohen [6, Sect. 3.4]. (In fact, for our applications, it suffices to have an algorithm to compute a square root in a finite field, for which we may use the algorithm of Tonelli and Shanks (see Cohen [6, Sect. 1.5.1])).

From this list of primes we easily obtain the factorization of $\mathfrak{a}$. Conversely, if one has an algorithm to factor ideals, then one may factor $a\mathbb{Z}_F$ into primes and computing norms we recover the prime factorization of a over $\mathbb{Z}$. □

Remark 7.17. Deterministically, already the problem of finding a nonsquare modulo a prime p is difficult; one unconditional result known is that the smallest quadratic nonresidue of a prime p is of size exponential in $\log p$; under condition of a generalized Riemann hypothesis, one can find a quadratic nonresidue which is of polynomial size in $\log p$.

We will also make use of one other lemma.

Lemma 7.18. *Let $\mathfrak{a}$ be an ideal of $\mathbb{Z}_F$ which is odd, not a square, and not a prime power. Let*

$$S = \left\{ b \in (\mathbb{Z}_F/\mathfrak{a})^\times : \textit{there exist } \mathfrak{p}^e, \mathfrak{q}^f \parallel \mathfrak{a} \textit{ with } \left(\frac{b}{\mathfrak{p}}\right)^e = -1 \textit{ and } \left(\frac{b}{\mathfrak{q}}\right)^f = 1 \right\}.$$

Then $\#S \geq \dfrac{1}{2}\#(\mathbb{Z}_F/\mathfrak{a})^\times$.

Proof. For an ideal $\mathfrak{b}$, let $\Phi(\mathfrak{b}) = \#(\mathbb{Z}_F/\mathfrak{b})^\times$. First consider the case where $\mathfrak{a} = \mathfrak{p}^e\mathfrak{q}^f$ is the product of two prime powers. Without loss of generality, we may assume e is odd. If f is even, then $b \in S$ if and only if $(b/\mathfrak{p}) = -1$, so $\#S = \Phi(\mathfrak{p}^e)/2 \cdot \Phi(\mathfrak{q}^f) = \Phi(\mathfrak{a})/2$. If f is odd, then $\#S = 2(\Phi(\mathfrak{p}^e)/2)(\Phi(\mathfrak{q}^f)/2) = \Phi(\mathfrak{a})/2$.

To conclude, write $\mathfrak{a} = \mathfrak{p}^e\mathfrak{q}^f\mathfrak{b}$ with $\mathfrak{b}$ coprime to $\mathfrak{p}\mathfrak{q}$ and e odd. Then by the preceding paragraph $\#S \geq (1/2)\Phi(\mathfrak{p}^e\mathfrak{q}^f)\Phi(\mathfrak{b}) = \Phi(\mathfrak{a})/2$. □

Proof of Theorem 7.15. Since one can factor ideals in probabilistic polynomial time given an algorithm to factor integers by Lemma 7.16, we may compute a maximal order as in the previous section as the resulting computations run in (deterministic) polynomial time.

Now we prove the converse. Suppose we have an algorithm to solve Problem (AlgebraMaxOrder). Let $a \in \mathbb{Z}_{>0}$ be the integer to be factored, which we may assume without loss of generality is odd, not a prime power, and not a square. We can in constant time (for fixed F) factor the absolute discriminant d_F, so we may also assume $\gcd(a, d_F) = 1$. It follows that the ideal $a\mathbb{Z}_F$ is also odd, not a prime power, and not a square.

We compute a random $b \in \mathbb{Z}_F/a\mathbb{Z}_F$ with $b \neq 0$. Since $\mathrm{N}(a\mathbb{Z}_F) = a^d$ where $d = [F : \mathbb{Q}]$, if $\mathrm{N}(b\mathbb{Z}_F)$ is not a power of a then dividing $\gcd(a^d, \mathrm{N}(b))$ by powers of a we obtain a factor of a. Otherwise, $\mathfrak{a} = a\mathbb{Z}_F + b\mathbb{Z}_F$ is a proper divisor of $a\mathbb{Z}_F$, and we repeat, computing a random $b \in \mathbb{Z}_F/\mathfrak{a}$—in at most d steps, we will either factor a or find an element b such that $a\mathbb{Z}_F + b\mathbb{Z}_F = \mathbb{Z}_F$. Note d depends only on F and not on B, so we find such a b in probabilistic polynomial time.

By Lemma 7.18, we can find in probabilistic polynomial time $b \in (\mathbb{Z}_F/a\mathbb{Z}_F)^\times$ such that $\mathfrak{p}^e, \mathfrak{q}^f \parallel a$ with $(b/\mathfrak{p})^e = -1$ and $(b/\mathfrak{q})^f = 1$, say. Let $B = \left(\dfrac{a, b}{F}\right)$. By hypothesis, calling an algorithm to solve (AlgebraMaxOrder) we may compute a maximal order $\mathcal{O} \subset B$.

We claim that $\mathfrak{p} \mid \mathfrak{d}(\mathcal{O})$ but $\mathfrak{q} \nmid \mathfrak{d}(\mathcal{O})$. Assuming this claim, we have that $\gcd(\mathrm{N}(\mathfrak{d}\ (\mathcal{O})), a)$ is a proper factor of a, and the proof is complete.

First we prove that $\mathfrak{p} \mid \mathfrak{d}(\mathcal{O})$. Since $\mathfrak{p}$ is prime to d_F, we know that $\mathfrak{p}$ is unramified in F, and since $\mathfrak{p}^e \parallel a\mathbb{Z}_F$ with e odd, the extension $F(\sqrt{a})/F$ is ramified at $\mathfrak{p}$. Since $(b/\mathfrak{p}) = -1$, by Corollary 5.5, the algebra B is ramified at $\mathfrak{p}$. Therefore by Lemma 7.5, $\mathfrak{p}$ divides the discriminant $\mathfrak{d}(\mathcal{O})$.

Now we show that $\mathfrak{q} \nmid \mathfrak{d}(\mathcal{O})$. If f is even, since $\mathfrak{q}^f \parallel a\mathbb{Z}_F$, we have that $F(\sqrt{a})/F$ is unramified at $\mathfrak{q}$; since also $(b/\mathfrak{q}) \neq 0$, by the same corollary, B is unramified at $\mathfrak{q}$. And if f is odd, then since $(b/\mathfrak{q})^f = 1$ we must have $(b/\mathfrak{q}) = 1$, and again by the corollary it follows that B is unramified. □

Relationship to conics. We return once again to the theme of rational points on conics.

We have seen that given an algorithm to factor integers, one can solve both Problems (IsMatrixRing), or equivalently (HasPoint), over a number field F in probabilistic polynomial time by factoring the discriminant and computing Hilbert symbols. We have also seen that (AlgebraMaxOrder) over a number field F is probabilistic polynomial time equivalent to the problem of factoring integers.

We are left to consider (ExhibitMatrixRing), or equivalently (ExhibitPoint). In the special case where $F = \mathbb{Q}$, one shows that again they are reducible to the problem of integer factorization.

Theorem 7.19 (Cremona-Rusin [8], Ivanyos-Szántó [15], Simon [36]). *There exists an explicit algorithm to solve* (ExhibitPoint) *over* $\mathbb{Q}$ *which runs in deterministic polynomial time given an oracle to factor integers.*

From our point of view, the algorithm(s) described in the above theorem can be rephrased in the following way: there exists an explicit algorithm which, given a order $\mathcal{O}$ over $\mathbb{Z}$ of discriminant 1 which is split at ∞, computes a zerodivisor $x \in \mathcal{O}$. This algorithm proceeds by computing a reduced basis of $\mathcal{O}$ with respect to the reduced norm nrd, a kind of indefinite LLL-algorithm.

Question 7.20. Does there exist an algorithm which, given an order $\mathcal{O}$ over $\mathbb{Z}_F$ of discriminant 1 which is split at all real places of F, computes a zerodivisor $x \in \mathcal{O}$?

One possible approach to this conjecture, then, is to provide an indefinite LLL algorithm over F in the special case of $\mathbb{Z}_F$-module of rank 4 and discriminant 1. Perhaps one can prove this at least in the case where $\mathbb{Z}_F$ is computably Euclidean?

We discuss the computational complexity of Problem (IsMatrixRing) over $\mathbb{Q}$ in the next section (and relate this to the problem of factoring integers). From the discussion above, it seems reasonable to conjecture the following.

Conjecture 7.21. *Problem* (ExhibitPoint) *over* $\mathbb{Q}$ *is (probabilistic) polynomial-time equivalent to the problem of factoring integers.*

Having treated the case of number fields in some detail, we note that over more general fields, the literature is much less complete.

Question 7.22. For which computable fields F is there an effective algorithm to solve Problems (HasPoint) and (ExhibitPoint)?

For example, one may ask for which fields F is there an effective version of the Hasse-Minkowski theorem? Of course, if one can solve (HasPoint), then given a conic which is known to have a solution one can always simply enumerate the points of $\mathbb{P}^2(F)$ until a solution is found.

8 Residuosity

In this final section, we return to Problem (IsMatrixRing) and characterize its computational complexity. Let F be a number field with ring of integers $\mathbb{Z}_F$.

For a nonzero ideal $\mathfrak{b}$ of $\mathbb{Z}_F$, let $\operatorname{sqrad}(\mathfrak{b})$ be the product of the prime ideals $\mathfrak{p}$ dividing $\mathfrak{b}$ to odd exponent, or equivalently the quotient of $\mathfrak{b}$ by the largest square ideal dividing $\mathfrak{b}$.

Problem (QUADRATICRESIDUOSITY). *Given an odd ideal $\mathfrak{b}$ and $a \in \mathbb{Z}_F$, determine if $a \in (\mathbb{Z}_F/\operatorname{sqrad}(\mathfrak{b}))^{\times 2}$, i.e., determine if a is a quadratic residue modulo* $\operatorname{sqrad}(\mathfrak{b})$.

Problem (QUADRATICRESIDUOSITY) reduces to the more familiar problem of quadratic residuosity when $\mathfrak{b}$ is a squarefree ideal, namely, to determine if $a \in (\mathbb{Z}_F/\mathfrak{b})^{\times 2}$. If $\mathfrak{b} = \mathfrak{p}$ is a prime ideal, one has $a \in (\mathbb{Z}_F/\mathfrak{p})^{\times 2}$ if and only if $(a/\mathfrak{p}) = 1$, and this Legendre symbol can be evaluated in deterministic polynomial time (as discussed above, by repeated squaring). In general, for $\mathfrak{b}$ squarefree, we have $a \in (\mathbb{Z}_F/\mathfrak{b})^{\times 2}$ if and only if $a \in (\mathbb{Z}_F/\mathfrak{p})^{\times 2}$ for all primes $\mathfrak{p} \mid \mathfrak{b}$. In particular, by this reduction if one can factor $\mathfrak{b}$, one can solve Problem (QUADRATICRESIDUOSITY). It is a terrific open problem in number theory to determine if the converse holds, even for the case $F = \mathbb{Q}$ and $\mathfrak{b}$ generated by pq with p, q distinct primes.

We first relate the Problems (ISMATRIXRING) and (QUADRATICRESIDUOSITY) as follows.

Proposition 8.1. *Problem* (ISMATRIXRING) *over F is deterministic polynomial-time reducible to Problem* (QUADRATICRESIDUOSITY) *over F.*

Proof. Let $B = \left(\dfrac{a,b}{F}\right)$ be a quaternion algebra over F. Scaling a, b by an integer square, we may assume $a, b \in \mathbb{Z}_F$. Recall that $B \cong \mathrm{M}_2(F)$ if and only if every place v of F is unramified in B, i.e., if $(a,b)_v = 1$ for all places v of F.

For fixed F, we can in constant (deterministic) time compute the set of even places of F. We then compute the Hilbert symbol $(a,b)_v$ for v real easily and for v even by Algorithm 6.6.

For the odd places, we first apply Lemma 5.5, which implies that we need only check primes $\mathfrak{p} \mid ab\mathbb{Z}_F$. We compute $\mathfrak{g} = a\mathbb{Z}_F + b\mathbb{Z}_F$ and then by small linear combinations we find $g \in \mathfrak{g}^{-1}$ such that $g\mathfrak{g}^{-1}$ is coprime to $a\mathbb{Z}_F$ and $b\mathbb{Z}_F$ and $(a+b)\mathbb{Z}_F$. Now $\left(\dfrac{a,b}{F}\right) \cong \left(\dfrac{a',b'}{F}\right)$ where $a' = a+b$ and $b' = -abg^2$. We claim that after repeating this eventually we will have a and b coprime. Indeed, if $\operatorname{ord}_{\mathfrak{p}}(a) = \operatorname{ord}_{\mathfrak{p}}(b)$ then already $\operatorname{ord}_{\mathfrak{p}}(-abg^2) = 0$, and if $\operatorname{ord}_{\mathfrak{p}}(a) > \operatorname{ord}_{\mathfrak{p}}(b) > 0$, say, then $\operatorname{ord}_{\mathfrak{p}}(-abg^2) = \operatorname{ord}_{\mathfrak{p}}(a) - \operatorname{ord}_{\mathfrak{p}}(b)$ and $\operatorname{ord}_{\mathfrak{p}}(a+b) = \operatorname{ord}_{\mathfrak{p}}(b)$, so then $\operatorname{ord}_{\mathfrak{p}}(a) + \operatorname{ord}_{\mathfrak{p}}(b) > \operatorname{ord}_{\mathfrak{p}}(a) = \operatorname{ord}_{\mathfrak{p}}(a') + \operatorname{ord}_{\mathfrak{p}}(b')$, and since this is a sequence of nonnegative integers eventually either we will have either $\operatorname{ord}_{\mathfrak{p}}(a) = 0$ or $\operatorname{ord}_{\mathfrak{p}}(b) = 0$.

Then for any prime $\mathfrak{p} \mid b\mathbb{Z}_F$, we have that $\mathfrak{p}$ is ramified in B if and only if $\mathfrak{p} \mid \operatorname{sqrad}(b\mathbb{Z}_F)$ and $(a/\mathfrak{p}) = -1$. We can test this latter condition for all $\mathfrak{p} \mid b\mathbb{Z}_F$ by calling the algorithm to solve (QUADRATICRESIDUOSITY) by determining if a is a quadratic residue modulo $\operatorname{sqrad}(b\mathbb{Z}_F)$. We then repeat this step with a, b interchanged, and we return TRUE if and only if both of these quadratic residuosity tests return TRUE. □

When $F = \mathbb{Q}$, in fact these problems are equivalent.

Theorem 8.2. *Problem* (ISMATRIXRING) *over $\mathbb{Q}$ is probabilistic polynomial-time equivalent to Problem* (QUADRATICRESIDUOSITY) *over $\mathbb{Q}$.*

Remark 8.3. Rónyai [29, 31] conditionally proves exactly Theorem 8.2 (under the assumption of the Generalized Riemann Hypothesis).

Before proving this theorem, we derive one preliminary result.

Lemma 8.4. *Let* $a, b \in \mathbb{Z}_{>0}$ *be such that* b *is odd and* $(a/b) = 1$. *Let* ℓ *be an odd prime such that* $\ell b \in (\mathbb{Z}/a\mathbb{Z})^{\times 2}$ *and* $\left(\dfrac{a}{\ell}\right) = 1$. *Then* $\left(\dfrac{a, \ell b}{\mathbb{Q}}\right) \cong \mathrm{M}_2(\mathbb{Q})$ *if and only if* a *is a square modulo* $\operatorname{sqrad}(b)$.

Proof. Again, we have $\left(\dfrac{a, \ell b}{\mathbb{Q}}\right) \cong \mathrm{M}_2(F)$ if and only if $(a, \ell b)_v = 1$ for all places v of $\mathbb{Q}$. Since $a > 0$, we know $(a, \ell b)_\infty = 1$. By hypothesis, for all odd $p \mid a$ we have $(\ell b/p) = 1$ hence $(a, \ell b)_p = 1$, and similarly $(a, \ell b)_\ell = 1$. Moreover, since $(a/b) = 1$, the number of primes $p \mid \operatorname{sqrad}(b)$ such that $(a/p) = -1$ must be even, and since the quaternion algebra $\left(\dfrac{a, \ell b}{\mathbb{Q}}\right)$ is ramified at an even number of places, we conclude that $(a, \ell b)_2 = 1$. Therefore $\left(\dfrac{a, \ell b}{\mathbb{Q}}\right) \cong \mathrm{M}_2(F)$ if and only if $(a, \ell b)_p = 1$ for all $p \mid \operatorname{sqrad}(b)$ if and only if a is a square modulo $\operatorname{sqrad}(b)$. □

The preceding lemma shows that the two problems in Theorem 8.2 can be linked by finding a suitable prime ℓ. The conditions on ℓ are congruence conditions, so by the theorem on primes in arithmetic progression, such primes are abundant. Explicitly, we rely on the specialization of a result from analytic number theory, stated by Adleman, Pomerance, and Rumely [2, Proposition 8] and attributed to the proof of Linnik's theorem by Bombieri (using results of Gallagher and related to a result of Tatuzawa); see their paper for further discussion.

Lemma 8.5. *There exist effectively computable (absolute) constants* $x_0, \delta \in \mathbb{R}_{>0}$ *such that whenever* $x \geq x_0$, *we have*

$$\left| \sum_{\substack{\ell \leq x \\ \ell \equiv b \pmod{q}}} \log \ell - \frac{x}{\phi(q)} \right| \leq \frac{x}{2\phi(q)}$$

for all q *with* $1 \leq q \leq x^\delta$ *and all* b *with* $\gcd(b, q) = 1$, *except possibly for those* q *which are multiples of a certain integer* $q_0(x) > (\log x)^{3/2}$.

Proof of Proposition 8.2. We must show that if we are able to solve (IsMatrixRing), then we can solve Problem (QuadraticResiduosity) in probabilistic polynomial time.

Let $x = \max((4b)^{1/\delta}, x_0)$, with x_0, δ as in Lemma 8.5. Let c be a random integer with $1 \leq c < b$. We compute $q \equiv ac^2 \pmod{4b}$ with $1 \leq q < 4b$ and $q \equiv 1 \pmod{4}$. Then q is a random element in $[1, 4b] \cap \mathbb{Z}$ such that $aq \in (\mathbb{Z}/b\mathbb{Z})^{\times 2}$ and $q \equiv 1 \pmod{4}$. Let

$$Q = \{1 \leq q < b : aq \in (\mathbb{Z}/b\mathbb{Z})^{\times 2} \text{ and } q \equiv 1 \pmod{4}\}.$$

From Lemma 8.5, we have $\sum_{\ell \leq x,\ \ell \equiv a \pmod{q}} \log \ell < x/(2\phi(q))$ only if q is divisible by $q_0(x) > (\log x)^{3/2}$; thus the set of such $q \in Q$ has cardinality at most $\#Q/(\log x)^{3/2}$. Using partial summation (a standard argument which can be found in Davenport [9, p. 112]), it follows that a random $q \in Q$ has probability $1 - 1/(\log x)^{3/2}$ of satisfying

$$\pi(x;q,b) = \#\{\ell \leq x : \ell \text{ prime},\ \ell \equiv b \pmod{q}\} < \frac{1}{2\phi(q)} \frac{x}{\log x}$$

whenever $\gcd(b,q) = 1$. We then compute a random integer $1 \leq \ell < x$ with $\ell \equiv b \pmod{q}$ and test if ℓ is prime, which can be done in (deterministic) polynomial time [1]. Combining these, in probabilistic polynomial time, we may assume that ℓ indeed is prime.

We conclude by calling the algorithm to solve (IsMatrixRing) on $B = \left(\dfrac{q, \ell b}{\mathbb{Q}}\right)$. We have

$$\left(\frac{q}{\ell}\right) = \left(\frac{\ell}{q}\right) = \left(\frac{b}{q}\right) = \left(\frac{q}{b}\right) = \left(\frac{a}{b}\right) = 1$$

since $q \equiv 1 \pmod{4}$, and $\ell b \equiv 1 \pmod{q}$. So by Lemma 8.4, we have $B \cong \mathrm{M}_2(\mathbb{Q})$ if and only if q is a square modulo $\operatorname{sqrad}(b)$, which holds only if a is a square modulo $\operatorname{sqrad}(b)$, as desired. □

We leave the natural generalization where $\mathbb{Q}$ is replaced by a number field F as an open question.

References

1. Manindra Agrawal, Neeraj Kayal, and Nitin Saxena, *PRIMES is in P*, Ann. of Math. (2) **160** (2004), no. 2, 781–793.
2. Leonard M. Adleman, Carl Pomerance, and Robert S. Rumely, *On distinguishing prime numbers from composite numbers*, Ann. of Math. (2) **117** (1983), no. 1, 173–206.
3. Wieb Bosma, John Cannon, and Catherine Playoust, *The Magma algebra system. I. The user language*, J. Symbolic Comput. **24** (1997), vol. 3–4, 235–265.
4. J. A. Buchmann and H. W. Lenstra, Jr., *Approximating rings of integers in number fields*, J. Théor. Nombres Bordeaux **6** (1994), no. 2, 221–260.
5. A. L. Chistov, *The complexity of the construction of the ring of integers of a global field*, Soviet Math. Dokl. **39** (1989), no. 3, 597–600.
6. Henri Cohen, *Computational algebraic number theory*, Grad. Texts in Math., vol. 193, Springer, Berlin, 2000.
7. Henri Cohen, *Advanced topics in computational algebraic number theory*, Grad. Texts in Math., vol. 193, Springer, Berlin, 2000.
8. J. E. Cremona and D. Rusin, *Efficient solution of rational conics*, Math. Comp. **72** (2003), no. 243, 1417–1441.

9. Harold Davenport, *Multiplicative number theory*, 3rd. ed., Graduate texts in mathematics, vol. 74, Springer-Verlag, Berlin, 2000.
10. Leonard Eugene Dickson, *Algebras and their arithmetics*, Dover, New York, 1960.
11. Carsten Friedrichs, *Berechnung von Maximalordnungen uber Dedekindringen*, Ph. D. dissertation, Technischen Universität Berlin, 2000.
12. A. Fröhlich, *Local fields*, in *Algebraic number theory*, J.W.S. Cassels and A. Fröhlich, eds., Thompson Book Company, Washington, 1967, 1–41.
13. Joachim von zur Gathen and Jürgen Gerhard, *Modern computer algebra*, 2nd edition, Cambridge University Press, Cambridge, 2003.
14. Florian Hess, *Computing Riemann-Roch spaces in algebraic function fields and related topics*, J. Symbolic Comput. **33** (2002), no. 4, 425–445.
15. Gábor Ivanyos and Ágnes Szántó, *Lattice basis reduction for indefinite forms and an application*, Proceedings of the 5th Conference on Formal Power Series and Algebraic Combinatorics (Florence, 1993), Discrete Math. **153** (1996), no. 1–3, 177–188.
16. Gábor Ivanyos and Lajos Rónyai, *Finding maximal orders in semisimple algebras over* $\mathbb{Q}$, Comput. Complexity **3** (1993), no. 3, 245–261.
17. Nathan Jacobson, *Finite-dimensional division algebras over fields*, Springer-Verlag, Berlin, 1996.
18. Markus Kirschmer and John Voight, *Algorithmic enumeration of ideal classes for quaternion orders*, SIAM J. Comput. (SICOMP) **39** (2010), no. 5, 1714–1747.
19. Max-Albert Knus, *Quadratic forms, Clifford algebras and spinors*, Seminários de Matemática, 1, Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Ciência da Computaç ã o, Campinas, 1988.
20. Max-Albert Knus, Alexander Merkurjev, and Jean-Pierre Tignol, *The book of involutions*, American Math. Soc. Colloquium Publications, vol. 44, AMS, Providence, RI, 1998.
21. T.Y. Lam, *A first course in noncommutative rings*, 2nd ed., Graduate texts in mathematics, vol. 131, American Math. Soc., Providence, 2001.
22. H.W. Lenstra, Jr., *Algorithms in algebraic number theory*, Bull. Amer. Math. Soc. (N.S.) **26** (1992), no. 2, 211–244.
23. H. W. Lenstra, Jr., *Computing Jacobi symbols in algebraic number fields*, Nieuw Arch. Wisk. (4) **13** (1995), no. 3, 421–426.
24. Gabriele Nebe and Allan Steel, *Recognition of division algebras*, J. Algebra **322** (2009), no. 3, 903–909.
25. Jürgen Neukirch, *Algebraic number theory*, Grundlehren der mathematischen Wissenschaften, vol. 322, Springer-Verlag, Berlin, 1999.
26. O. Timothy O'Meara, *Introduction to quadratic forms*, Classics in Mathematics, Springer-Verlag, Berlin, 2000.
27. Michael Pohst and Hans Zassenhaus, *Algorithmic algebraic number theory*, Revised reprint, Encyclopedia of Mathematics and its Applications, vol. 30, Cambridge University Press, Cambridge, 1997.
28. Irving Reiner, *Maximal orders*, Clarendon Press, Oxford, 2003.
29. Lajos Rónyai, *Zero divisors in quaternion algebras*, J. Algorithms **9** (1988), 494–506.
30. Lajos Rónyai, *Algorithmic properties of maximal orders in simple algebras over* $\mathbb{Q}$, Comput. Complexity **2** (1992), no. 3, 225–243.
31. Lajos Rónyai, *Simple algebras are difficult*, Proceedings, 19th ACM Symp. on Theory of Computing, 1990, 398–408.
32. Lajos Rónyai, *Computing the structure of finite algebras*, J. Symbolic Computation **9** (1990), 355–373.
33. Winfried Scharlau, *Quadratic and Hermitian forms*, Springer-Verlag, Berlin, 1985.
34. Viggo Stoltenberg-Hansen and John V. Tucker, Computable rings and fields, *Handbook of computability theory*, ed. Edward R. Griffor, North-Holland, Amsterdam, 1999, 336–447.
35. Dénis Simon, *Equations dans les corps de nombres et discriminants minimaux*, thèse, Universit Bordeaux I, 1998.

36. Dénis Simon, *Solving quadratic equations using reduced unimodular quadratic forms*, Math. Comp. **74** (2005), no. 251, 1531–1543.
37. Christiaan van de Woestijne, *Deterministic equation solving over finite fields*, ISSAC'05, ACM, New York, 2005, 348–353.
38. Marie-France Vignéras, *Arithmétique des algèbres de quaternions*, Lecture notes in mathematics, vol. 800, Springer, Berlin, 1980.
39. John Voight, *Quadratic forms and quaternion algebras: Algorithms and arithmetic*, Ph.D. thesis, University of California, Berkeley, 2005.
40. John Voight, *Rings of low rank with a standard involution*, Illinois J. Math. **55** (2011), no. 3, 1135–1154.
41. John Voight, *Characterizing quaternion rings over an arbitrary base*, J. Reine Angew. Math. **657** (2011), 113–134