GENERATING SETS FOR MAXIMAL ORDERS IN RATIONAL QUATERNION ALGEBRAS

KIRSTEN EISENTRÄGER, EYAL Z. GOREN, ANNAMARIA IEZZI, HARUN KIR, EDA KIRIMLI, JONATHAN R. LOVE, WILLIAM E. MAHANEY, JENNIFER PARK, AND MARIA SABITOVA

Abstract. Given a maximal order $\mathfrak{O}$ in a rational definite quaternion algebra, and a prime ℓ coprime to the discriminant of $\mathfrak{O}$, this paper considers subsets of $\mathfrak{O}$ consisting of elements with ℓ-power norms that together generate $\mathfrak{O}$ as a $\mathbb{Z}$-algebra. We prove two theorems about the existence of such sets: the first states that $\mathfrak{O}$ is generated by elements of norm ℓ^k for any k larger than an explicit bound, and the second states that there is a generating set for $\mathfrak{O}$ consisting of at most three elements, each with norm a power of ℓ. We discuss implications for the study of supersingular isogeny graphs. As steps towards these theorems, for quaternion orders $\mathcal{O}$ that are not necessarily maximal, we also prove structural results about the algebra of Brandt matrices for $\mathcal{O}$ and explicit bounds on the coefficients of the theta function of $\mathcal{O}$. Computational experiments are also discussed.

Contents

1. Introduction

The principal results of this paper concern generating sets for certain orders in definite quaternion algebras, where essential use is made of the theta functions of these orders. As we shall explain, the project is closely related to questions of current interest in mathematical cryptography, but it equally has its roots in several other areas of mathematics: lattices and their generating sets, counting representations of integers by quadratic forms, quaternion algebras, and the theory of Brandt modules, which allows us to relate the ideal theory of quaternion orders to modular forms. This paper contributes to all of these areas.

For ease of exposition, we will focus in the introduction on the case that $\mathcal{A}_p$ is a definite quaternion algebra over $\mathbb{Q}$ ramified only at p and ∞, and $\mathfrak{O}$ is a maximal order in $\mathcal{A}_p$; such orders arise as endomorphism rings of supersingular elliptic curves over an algebraic

Date: September 17, 2026.

closure $\overline{\mathbb{F}}_p$ of a finite field $\mathbb{F}_p$. Fix a prime $\ell \neq p$. In [GL25b, Theorem 1.1] it is proven that there exists a subset $C \subseteq \mathfrak{O}$ such that the norm of each $x \in C$ is a power of ℓ and C generates $\mathfrak{O}$ as a $\mathbb{Z}$-algebra:[1] that is, every element of $\mathfrak{O}$ can be written as a $\mathbb{Z}$-linear combination of products of elements in C, including the empty product $1 \in \mathfrak{O}$. In this paper we prove effective bounds on these generating sets. Our main result is a bound on the norms of the generators.

Theorem 1.1. *Let p and ℓ be distinct primes, and k a positive integer with $\ell^{k/2-1} > 3pk$. For any maximal order $\mathfrak{O} \subset \mathcal{A}_p$, the elements of norm ℓ^k in $\mathfrak{O}$ generate $\mathfrak{O}$ as a $\mathbb{Z}$-algebra.*

For example, these assumptions hold if $p > \ell^{17}$ and $k > 3\log_\ell(p)$. A more general version of this result — applicable to maximal orders in any definite quaternion algebra over $\mathbb{Q}$ — is proven in Theorem 7.2.

Along with bounding the norms of elements of the generating set C, we also prove a sharp bound on the number of elements in C.

Theorem 1.2. *For any maximal order $\mathfrak{O}$ in $\mathcal{A}_p$, there exists a set of at most three elements, each with norm in $\{\ell^k : k \geq 1\}$, that generates $\mathfrak{O}$ as a $\mathbb{Z}$-algebra.*

This is a special case of Theorem 8.2, which generalizes this result to arbitrary definite quaternion algebras $\mathcal{A}$ over $\mathbb{Q}$ and replaces the set of allowable norms $\{\ell^k : k \geq 1\}$ with any infinite set of positive integers coprime to the discriminant of $\mathcal{A}$.

1.1. **Motivation: cryptography and otherwise.** A primary motivation for these results arises from the perspective of cryptographic protocols based on supersingular isogenies and the associated isogeny graphs. These directed graphs are constructed as follows: one fixes two distinct primes $\ell \neq p$. The vertices of the graphs are supersingular j-invariants in characteristic p, of which there are about $p/12$, all of which are in $\mathbb{F}_{p^2}$. Outgoing edges from a vertex j_0 correspond to isogenies of degree ℓ from a preselected elliptic curve with j-invariant j_0, considering two isogenies as the same if they have the same kernel. This is thus a directed graph, which is $(\ell+1)$-regular for outgoing edges. This defines the supersingular ℓ-isogeny graph $G(p,\ell)$. These graphs and their various generalizations have been studied intensively in recent years, largely motivated by their cryptographic applications ([GL25a] contains references to the relevant literature). However, supersingular graphs were studied prior to their use in cryptography, notably by Mestre [Mes86] in the context of calculating spaces of modular forms and Pizer [Piz90] in the context of Ramanujan graphs.

In [CLG09], Charles, Goren, and Lauter introduced a cryptographic hash function (henceforth the "CGL hash function") based on supersingular isogeny graphs, which was later suggested for use in post-quantum cryptography. While the jury is still out on the general applicability of this construction, there is no doubt that a host of fruitful mathematical questions arose, and many interesting developments concerning elliptic curves, complex multiplication and quaternion algebras have been made in that context. This paper contributes further to this growing field.

The basic construction of the CGL hash function is as follows: the input message, a string M of 0's and 1's of large length, is to be converted into a shorter string, the hash value. One chooses a large prime p and a small prime ℓ; for instance $\ell = 2$ and (currently) $p \approx 2^{500}$ are good choices. The original string M is converted in some deterministic fashion into directions for walking the supersingular isogeny graph $G(p,\ell)$ starting at some publicly known initial vertex j_0 that is part of the protocol. The hash value is the final vertex j the walk reaches, which in turn can be encoded in some fashion as a string of 0's and 1's (of length about 1000, for the parameters mentioned above).

[1] In fact the cited result proves that C generates $\mathfrak{O}$ as a $\mathbb{Z}$-module, but in this paper we focus primarily on generating sets of $\mathbb{Z}$-algebras.

The key underlying hardness assumption for the above hash function and other cryptographic constructions based on supersingular isogenies is the computational difficulty of finding isogenies between supersingular elliptic curves, often referred to as the *Isogeny Problem.* This problem was shown to be equivalent, under polynomial time reductions, to the *Supersingular Endomorphism Ring Problem* [HW26], building on earlier work [KLPT14, Wes21]. Given a prime p and a supersingular elliptic curve E defined over $\mathbb{F}_{p^2}$, the Supersingular Endomorphism Ring Problem asks to find four endomorphisms of E in an efficient representation that generate the endomorphism ring of E as a lattice. Each cycle in a supersingular ℓ-isogeny graph through a vertex j_0 produces an endomorphism of the associated elliptic curve E_0 by composing the isogenies associated to each edge in the cycle. Modulo certain details concerning curves with extra automorphisms, one can build a dictionary between cycles in $G(p, \ell)$ and endomorphisms of ℓ-power degree. In particular, one can ask whether such cycles generate the endomorphism ring of E_0, and if so, how many of them are needed.

The Supersingular Endomorphism Ring Problem is also equivalent to that of determining, given E and p as above, the maximal order $\mathfrak{O}$ in the rational quaternion algebra $\mathcal{A}_p$ ramified only at p and ∞ which is isomorphic to the endomorphism ring of E. These different problems related to isogenies and the relevant literature on them are discussed in [CFG+26]. The computation of endomorphism rings was first studied by Kohel in his thesis [Koh96], where he proposed an approach based on finding cycles in $G(p, \ell)$. Later work determined necessary and sufficient conditions for two such cycles to generate a finite-index subring of $\mathfrak{O}$ [BCNE+19]. There are algorithms that use such a finite-index subring to compute the entirety of $\mathfrak{O}$ (see e.g. [EHL+20]), but the output of these algorithms is typically not a description of $\mathfrak{O}$ purely in terms of cycles of $G(p, \ell)$. In [GL25b] it was shown that such cycles do generate the entirety of $\mathfrak{O}$, but the results were not effective: no bound on the number of cycles required or the lengths of such cycles was given. Theorem 1.1 and Theorem 1.2 give explicit bounds.

More broadly, one is led to a natural question in the theory of lattices: given an integral lattice Λ and a set S of positive integers, is Λ generated (as a $\mathbb{Z}$-module) by vectors whose norms lie in S? General results concerning this can be found in [GL25b, GL25a]. To approach this problem one may consider the sublattice of Λ spanned by elements with norm in S, and ask whether it equals Λ. This in turn leads to a natural follow-up question. Given a pair of integral lattices $\Lambda_1 \supsetneq \Lambda_2$ of the same rank with associated theta functions $\theta_{\Lambda_1}, \theta_{\Lambda_2}$, their difference $\theta_{\Lambda_1} - \theta_{\Lambda_2}$ is a nonzero modular form of a certain level and weight determined effectively by properties of the lattices with non-negative Fourier coefficients. One can then ask, how large are the coefficients of the terms of this form with exponent in S? While this question may be posed for arbitrary lattices, our results in this direction rely on the multiplicative structure of an order in a quaternion algebra, a property general lattices lack. In the case of a singleton $S = \{\ell^k\}$ for an appropriate value of k, we prove that the associated coefficient is positive when Λ_1 is a maximal order in a rational definite quaternion algebra and Λ_2 is any suborder (Proposition 1.3 below); this is the key ingredient in the proof of Theorem 1.1.

1.2. **Discussion of results.** We now turn to a more in depth discussion of the results of this paper. Let $\mathcal{A}$ be any definite quaternion algebra over $\mathbb{Q}$, and $\mathfrak{O}$ a maximal order in $\mathcal{A}$. Our aim is to prove effective results on generating sets of $\mathfrak{O}$ consisting of elements with norm a power of ℓ.

1.2.1. *Bounding norms.* Our main result is the existence of a generating set C for $\mathfrak{O}$ with elements whose norm is an explicitly bounded power of a prime ℓ (Theorem 7.2). To obtain this, we prove explicit bounds on $a_n(\mathcal{O})$, the number of elements of norm n in an order $\mathcal{O} \subseteq \mathcal{A}$. Let ϕ denote Euler's totient function.

Proposition 1.3 (cf. Proposition 7.1)**.** *Let $\mathcal{O} \subset \mathcal{A}$ be an order, $\ell \nmid \operatorname{disc}(\mathcal{A})$ a prime, and $k \geq 1$ be an integer. If $\mathcal{O}$ is maximal then*

$$a_{\ell^k}(\mathcal{O}) \geq \frac{24}{\phi(\operatorname{disc}(\mathcal{A}))} \frac{\ell^{k+1}-1}{\ell-1} - 12k\ell^{(k+2)/2},$$

and if $\mathcal{O}$ is not maximal then

$$a_{\ell^k}(\mathcal{O}) \leq \frac{16}{\phi(\operatorname{disc}(\mathcal{A}))} \frac{\ell^{k+1}-1}{\ell-1} + 12k\ell^{(k+2)/2}.$$

Theorem 7.2 (and Theorem 1.1) follows fairly directly from these bounds. Namely, let $\mathbb{Z}[C]$ denote the $\mathbb{Z}$-algebra generated by the set C consisting of all the elements of $\mathfrak{O}$ of norm ℓ^k. The assumptions on k ensure that $\mathfrak{O}$ contains strictly more elements of norm ℓ^k than any non-maximal order, so the order $\mathbb{Z}[C]$ must be maximal, that is, $\mathbb{Z}[C] = \mathfrak{O}$.

We have thus reduced to a question about bounding $a_n(\mathcal{O})$, the number of elements of norm n in an order $\mathcal{O} \subset \mathcal{A}$. To motivate our approach, we first describe a classical method for counting the number of elements of each norm in an integral lattice, and then explain where the approach breaks down and the modifications we make. We begin by considering the *theta function* of $\mathcal{O}$,

$$\theta_{\mathcal{O}}(q) := \sum_{n=0}^{\infty} a_n(\mathcal{O}) q^n.$$

If we evaluate this expression at $q = e^{2\pi i z}$ for z in the upper half plane, the result is a modular form of weight 2 and level $\Gamma_0(N)$ for some positive integer N which depends only on the structure of $\mathcal{O}$ as a lattice. Classical results about modular forms tell us that there exists a decomposition

$$\theta_{\mathcal{O}}(q) = E(q) + \sum_{i=1}^{m} c_i f_i(q^{d_i})$$

where E is an *Eisenstein series*, and for each $i = 1, \ldots, m$, f_i is a *newform* of some level M_i, $M_i \mid N$, $d_i \mid (N/M_i)$, and c_i is a constant. We can therefore compute $a_n(\mathcal{O})$ by computing the coefficients of each modular form on the right. Eisenstein series have explicitly computable coefficients, and while the coefficients of newforms are more mysterious, their growth is precisely controlled by the *Deligne bound*. This approach allows us to write $a_n(\mathcal{O})$ as the sum of a main term, namely the coefficient of q^n in $E(q)$, and an error term coming from the newforms. When we restrict to a set of positive integers n satisfying certain mild constraints (for example, being coprime to N is sufficient), we can then prove that the main term dominates the error term for sufficiently large n, giving us asymptotic formulas for $a_n(\mathcal{O})$.

The downside to this approach is that a priori, we have no control over the coefficients c_i, which means that it is impossible to obtain effective bounds using this approach; this problem arises because we only used the fact that $\theta_{\mathcal{O}}$ lies *somewhere* in the vector space of modular forms with given weight and level. Instead we consider a slightly modified approach, which takes advantage of the algebraic structure of quaternion orders in order to show that $\theta_{\mathcal{O}}$ lies fairly close to the origin in this space of modular forms.

To do this we define the *Brandt module* $\mathfrak{M}(\mathcal{O})$ of $\mathcal{O}$, which is a vector space with basis vectors $v_1, \ldots, v_h$ corresponding to invertible right-ideal classes $[I_1], \ldots, [I_h]$ of $\mathcal{O}$; we typically order the basis so that $[I_1]$ is the principal right-ideal class. The Brandt module comes equipped with *Brandt matrices* B_n for all $n \geq 0$, which are linear maps on $\mathfrak{M}(\mathcal{O})$ that measure relative containment of ideal classes (Section 2.5). The structure of the Brandt module together with its Brandt matrices may be thought of as a noncommutative analogue of the ideal class group of an order in a number field. In the case that $\mathcal{O}$ is a maximal order in $\mathcal{A}_p$, the Brandt matrix B_ℓ for prime $\ell \neq p$ is an adjacency matrix for $G(p, \ell)$. For more information see for example [Gro87]. An important result we need about the Brandt algebra is the following.

Theorem 1.4 (Theorem 3.1)**.** *Let $\mathcal{O}$ be an order in a quaternion algebra over $\mathbb{Q}$. The Brandt matrices B_n for all $n \geq 0$ pairwise commute.*

Remark 1.5. We could not find this result in the literature; most sources only prove this for maximal orders $\mathcal{O}$, or for operators B_n with n coprime to the discriminant of $\mathcal{O}$. We note that the behavior of Brandt matrices B_ℓ for primes ℓ dividing the discriminant can be quite subtle. Many existing approaches to proving results of this form do not apply in the cases we need: see Section 3.1.

We then define a symmetric bilinear pairing Θ on $\mathfrak{M}(\mathcal{O})$ valued in the space $\mathcal{M}_2(\Gamma_0(N))$ of modular forms of weight 2 and level $\Gamma_0(N)$. This map generalizes the notion of theta function to the entire Brandt module: $2\Theta(v_i, v_j)$ for each i, j is the theta function of some lattice in $\mathcal{A}$, with $2\Theta(v_1, v_1) = \theta_{\mathcal{O}}$ (Lemma 5.5). Under this map, simultaneous eigenvectors for all Brandt operators are sent to modular forms that are eigenvectors for all Hecke operators T_n with n coprime to N (cf. Lemma 5.6). We can therefore prove the following decomposition, which should be thought of as a modification of the classical decomposition into Eisenstein series and newforms.

Theorem 1.6 (Theorem 6.1)**.** *Let $\mathcal{O}$ be an order in a definite quaternion algebra over $\mathbb{Q}$, and let h be the class number of $\mathcal{O}$. For some positive integer N, there exist modular forms $f_1, \ldots, f_h \in \mathcal{M}_2(\Gamma_0(N))$ such that*

- *each f_i is an eigenvector for the Hecke operator T_n for all n coprime to N,*
- *the coefficient of q in each f_i is equal to 1, and*
- *there exist $c_1, \ldots, c_h \geq 0$ with $c_1 + \cdots + c_h = |\mathcal{O}^\times|$ and*

$$\theta_{\mathcal{O}}(q) = c_1 f_1(q) + \cdots + c_h f_h(q).$$

The new key input here is the non-negativity of the coefficients c_i, from which we can conclude that the cuspidal contribution to this sum is effectively bounded.

Unfortunately, because the forms f_i occurring in this decomposition are not necessarily Eisenstein series or newforms, we do not have as precise control over their q-expansions. To handle this, we restrict our attention to maximal orders as well as *submaximal orders*, those orders $\mathcal{O}$ that are contained in a maximal order $\mathfrak{O}$ with no other order strictly between the two. We classify the possible structures of submaximal orders (Lemma 4.2) and then we consider the possible q-expansions that the forms f_i from Theorem 1.6 may take in each case. This part of the proof is by far the most technical, but the end result is the recovery of a similar decomposition as the one found in the classical case: some forms are "Eisenstein-like" in that their coefficients are explicitly computable (Section 6.2), and others are "newform-like" in that their coefficients can be effectively bounded (Section 6.3). Putting these results together yields explicit lower bounds on $a_n(\mathfrak{O})$ when $\mathfrak{O}$ is maximal, and explicit upper bounds on $a_n(\mathcal{O})$ when $\mathcal{O}$ is submaximal; because any non-maximal order is contained in a submaximal order, this upper bound then applies to arbitrary non-maximal orders, which gives the desired bounds as in Proposition 1.3.

1.2.2. *Bounding number of generators.* Finally, we turn to bounding the number of elements in a generating set, namely Theorem 8.2. For this proof we largely follow the methods developed in [GL25a], namely the use of a local-global principle for lattice cosets [GL25a, Theorem 6.6.1]. A direct application of this local-global principle can easily produce a $\mathbb{Z}$-algebra generating set C for $\mathfrak{O}$ with $|C| \leq 6$. To get the bound down to three (the best possible, see Remark 8.3) requires one key new idea, which is to produce two elements $\alpha, \beta \in \mathfrak{O}$ which generate quadratic orders with coprime discriminants; after this, the third element can be selected so that locally, prime by prime, it fills out the rest of $\mathfrak{O}$. The main tool that allows this process to work is Lemma 8.8, which says that given a local order $\mathfrak{O}_\ell$ and an element $x \in \mathfrak{O}_\ell$ such that $1, x$ can be completed to a basis of $\mathfrak{O}_\ell$, it is in fact possible to complete it to a basis of the form $1, x, y, xy$ where y has (almost) any desired norm. We note also that the proof shows that $\mathfrak{O}$ is generated as a $\mathbb{Z}$-module by at most seven elements; see the discussion following Theorem 8.2.

1.3. **Outline.** The proof of Theorem 7.2 (of which Theorem 1.1 is a special case) is the focus of Section 2–Section 7. In Section 2 we introduce the tools needed from the theory of quadratic forms, lattices, and quaternion orders, building up to the definition of Brandt matrices. Section 3 proves commutativity of the Brandt algebra. Section 4 classifies submaximal orders, and then proves special properties about the Brandt algebras of specific types of submaximal orders.

In Section 5 we turn to the theory of modular forms, including an introduction of the theta operator and its properties. Section 6 proves Theorem 1.6 and studies the Fourier coefficients of modular forms coming from simultaneous eigenvectors of Brandt matrices of maximal or submaximal orders. The above work culminates in Section 7, which proves bounds on the number of elements of norm ℓ^k in a quaternion order and uses this to prove Theorem 7.2.

We note that Section 4.3 and Section 6.3 are devoted to careful analysis of special cases: these sections investigate Brandt algebras of specific submaximal orders and the q-expansions of associated theta functions. The reader aiming to understand the big picture can come back to these sections as needed.

In Section 8 we prove Theorem 8.2. In Section 9 we report on some computational experiments.

1.4. **Acknowledgments.** This project began at the Banff International Research Station (BIRS) Workshop 25w5500, *Isogeny Graphs in Cryptography.* The authors are very grateful to the organizers of this workshop: Travis Morrison, Victoria de Quehen, Jason LeGrow, and Chloe Martindale. The project was led by Jonathan Love, who was supported by ERC Starting Grant 101076941 ('GAGARIN'). Eisenträger was partially supported by National Science Foundation award CNS-2521084. Goren was partially supported by NSERC grant 223148. Iezzi was partially supported by the French National Research Agency (ANR) under the "France 2030" program (ANR-15-IDEX-02), within the LabEx PERSYVAL, and under the grant ANR-20-CE40-0013. Kir was supported by the Agence Nationale de la Recherche under grants ANR-22-PETQ-0008 (PQ-TLS). Kirimli was partially supported by EPSRC through grant number EP/V011324/1. Mahaney was supported by the Julian Chin Fellowship from the National Security Institute at Virginia Tech. Park was partially supported by DMS-2152182. Sabitova was supported in part by PSC-CUNY Award No. TRADB-56-58, jointly funded by The Professional Staff Congress and The City University of New York.

1.5. **AI disclosure.** An initial complete draft of the paper was written without AI usage, with the exception of the use of VSCode autocomplete for writing code to run computational experiments. Before releasing the paper, we uploaded the manuscript to Claude and ChatGPT for a final round of proofreading. In addition to some spelling and typesetting corrections, this process identified a logical gap in the proof of Proposition 4.22; we corrected this gap ourselves.

1.6. **Notation.** Here we collect any notation that is used in more than one section, together with either our running definition or convention, or a reference to the place where it is defined.

- ℓ is a prime number
- $\mathbb{F}_\ell$ is a finite field with ℓ elements
- $\overline{\mathbb{F}}_\ell$ is an algebraic closure of $\mathbb{F}_\ell$
- $\mathbb{Q}_\ell$ is the field of ℓ-adic numbers, with valuation ring $\mathbb{Z}_\ell$
- $\mathbb{Q}_{\ell^2}$ is the (unique up to isomorphism) quadratic unramified extension of $\mathbb{Q}_\ell$, with valuation ring $\mathbb{Z}_{\ell^2}$

- $\mathcal{A}$ is a definite quaternion algebra over $\mathbb{Q}$
- Λ is a lattice in $\mathcal{A}$; we write $\mathcal{O}$ for orders, I, J for invertible right $\mathcal{O}$-ideals, and $\mathfrak{O}$ for maximal orders.
- p is a prime number dividing $\operatorname{disc}(\mathcal{A})$ (Section 2.2)
- r is a prime number dividing $\operatorname{disc}(\mathcal{O})$ (Section 2.2) but not dividing $\operatorname{disc}(\mathcal{A})$
- $\mathcal{A}_\ell := \mathcal{A} \otimes_{\mathbb{Q}} \mathbb{Q}_\ell$ and $\Lambda_\ell := \Lambda \otimes_{\mathbb{Z}} \mathbb{Z}_\ell$ are completions at ℓ. One mild abuse of notation is that we also use subscripts $I_1, I_2, \dots$ for lists of invertible right $\mathcal{O}$-ideals, in which case we write $(I_i)_\ell$ for the completion. We avoid using both unindexed and indexed ideals in the same section, so the meaning of I_2, for instance, should be clear from context.
- h is the class number of $\mathcal{O}$ (Section 2.4)
- For $A \in M_n(R)$ (the ring of $n \times n$ matrices with entries in the commutative ring R), $A(i,j)$ is the entry of A in row i and column j
- $q = e^{2\pi i z}$ for z in the upper half plane of $\mathbb{C}$ (Eq. (14))
- $\langle \cdot , \cdot \rangle$: Eq. (1) if the inputs are in $\mathcal{A}$, Eq. (3) if the inputs are in $\mathfrak{M}(\mathcal{O})$
- $\lambda_n(u)$: the eigenvalue of an eigenvector $u \in \mathfrak{M}(\mathcal{O})$ under the action of B_n
- $u_1, \dots, u_h \in \mathfrak{M}(\mathcal{O})$: a normalized Brandt eigenbasis (start of Section 6)
- S-eigenvector: an eigenvector for every linear operator in S
- $N_{K/F}$ is the field norm of a finite field extension K/F
- $\operatorname{disc}(x) = \operatorname{trd}(x)^2 - 4\operatorname{nrd}(x)$, $x \in \mathcal{A}$
- Section 1.1: $G(p, \ell)$
- Section 2.1: $\Lambda^\sharp$, $\operatorname{Span}_R(C)$
- Section 2.2: $\overline{\alpha}$, $\operatorname{nrd}(\alpha)$, $\operatorname{trd}(\alpha)$, $\operatorname{nrd}(\Lambda)$, $\langle \alpha, \beta \rangle_\Lambda$
- Section 2.3: $R[C]$
- Section 2.4: $\mathcal{O}_L(I)$, $\mathcal{O}_R(I)$, $[I]$, $\operatorname{Gen}(I)$, $J \subseteq_n I$
- Section 2.5: $I_1 = \mathcal{O}, I_2, \dots, I_h$, $\mathfrak{M}(\mathcal{O})$, $v_1, \dots, v_h$, B_n, $\mathfrak{B}(\mathcal{O})$
- Section 3.2: $\mathfrak{b}_n(i,j)$, $\mathfrak{N}_n(J, I)$
- Section 3.3: $\mathfrak{N}_{\ell^n}(x, \mathcal{O}_\ell)$, $N_{\ell^n}(x)$
- Section 4.2: $\mathfrak{M}_{\mathrm{G}}(\mathcal{O})$, e, e'
- Section 4.3: $W_{\mathfrak{r}}$, E
- Section 5.1: $\mathcal{M}_2(\Gamma_0(N))$, $\mathcal{S}_2(\Gamma_0(N))$, $\mathcal{S}_2^{\mathrm{new}}(\Gamma_0(N))$, T_n, $\mathbb{T}$, $\mathbb{T}^{\mathrm{an}}$
- Section 5.2: $\theta_{\mathcal{O}}$, $a_n(\mathcal{O})$, Θ
- Corollary 6.3: g
- Proposition 6.5: $g_n(\mathcal{O})$

2. Background: quaternions

For more information on the topics in this section, the authors refer the reader to [Voi21]. Throughout we use F to refer to a field of characteristic 0, and $R \subseteq F$ a principal ideal domain with field of fractions F. We will specifically consider $F = \mathbb{Q}$ and $R = \mathbb{Z}$, or $F = \mathbb{Q}_\ell$ and $R = \mathbb{Z}_\ell$ for some prime ℓ, or $F = R = \mathbb{R}$.

2.1. **Lattices.** Let $V \simeq F^n$ be a finite-dimensional F-vector space. An R-submodule $\Lambda \subseteq V$ of rank n is called a **lattice over** R (or an "R-lattice"). In this case we have $\Lambda \otimes_R F = V$. In the case $R = \mathbb{Z}$ we simply call Λ a "lattice." Given a subset $C \subseteq \Lambda$, we write $\operatorname{Span}_R(C)$ for the R-lattice generated by C; that is, the minimal R-sublattice of Λ that contains C.

Let $\langle \cdot , \cdot \rangle : \Lambda \times \Lambda \to F$ be a nondegenerate symmetric bilinear form on Λ (which extends uniquely to a bilinear form on V by tensoring with F). We say that Λ is **integral** with respect to $\langle \cdot , \cdot \rangle$ if $\langle x, y \rangle \in R$ for all $x, y \in \Lambda$, and that Λ is **even integral** with respect to $\langle \cdot , \cdot \rangle$ if $\langle x, x \rangle \in 2R$ for all $x \in \Lambda$. Note that any even integral lattice is integral because $2\langle x, y \rangle = \langle x+y, x+y \rangle - \langle x, x \rangle - \langle y, y \rangle$. If Λ and Λ' are R-lattices with respective bilinear forms $\langle \cdot , \cdot \rangle$ and $\langle \cdot , \cdot \rangle'$, we say a map $\phi : \Lambda \to \Lambda'$ is an **isometry** if it is a bijection with $\langle \phi(x), \phi(y) \rangle' = \langle x, y \rangle$ for all $x, y \in \Lambda$.

Definition 2.1. Given an R-lattice $\Lambda \subset V$ equipped with a nondegenerate symmetric bilinear form $\langle \cdot, \cdot \rangle$, the **dual lattice** of Λ (with respect to $\langle \cdot, \cdot \rangle$) is

$$\Lambda^\sharp = \{\alpha \in V \;:\; \langle \alpha, x \rangle \in R \text{ for all } x \in \Lambda\}.$$

If $R = \mathbb{Z}$ or $R = \mathbb{Z}_\ell$ for a prime ℓ, we define the **level** of Λ (also called the "Stufe" by some authors) to be the minimal positive integer N such that $\Lambda^\sharp$ is an even integral lattice with respect to the bilinear form $(x, y) \mapsto N\langle x, y \rangle$.

Let V be a vector space over $\mathbb{Q}$. For a lattice $\Lambda \subset V$ and a prime $\ell \in \mathbb{Z}$, we define the **completion** of Λ at ℓ to be the $\mathbb{Z}_\ell$-lattice $\Lambda_\ell := \Lambda \otimes_{\mathbb{Z}} \mathbb{Z}_\ell$. A bilinear form $\Lambda \times \Lambda \to \mathbb{Q}$ induces a bilinear form $\Lambda_\ell \times \Lambda_\ell \to \mathbb{Q}_\ell$ for each prime ℓ. The **local-global dictionary** for lattices in V states that given any fixed lattice $\Lambda^0 \subset V$, there is a bijection between lattices $\Lambda \subset V$ and tuples of lattices (Λ_ℓ) over primes ℓ that satisfy $\Lambda_\ell = \Lambda^0_\ell$ for all but finitely many ℓ [Voi21, Theorem 9.1.1, Lemma 9.5.3].[2] A property of a lattice Λ is **local** if it holds for Λ if and only if it holds for Λ_ℓ for all primes ℓ. For instance, being an integral or an even integral lattice is a local property. The dual and level of a lattice can be computed locally, in the sense that they are uniquely determined by the duals and levels of the completions at all primes.

We finally note one elementary but useful property about lattices that will be used to check whether a certain set generates the lattice.

Lemma 2.2. *Let Λ be a lattice, $\Lambda' \subseteq \Lambda$ a sublattice, and ℓ a prime. If Λ' surjects onto $\Lambda/\ell\Lambda$ under reduction mod $\ell\Lambda$ then $\Lambda'_\ell = \Lambda_\ell$.*

Proof. Follows by Nakayama's Lemma or by a matrix determinant computation. □

2.2. **Lattices in quaternion algebras.** A **quaternion algebra** $\mathcal{A}$ over F is a (noncommutative) central simple algebra of rank 4 over F. All quaternion algebras over F can be given an F-basis $\{1, \mathrm{i}, \mathrm{j}, \mathrm{ij}\}$ such that $\mathrm{i}^2 = a, \mathrm{j}^2 = b, \mathrm{ij} = -\mathrm{ji}$ for some nonzero $a, b \in F$. With respect to such a basis we can define the **standard involution** on $\mathcal{A}$ by $\overline{t + x\mathrm{i} + y\mathrm{j} + z\mathrm{ij}} = t - x\mathrm{i} - y\mathrm{j} - z\mathrm{ij}$; we then define the **(reduced) trace** of a quaternion element $\alpha \in \mathcal{A}$ as $\mathrm{trd}(\alpha) = \alpha + \overline{\alpha}$ and the **(reduced) norm** as $\mathrm{nrd}(\alpha) = \alpha\overline{\alpha}$. Since no other notion of trace or norm appears in this paper, we will usually drop the word "reduced" for convenience.

Using the trace, we have a nondegenerate symmetric bilinear form on $\mathcal{A}$,

$$\langle \alpha, \beta \rangle := \mathrm{trd}(\alpha\overline{\beta}) = \mathrm{nrd}(\alpha + \beta) - \mathrm{nrd}(\alpha) - \mathrm{nrd}(\beta), \tag{1}$$

which we call the **trace form** on $\mathcal{A}$. This bilinear form induces a quadratic form $Q\colon \mathcal{A} \to F$ given by the reduced norm, $Q(\alpha) = \mathrm{nrd}(\alpha) = \frac{1}{2}\langle \alpha, \alpha \rangle$.

Now suppose $\mathcal{A}$ is a quaternion algebra over $\mathbb{Q}$. For any field extension $F/\mathbb{Q}$, $\mathcal{A} \otimes_{\mathbb{Q}} F$ is a quaternion algebra over F. We say $\mathcal{A}$ is **definite** if $\mathcal{A} \otimes_{\mathbb{Q}} \mathbb{R}$ is isomorphic to a division algebra over $\mathbb{R}$ (the classical Hamilton quaternions). Otherwise $\mathcal{A}$ is **indefinite**, in which case $\mathcal{A} \otimes_{\mathbb{Q}} \mathbb{R} \cong M_2(\mathbb{R})$. In a similar way, we say a prime $p \in \mathbb{Z}$ **ramifies** in $\mathcal{A}$ if the algebra $\mathcal{A} \otimes_{\mathbb{Q}} \mathbb{Q}_p$ is a division algebra, otherwise p **splits** in $\mathcal{A}$ and $\mathcal{A} \otimes_{\mathbb{Q}} \mathbb{Q}_p \cong M_2(\mathbb{Q}_p)$ is a matrix algebra. The **discriminant** $\mathrm{disc}(\mathcal{A})$ of $\mathcal{A}$ is the product of primes that ramify in $\mathcal{A}$.

If $\Lambda \subseteq \mathcal{A}$ is a lattice, the set $\{\mathrm{nrd}(x) : x \in \Lambda\} \subseteq \mathbb{Q}$ generates a finitely generated $\mathbb{Z}$-module, which is cyclic because $\mathbb{Z}$ is a principal ideal domain. We define the **norm** of Λ, $\mathrm{nrd}(\Lambda)$, to be the positive generator of this $\mathbb{Z}$-module.

Definition 2.3. Given a lattice Λ in a quaternion algebra $\mathcal{A}$ over $\mathbb{Q}$, we define the **normalized trace form** on Λ by

$$\langle \alpha, \beta \rangle_\Lambda := \frac{\mathrm{trd}(\alpha\overline{\beta})}{\mathrm{nrd}(\Lambda)}, \qquad \alpha, \beta \in \Lambda.$$

[2] All lattices in V have the same completion at the infinite place, $\Lambda \otimes \mathbb{R} = V \otimes \mathbb{R}$, so it suffices to consider the finite places.

For any $\alpha \in \Lambda$, $\operatorname{trd}(\alpha\bar{\alpha}) = 2\operatorname{nrd}(\alpha)$ is an even integer multiple of $\operatorname{nrd}(\Lambda)$, so the normalized trace form equips Λ with the structure of an even integral lattice. Given a prime ℓ, we may equip $\Lambda_\ell = \Lambda \otimes_{\mathbb{Z}} \mathbb{Z}_\ell$ with the bilinear form induced from the normalized trace form on Λ.

Lemma 2.4. *If (b_1, b_2, b_3, b_4) is a basis for a lattice Λ in a quaternion algebra $\mathcal{A}$ over $\mathbb{Q}$, then*

$$\det(\langle b_i, b_j\rangle_\Lambda)_{1\le i,j\le 4} \in \mathbb{Z} \tag{2}$$

is a perfect square that does not depend on the choice of basis.

Proof. This value is an integer because Λ is an integral lattice with respect to $\langle\cdot,\cdot\rangle_\Lambda$, so it suffices to prove that it is a square in $\mathbb{Q}$. Let $\Lambda_0 \subset \mathcal{A}$ be the lattice with basis $(b_1', b_2', b_3', b_4') = (1, \mathrm{i}, \mathrm{j}, \mathrm{ij})$. There exists a positive integer m such that $m\Lambda \subseteq \Lambda_0$, and

$$\begin{aligned}
\det(\langle b_i, b_j\rangle_\Lambda)_{1\le i,j\le 4} &= \frac{1}{m^8 \operatorname{nrd}(\Lambda)^4} \det(\operatorname{trd}(mb_i\overline{mb_j}))_{1\le i,j\le 4}\\
&= \frac{[\Lambda_0 : m\Lambda]^2}{m^8 \operatorname{nrd}(\Lambda)^4} \det(\operatorname{trd}(b_i'\overline{b}_j'))_{1\le i,j\le 4}\\
&= \frac{[\Lambda_0 : m\Lambda]^2}{m^8 \operatorname{nrd}(\Lambda)^4} (4ab)^2,
\end{aligned}$$

with the second equality holding by the proof of [Voi21, Lemma 15.2.5]. □

The **(reduced) discriminant** of Λ, denoted $\operatorname{disc}(\Lambda) \in \mathbb{Z}$, is the positive square root of the quantity (2) in Lemma 2.4.

2.3. **Quaternion orders.** Let $\mathcal{A}$ be a quaternion algebra over F. An **order** in $\mathcal{A}$ is an R-lattice $\mathcal{O} \subset \mathcal{A}$ (Section 2.1) that contains 1 and is closed under multiplication. For any $x \in \mathcal{O}$ we have $\operatorname{nrd}(x), \operatorname{trd}(x) \in R$. Given a subset $C \subseteq \mathcal{O}$, we write $R[C]$ for the R-algebra generated by C; that is, the minimal suborder of $\mathcal{O}$ that contains C.

Let $\mathcal{O}$ be an order in a quaternion algebra $\mathcal{A}$ over $\mathbb{Q}$. Since $\operatorname{nrd}(\mathcal{O}) = 1$, the normalized trace form on $\mathcal{O}$ agrees with the trace form on $\mathcal{A}$. In particular, any order $\mathcal{O}$ is an even integral lattice with respect to the trace form. We call the dual of $\mathcal{O}$ with respect to the trace form the **codifferent** $\mathcal{O}^\sharp$ of $\mathcal{O}$. Note that $\mathcal{O}^\sharp \supseteq \mathcal{O}$.

We note here an equivalent definition of level, specific for orders in quaternion algebras over $\mathbb{Q}$ (this is the definition given in [Brz83] and [Eic55], for example).

Lemma 2.5. *For an order $\mathcal{O}$ in a quaternion algebra $\mathcal{A}$ over $\mathbb{Q}$, the level of $\mathcal{O}$ as a lattice (with respect to the trace form) is equal to $\operatorname{nrd}(\mathcal{O}^\sharp)^{-1}$.*

Proof. For any integer n, the bilinear form $(x, y) \mapsto n\operatorname{trd}(x\overline{y})$ gives $\mathcal{O}^\sharp$ the structure of an even integral lattice if and only if $n \operatorname{nrd}(x) \in \mathbb{Z}$ for all $x \in \mathcal{O}^\sharp$. This holds if and only if

$$n \operatorname{nrd}(\mathcal{O}^\sharp)\mathbb{Z} = n\{\operatorname{nrd}(x) : x \in \mathcal{O}^\sharp\} \subseteq \mathbb{Z},$$

which in turn holds if and only if $n \in \operatorname{nrd}(\mathcal{O}^\sharp)^{-1}\mathbb{Z}$. Since $\mathcal{O}^\sharp$ contains 1, $\operatorname{nrd}(1) = 1$ is an integer multiple of $\operatorname{nrd}(\mathcal{O}^\sharp)$ by definition, so $\operatorname{nrd}(\mathcal{O}^\sharp)$ must be the reciprocal of a positive integer. Thus $N = \operatorname{nrd}(\mathcal{O}^\sharp)^{-1}$ is the smallest positive integer in $\operatorname{nrd}(\mathcal{O}^\sharp)^{-1}\mathbb{Z}$. □

Definition 2.6. Let $\mathcal{O}$ be an order in a quaternion algebra $\mathcal{A}$ over F. We say $\mathcal{O}$ is **maximal** if it is not strictly contained in another order of $\mathcal{A}$. We say $\mathcal{O}$ is **Eichler** if $\mathcal{O} = \mathcal{O}_1 \cap \mathcal{O}_2$ is an intersection of two (not necessarily distinct) maximal orders $\mathcal{O}_1$ and $\mathcal{O}_2$.

For a lattice Λ in a quaternion algebra $\mathcal{A}$ over $\mathbb{Q}$, the properties of being an order, a maximal order, or an Eichler order are all local properties. The discriminant and norm of Λ can be computed locally. For an order $\mathcal{O} \subseteq \mathcal{A}$, the (reduced) discriminant of $\mathcal{O}$ is equal to $[\mathfrak{O} : \mathcal{O}] \operatorname{disc}(\mathcal{A})$, where $\mathfrak{O}$ is any maximal order containing $\mathcal{O}$.

2.4. **Ideals and ideal classes.** In this paper we work primarily with right orders and right ideals, but all definitions given in this section have analogous left-sided versions.

Let $\mathcal{A}$ be a quaternion algebra over F. The **right order** of an R-lattice I in $\mathcal{A}$ is the order $\mathcal{O}_{\mathrm{R}}(I) = \{\alpha \in \mathcal{A} \mid I\alpha \subseteq I\}$. If $\mathcal{O}$ is an order, we say that a lattice I is a **right $\mathcal{O}$-ideal** if $\mathcal{O}_{\mathrm{R}}(I) \supseteq \mathcal{O}$.

Now suppose $\mathcal{O}$ is an order in a quaternion algebra $\mathcal{A}$ over $\mathbb{Q}$. A right $\mathcal{O}$-ideal I is **invertible** if I is a locally principal right $\mathcal{O}$-ideal, i.e. for all primes ℓ there is an $\alpha_\ell \in \mathcal{A}_\ell^\times$ such that $I_\ell = \alpha_\ell \mathcal{O}_\ell$. This is equivalent to the existence of a lattice J such that $JI = \mathcal{O}_{\mathrm{R}}(I) = \mathcal{O}$ and $IJ = \mathcal{O}_{\mathrm{L}}(I)$, where $\mathcal{O}_{\mathrm{L}}(I)$ is the left order of I.

Definition 2.7. Let $\mathcal{O}$ be an order in a quaternion algebra $\mathcal{A}$ over $\mathbb{Q}$, and suppose that I is an invertible right $\mathcal{O}$-ideal. We say that J is *contained in* I *with relative norm* n, or $J \subseteq_n I$, if J is an invertible right $\mathcal{O}$-ideal with $J \subseteq I$ and $\mathrm{nrd}(J) = n\,\mathrm{nrd}(I)$.

Two right $\mathcal{O}$-ideals I, J are in the same **right ideal class**, written $[I] = [J]$, if there exists $\alpha \in \mathcal{A}^\times$ such that $I = \alpha J$. Ideals in the same right ideal class are isometric as lattices equipped with their associated normalized trace forms (Definition 2.3): since $\mathrm{trd}(ab) = \mathrm{trd}(ba)$ for $a, b \in \mathcal{A}$, for any $\alpha \in \mathcal{A}^\times$ and $x, y \in I$ we have:

$$\langle \alpha x, \alpha y\rangle_{\alpha I} = \frac{\mathrm{trd}((\alpha x)(\bar{y}\bar{\alpha}))}{\mathrm{nrd}(\alpha I)} = \frac{\mathrm{trd}(\bar{\alpha}\alpha x\bar{y})}{\mathrm{nrd}(\alpha)\mathrm{nrd}(I)} = \frac{\mathrm{trd}(x\bar{y})}{\mathrm{nrd}(I)} = \langle x, y\rangle_I.$$

In particular $x \mapsto \alpha x$ is an isometry from I to αI.

We say that two invertible right $\mathcal{O}$-ideals I, J are in the same **genus**, and write $J \in \mathrm{Gen}(I)$, if the completions I_ℓ and J_ℓ, with bilinear forms induced by the normalized trace forms on I and J respectively, are isometric as $\mathbb{Z}_\ell$-lattices for every prime ℓ. Ideals in the same class are automatically in the same genus.

The **right ideal class set** of $\mathcal{O}$ is the set $\mathrm{Cls}(\mathcal{O})$ consisting of equivalence classes of invertible right $\mathcal{O}$-ideals. It is well known (see e.g. [Voi21, Theorem 17.1.1]) that the right ideal class set of an order in a quaternion algebra over $\mathbb{Q}$ (more generally over any number field) is finite, and we call its cardinality the **class number** of $\mathcal{O}$.

2.5. **Brandt modules and algebras.** Let $\mathcal{O}$ be an order in a definite quaternion algebra over $\mathbb{Q}$. Let $I_1, \ldots, I_h$ be a list of representatives of the right ideal classes of $\mathcal{O}$, where h is the class number of $\mathcal{O}$ and each I_j is an *invertible* right $\mathcal{O}$-ideal. Without loss of generality we may take $I_1 = \mathcal{O}$. We define the (real) **Brandt module**

$$\mathfrak{M}(\mathcal{O}) := \{f : \{I_1, \ldots, I_h\} \to \mathbb{R}\}.$$

Thus $\mathfrak{M}(\mathcal{O}) \simeq \mathbb{R}^h$, and we have a fixed standard basis $v_1, \ldots, v_h$ given by Kronecker delta functions, $v_i(I_j) = \delta_{ij}$. We usually think of $\mathfrak{M}(\mathcal{O})$ as column vectors of length h with entries in $\mathbb{R}$.

The Brandt module comes equipped with an inner product for which the standard basis is orthogonal. For each i, let $\mathcal{O}_{\mathrm{L}}(I_i)$ denote the left order of I_i, and let $w_i := |\mathcal{O}_{\mathrm{L}}(I_i)^\times/\{\pm 1\}|$. Then we define

$$\langle v_i, v_j\rangle = \begin{cases} w_i & \text{if } i = j, \\ 0 & \text{otherwise.}\end{cases} \tag{3}$$

The Brandt module also comes equipped with matrices B_n for all $n \geq 0$ acting on the left on $\mathfrak{M}(\mathcal{O})$, defined entry-wise with respect to the standard basis $\{v_1, \ldots, v_h\}$ by

$$\begin{aligned} B_n(i,j) &:= \frac{1}{2w_i}\left|\left\{\alpha \in I_j I_i^{-1} : \mathrm{nrd}(\alpha)\frac{\mathrm{nrd}(I_i)}{\mathrm{nrd}(I_j)} = n\right\}\right| \\ &= \frac{1}{2w_i}\left|\left\{\alpha \in I_j I_i^{-1} : \tfrac{1}{2}\langle \alpha, \alpha\rangle_{I_j I_i^{-1}} = n\right\}\right|, \end{aligned} \tag{4}$$

where in the second line we use the normalized trace form on $I_j I_i^{-1}$ (Definition 2.3). For $n \geq 1$ we have the equivalent form

$$B_n(i,j) = \left|\{J \subseteq_n I_j : [J] = [I_i]\}\right| \tag{5}$$

by [Voi21, Lemma 41.2.7(a)], where $J \subseteq_n I_j$ means that J is an invertible right $\mathcal{O}$-ideal with $\mathrm{nrd}(J) = n\,\mathrm{nrd}(I_j)$ (Definition 2.7). Note that B_1 is the identity matrix, and B_0 is the matrix where every entry in row i is $\frac{1}{2w_i}$.

Definition 2.8. The **Brandt algebra** $\mathfrak{B}(\mathcal{O})$ is the $\mathbb{Z}$-subalgebra of $\mathrm{End}(\mathfrak{M}(\mathcal{O})) \cong M_{h\times h}(\mathbb{R})$ generated by the Brandt matrices $\{B_n : n \geq 0\}$.

Lemma 2.9. *For all $n \geq 0$, B_n is self-adjoint with respect to the inner product.*

Proof. Let D be the diagonal matrix with $D(i,i) = w_i$, so the inner product is given by $\langle v, w\rangle = v^t D w$. The self-adjointness of B_n with respect to $\langle v, w\rangle$ is equivalent to checking that $DB_n = B_n^t D$. Since $B_n^t D = (DB_n)^t$, it suffices to check that DB_n is a symmetric matrix. We have

$$(DB_n)(i,j) = w_i B_n(i,j) = \frac{1}{2}\left|\left\{\alpha \in I_j I_i^{-1} : \tfrac{1}{2}\langle \alpha, \alpha\rangle_{I_j I_i^{-1}} = n\right\}\right|,$$
$$(DB_n)(j,i) = w_j B_n(j,i) = \frac{1}{2}\left|\left\{\alpha \in I_i I_j^{-1} : \tfrac{1}{2}\langle \alpha, \alpha\rangle_{I_i I_j^{-1}} = n\right\}\right|.$$

Since I_i and I_j are invertible right $\mathcal{O}$-ideals, we have $\overline{I_i} I_i = \mathrm{nrd}(I_i)\mathcal{O}$ and $\overline{I_j} I_j = \mathrm{nrd}(I_j)\mathcal{O}$ by [Voi21, 16.6.14], so $\alpha \mapsto \overline{\alpha}\frac{\mathrm{nrd}(I_i)}{\mathrm{nrd}(I_j)}$ defines an isometry from $I_j I_i^{-1}$ to $I_i I_j^{-1}$, each equipped with their normalized trace forms. This proves that the two sets above have the same size. $\square$

3. Commutativity of the Brandt algebra

Let $\mathcal{O}$ be an order in a definite rational quaternion algebra $\mathcal{A}$. Recall from Definition 2.8 that $\mathfrak{B}(\mathcal{O})$ is the $\mathbb{Z}$-algebra generated by Brandt matrices B_n for all $n \geq 0$. The main result of this section is the following.

Theorem 3.1. *For an order $\mathcal{O}$ in a rational definite quaternion algebra, $\mathfrak{B}(\mathcal{O})$ is commutative.*

We begin with a few notes on the subtleties underlying this result, and why it does not directly follow from certain well-known results in the literature on establishing commutativity of Hecke algebras. The proof begins in Section 3.2.

3.1. **Failure of commutativity in Hecke algebras.** We first recall that commutativity of $\mathfrak{B}(\mathcal{O})$ is well known for maximal orders. More generally, it follows immediately from [Eic73, Chapter II Theorem 2] that for an arbitrary order $\mathcal{O}$ contained in a maximal order $\mathfrak{O}$, the Brandt matrices B_n with n coprime to $[\mathfrak{O} : \mathcal{O}]$ generate a commutative subalgebra. However, for this paper we will need to use a stronger result, namely the full Brandt algebra is commutative for arbitrary *submaximal* orders (Definition 4.1).

For an arbitrary rational quaternion order $\mathcal{O}$, the Brandt algebra $\mathfrak{B}(\mathcal{O})$ may be naturally interpreted as a subalgebra of a *quaternionic Hecke algebra* $\mathfrak{H}(\mathcal{O})$. One way to define this Hecke algebra is to consider the adelic counterparts $\widehat{\mathcal{A}}$ and $\widehat{\mathcal{O}}$ of the quaternion algebra $\mathcal{A}$ and order $\mathcal{O}$, and define $\mathfrak{H}(\mathcal{O})$ to be the convolution algebra of functions on $\widehat{\mathcal{A}}^\times$ that are bi-invariant under $\widehat{\mathcal{O}}^\times$. This structure arises, for instance, in the Jacquet-Langlands correspondence, which compares $\mathfrak{H}(\mathcal{O})$ to the algebra of Hecke operators for GL_2. One may identify $\mathfrak{B}(\mathcal{O})$ with a subalgebra of $\mathfrak{H}(\mathcal{O})$ by associating each Brandt operator B_n with the characteristic function of elements of norm n in $\widehat{\mathcal{A}}^\times$.

If $\mathcal{O}$ is maximal then $\mathfrak{H}(\mathcal{O}) = \mathfrak{B}(\mathcal{O})$ is commutative, but for a non-maximal order, the full Hecke algebra is *not* commutative in general. Theorem 3.1 shows nonetheless that $\mathfrak{B}(\mathcal{O})$ is a commutative subalgebra of $\mathfrak{H}(\mathcal{O})$ for an arbitrary $\mathcal{O}$.

There are many well-established techniques for proving commutativity of subalgebras of Hecke algebras. One of these methods, often called "Gelfand's trick," can be used to prove commutativity of any subalgebra generated by double cosets of $\widehat{\mathcal{O}}^\times$ that are preserved by an involutive anti-automorphism (such as the standard involution on $\widehat{\mathcal{A}}$). See [Shi71, Proposition 3.8] for a general version of this method. Unfortunately, while the set of elements of norm n is stable under the standard involution, the individual $\widehat{\mathcal{O}}^\times$-double cosets comprising this set are not, in general.

The construction in the following proposition shows that there is no direct analogue of Gelfand's trick in this scenario. That is, even if a subalgebra of $\mathfrak{H}(\mathcal{O})$ is generated by operators that are each stable under the standard involution, the subalgebra may be non-commutative. Thus any proof of commutativity of $\mathfrak{B}(\mathcal{O})$ must rely on properties of Brandt matrices beyond just stability under the standard involution.

Proposition 3.2. *Let $\epsilon \geq 2$, ℓ a split prime of $\mathcal{A}$, and $\mathcal{O} \subseteq \mathcal{A}$ an Eichler order of index ℓ^ϵ in a maximal order. Then there exist two operators $A, B \in \mathfrak{H}(\mathcal{O})$ satisfying $\overline{A} = A$ and $\overline{B} = B$, but with $AB \neq BA$.*

Proof. By standard arguments we may reduce to considering the local Hecke algebra at ℓ. Without loss of generality we may take $\mathcal{O}_\ell$ to be the standard Eichler $\mathbb{Z}_\ell$-order of level ℓ^ϵ:

$$\mathcal{O}_\ell = \left\{ \begin{pmatrix} a & b \\ \ell^\epsilon c & d \end{pmatrix} : a, b, c, d \in \mathbb{Z}_\ell \right\},$$

with $\mathcal{O}_\ell^\times$ consisting of those matrices in $\mathcal{O}_\ell$ with $a, d \in \mathbb{Z}_\ell^\times$. Let

$$\alpha = \begin{pmatrix} \ell^2 & 1 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad \beta = \begin{pmatrix} \ell & 0 \\ 0 & 1 \end{pmatrix}.$$

We will check by direct computation that

$$\overline{\beta}\alpha = \begin{pmatrix} \ell^2 & 1 \\ 0 & \ell \end{pmatrix} \notin \left(\mathcal{O}_\ell^\times \alpha \mathcal{O}_\ell^\times \cup \mathcal{O}_\ell^\times \overline{\alpha} \mathcal{O}_\ell^\times \right) \cdot \left(\mathcal{O}_\ell^\times \beta \mathcal{O}_\ell^\times \cup \mathcal{O}_\ell^\times \overline{\beta} \mathcal{O}_\ell^\times \right). \tag{6}$$

We first compute the double coset $\mathcal{O}_\ell^\times \overline{\beta}\alpha \mathcal{O}_\ell^\times$. For $b, c, f, g \in \mathbb{Z}_\ell$ and $a, d, e, h \in \mathbb{Z}_\ell^\times$, we perform the following computation, using the assumption $\epsilon \geq 2$:

$$\begin{pmatrix} a & b \\ \ell^\epsilon c & d \end{pmatrix} \begin{pmatrix} \ell^2 & 1 \\ 0 & \ell \end{pmatrix} \begin{pmatrix} e & f \\ \ell^\epsilon g & h \end{pmatrix} \equiv \begin{pmatrix} 0 & ah + bh\ell \\ 0 & dh\ell \end{pmatrix} \pmod{\ell^2 M_2(\mathbb{Z}_\ell)}.$$

In particular, since $a, d, e, h \in \mathbb{Z}_\ell^\times$, the top-right entry has valuation 0 and the bottom-right entry has valuation 1. (A similar computation can be used to verify that $\overline{\alpha} \notin \mathcal{O}_\ell^\times \alpha \mathcal{O}_\ell^\times$ and $\overline{\beta} \notin \mathcal{O}_\ell^\times \beta \mathcal{O}_\ell^\times$.) On the other hand, setting $u = \begin{pmatrix} w & x \\ \ell^\epsilon y & z \end{pmatrix}$ with $x, y \in \mathbb{Z}_\ell$ and $w, z \in \mathbb{Z}_\ell^\times$, we have the following congruences modulo $\ell^2 M_2(\mathbb{Z}_\ell)$:

$$\alpha u \beta \equiv \begin{pmatrix} 0 & z \\ 0 & z \end{pmatrix}, \qquad \alpha u \overline{\beta} \equiv \begin{pmatrix} 0 & z\ell \\ 0 & z\ell \end{pmatrix},$$

$$\overline{\alpha} u \beta \equiv \begin{pmatrix} w\ell & x - z \\ 0 & 0 \end{pmatrix}, \qquad \overline{\alpha} u \overline{\beta} \equiv \begin{pmatrix} w & (x - z)\ell \\ 0 & 0 \end{pmatrix}.$$

No matter what $u \in \mathcal{O}_\ell^\times$ we take, none of these elements have a top-right entry of valuation 0 and a bottom-right entry of valuation 1, so they do not lie in $\mathcal{O}_\ell^\times \overline{\beta}\alpha \mathcal{O}_\ell^\times$. We can conclude that

$$\overline{\beta}\alpha \notin \mathcal{O}_\ell^\times \alpha \mathcal{O}_\ell^\times \beta \mathcal{O}_\ell^\times \cup \mathcal{O}_\ell^\times \alpha \mathcal{O}_\ell^\times \overline{\beta} \mathcal{O}_\ell^\times \cup \mathcal{O}_\ell^\times \overline{\alpha} \mathcal{O}_\ell^\times \beta \mathcal{O}_\ell^\times \cup \mathcal{O}_\ell^\times \overline{\alpha} \mathcal{O}_\ell^\times \overline{\beta} \mathcal{O}_\ell^\times,$$

establishing Eq. (6).

Now take A and B to be the characteristic functions of the subsets $\mathcal{O}_\ell^\times \alpha \mathcal{O}_\ell^\times \cup \mathcal{O}_\ell^\times \overline{\alpha} \mathcal{O}_\ell^\times$ and $\mathcal{O}_\ell^\times \beta \mathcal{O}_\ell^\times \cup \mathcal{O}_\ell^\times \overline{\beta} \mathcal{O}_\ell^\times$ of $\mathcal{A}_\ell^\times$, respectively. Observe that these sets are stable under the standard involution, so $\overline{A} = A$ and $\overline{B} = B$. The convolution product BA evaluates to a nonzero value at $\overline{\beta}\alpha$, but by Eq. (6), AB evaluates to 0 at $\overline{\beta}\alpha$. This proves that $AB \neq BA$. □

3.2. **Reducing to a local computation.** We return to the proof of Theorem 3.1.

Lemma 3.3. *For $n \geq 1$ and $1 \leq j \leq h$, we have*

$$\sum_{k=1}^{h} B_n(k,j) = \left|\{J \subseteq_n I_j\}\right|,$$

and this value is independent of j.

Proof. The equality follows from the definition of $B_n(k,j)$, since every invertible right $\mathcal{O}$-ideal belongs to a unique right ideal class $[I_k]$ for some $1 \leq k \leq h$. Since each invertible right $\mathcal{O}$-ideal I_j is locally principal, the independence from j follows from the local-global dictionary (cf. Section 2.1). □

Lemma 3.4. *B_0 commutes with B_n for every $n \geq 1$.*

Proof. Recall that $B_0(i,j) = \frac{1}{2w_i}$, for any i,j. Since B_n is self-adjoint by Lemma 2.9, we have

$$\begin{aligned}(B_nB_0)(i,j) &= \sum_{k=1}^{h} \frac{1}{2w_k} B_n(i,k) = \sum_{k=1}^{h} \frac{1}{2w_i} B_n(k,i),\\ (B_0B_n)(i,j) &= \sum_{k=1}^{h} \frac{1}{2w_i} B_n(k,j).\end{aligned}$$

Thus $(B_nB_0)(i,j) = (B_0B_n)(i,j)$ by Lemma 3.3. □

Given Lemma 3.4, we can restrict our attention to the Brandt matrices B_n with $n \geq 1$. For $n \geq 1$, define

$$\mathfrak{b}_n(i,j) = \left\{J \subseteq_n I_j \ : \ [J] = [I_i]\right\}.$$

Then $B_n(i,j) = |\mathfrak{b}_n(i,j)|$ by Eq. (5).

We also define

$$\mathfrak{N}_n(J,I) := \left\{J' : J \subseteq J' \subseteq_n I\right\},$$

as the set of invertible right $\mathcal{O}$-ideals J' containing J which are contained in I with relative norm n (Definition 2.7).

Proposition 3.5. *For any $i,j \in \{1,\dots,h\}$ and $m,n \geq 1$,*

$$(B_mB_n)(i,j) = \sum_{J \in \mathfrak{b}_{mn}(i,j)} \left|\mathfrak{N}_n(J,I_j)\right|.$$

Proof. The number $(B_mB_n)(i,j) = \sum_{k=1}^{h} B_m(i,k)B_n(k,j)$ equals the sum over all k of the sizes of the product sets $\mathfrak{b}_m(i,k) \times \mathfrak{b}_n(k,j)$. First we fix $1 \leq k \leq h$ as well as $J' \in \mathfrak{b}_n(k,j)$. Pick $\alpha_{J'} \in \mathcal{A}^\times$ with $J' = \alpha_{J'} I_k$. The map $J'' \mapsto \alpha_{J'} J''$ then induces a bijection

$$\mathfrak{b}_m(i,k) = \left\{J'' \subseteq_m I_k : [J''] = [I_i]\right\} \quad \to \quad \left\{J \subseteq_m J' : [J] = [I_i]\right\}.$$

Now we still keep k fixed but allow J' to vary: the map $(J'', J') \mapsto (\alpha_{J'} J'', J')$ defines a bijection from $\mathfrak{b}_m(i,k) \times \mathfrak{b}_n(k,j)$ to

$$\left\{(J, J') : J \subseteq_m J' \subseteq_n I_j,\ [J] = [I_i],\ [J'] = [I_k]\right\},$$

where J and J' both vary over invertible right $\mathcal{O}$-ideals. If we now take the union over all k, we find that

$$\left\{(J, J') : J \subseteq_m J' \subseteq_n I_j,\ [J] = [I_i]\right\}$$

has size $(B_m B_n)(i,j)$. Writing this final set as a union over $J \in \mathfrak{b}_{mn}(i,j)$ gives us the desired result. $\square$

If i, j, and $J \in \mathfrak{b}_{mn}(i,j)$ are fixed, the set $\mathfrak{N}_n(J, I_j)$ can be determined locally. In particular we can use this to prove the following.

Lemma 3.6 ([Voi21, Proposition 41.3.1])**.** *If* $\gcd(m,n) = 1$ *then* $B_m B_n = B_n B_m = B_{mn}$.

Proof. For each $J \in \mathfrak{b}_{mn}(i,j)$, there is a unique $J \subseteq J' \subseteq_n I_j$, defined by setting $J'_\ell = J_\ell$ for $\ell \mid n$ and $J'_\ell = (I_j)_\ell$ for $\ell \nmid n$. So $B_n B_m = B_{mn}$ by Proposition 3.5. $\square$

Since we have also checked that B_0 commutes with every Brandt matrix, this allows us to make the following reduction.

Lemma 3.7. $\mathfrak{B}(\mathcal{O})$ *is commutative if and only if the* $\mathbb{Z}$*-subalgebra of* $\mathfrak{B}(\mathcal{O})$ *generated by* $\{B_{\ell^k} : k \geq 0\}$ *is commutative for every prime* ℓ.

3.3. **Local analysis.** From now on we fix a prime ℓ and we denote $\mathcal{O}_\ell = \mathcal{O} \otimes_{\mathbb{Z}} \mathbb{Z}_\ell$. For $a \in \mathbb{Z}_\ell$ we let $\mathrm{val}(a)$ denote the ℓ-adic valuation, so that $a \in \ell^{\mathrm{val}(a)} \mathbb{Z}_\ell^\times$. We declare $\mathrm{val}(0) = \infty$.

For $x \in \mathcal{O}_\ell$ and $m \in \mathbb{N}$, let $\mathfrak{N}_{\ell^m}(x) = \mathfrak{N}_{\ell^m}(x, \mathcal{O}_\ell)$ denote the set of principal right ideals $I \subset \mathcal{O}_\ell$ of norm ℓ^m with $x \in I$. Let $N_{\ell^m}(x) = N_{\ell^m}(x, \mathcal{O}_\ell) = |\mathfrak{N}_{\ell^m}(x, \mathcal{O}_\ell)|$ be the cardinality of $\mathfrak{N}_{\ell^m}(x, \mathcal{O}_\ell)$.

The idea of the proof of Theorem 3.1 is to relate the sets $\mathfrak{N}_{\ell^m}(x, \mathcal{O}_\ell)$ of local ideals to the sets $\mathfrak{N}_{\ell^m}(J, I_j)$ of global ideals that appear in Proposition 3.5, justifying the abuse of notation. We will prove the identity $N_{\ell^m}(x) = N_{\ell^m}(\bar{x})$ for all $x \in \mathcal{O}_\ell$ with $\mathrm{nrd}(x) \neq 0$ (Proposition 3.10), and use this to prove commutativity of Brandt matrices. We begin with some simple relations satisfied by $N_{\ell^m}(x)$.

Lemma 3.8. *Let* $x, x' \in \mathcal{O}_\ell$ *and* $m \in \mathbb{N}$.
(1) *If* $x - x' \in \ell^m \mathcal{O}_\ell$, *then* $N_{\ell^m}(x) = N_{\ell^m}(x')$.
(2) *If* $n = \mathrm{val}(\mathrm{nrd}(x))$ *and* $m \leq n < \infty$, *then* $N_{\ell^m}(x) = N_{\ell^{n-m}}(\bar{x})$.

Proof. For (1), any right $\mathcal{O}_\ell$-ideal I of norm ℓ^m contains $\ell^m \mathcal{O}_\ell$, so $x \in I$ if and only if $x' \in I$. That is, $\mathfrak{N}_{\ell^m}(x) = \mathfrak{N}_{\ell^m}(x')$ and consequently $N_{\ell^m}(x) = N_{\ell^m}(x')$.

For (2), observe that

$$N_{\ell^m}(x) = \left| \left\{ (y,z) \in \mathcal{O}_\ell \times \mathcal{O}_\ell : yz = x,\ \mathrm{nrd}(y) \in \ell^m \mathbb{Z}_\ell^\times \right\} / {\sim} \right|$$

with the equivalence $(y,z) \sim (yu, u^{-1}z)$ for any $u \in \mathcal{O}_\ell^\times$. Indeed, to each (y,z) we can associate the ideal $y\mathcal{O}_\ell \in \mathfrak{N}_{\ell^m}(x, \mathcal{O}_\ell)$ with two pairs determining the same ideal if and only if they are equivalent, and every ideal in $\mathfrak{N}_{\ell^m}(x, \mathcal{O}_\ell)$ can be obtained in this way. From this we see that $(y,z) \mapsto (\bar{z}, \bar{y})$ induces a bijection $\mathfrak{N}_{\ell^m}(x) \to \mathfrak{N}_{\ell^{n-m}}(\bar{x})$. $\square$

Say $x \in \mathcal{O}_\ell$ is m**-norm-reduced** if $\mathrm{nrd}(x) \neq 0$ and

$$\mathrm{val}(\mathrm{nrd}(x)) \leq m + \mathrm{val}(\mathrm{trd}(x)). \tag{7}$$

Note that every x with nonzero norm is m-norm-reduced for sufficiently large m, so this is only a nontrivial constraint when m is fixed. We will show in the proof of Proposition 3.10 that whenever x is m-norm-reduced, we can define an explicit bijection $\mathfrak{N}_{\ell^m}(x) \to \mathfrak{N}_{\ell^m}(\bar{x})$.

In order to prove $N_{\ell^m}(x) = N_{\ell^m}(\bar{x})$ more generally, we first need a way to reduce to the m-norm-reduced case.

Lemma 3.9. *Let $\mathcal{O}_\ell$ be a quaternion order over $\mathbb{Z}_\ell$. Let $x \in \mathcal{O}_\ell$ with $\operatorname{nrd}(x) \neq 0$. Let $m \in \mathbb{N}$ and assume that $\operatorname{val}(\operatorname{trd}(x)) \neq m$. Then there exists an m-norm-reduced $x' \in \mathcal{O}_\ell$ with $x - x' \in \ell^m \mathcal{O}_\ell$.*

Proof. Let $t := \operatorname{val}(\operatorname{trd}(x))$ and $n := \operatorname{val}(\operatorname{nrd}(x))$. If $n \leq m + t$ then $x' = x$ is already m-norm-reduced, so suppose instead that $n > m + t$. Since we are assuming $m \neq t$, the valuation of

$$\operatorname{nrd}(x + \ell^m) = \operatorname{nrd}(x) + \ell^m \operatorname{trd}(x) + \ell^{2m}$$

is equal to $m + \min(m, t)$, because exactly one of the three terms attains this minimal valuation. In particular, this implies $\operatorname{nrd}(x + \ell^m) \neq 0$. On the other hand,

$$\operatorname{trd}(x + \ell^m) = \operatorname{trd}(x) + 2\ell^m$$

has valuation greater than or equal to $\min(m, t)$, so

$$\operatorname{val}(\operatorname{nrd}(x + \ell^m)) = m + \min(m, t) \leq \operatorname{val}(\ell^m \operatorname{trd}(x + \ell^m)).$$

Thus $x + \ell^m$ is m-norm-reduced. $\square$

Proposition 3.10. *For any quaternion order $\mathcal{O}_\ell$ over $\mathbb{Z}_\ell$, any $m \in \mathbb{N}$ and any $x \in \mathcal{O}_\ell$ with $\operatorname{nrd}(x) \neq 0$, we have $N_{\ell^m}(x) = N_{\ell^m}(\bar{x})$.*

Proof. We first prove the result assuming $\operatorname{val}(\operatorname{trd}(x)) \neq m$. By Lemma 3.8 we can replace x with any $x' \equiv x \bmod \ell^m \mathcal{O}_\ell$ without changing the values of $N_{\ell^m}(x)$ and $N_{\ell^m}(\bar{x})$. So by Lemma 3.9, we may assume without loss of generality that x is m-norm-reduced.

Under this assumption we will prove that $y\mathcal{O}_\ell \mapsto \bar{x}x^{-1}y\mathcal{O}_\ell$ is a well-defined map $\mathfrak{N}_{\ell^m}(x) \to \mathfrak{N}_{\ell^m}(\bar{x})$. Given $y\mathcal{O}_\ell \in \mathfrak{N}_{\ell^m}(x)$, we have

$$x \in y\mathcal{O}_\ell \quad \Leftrightarrow \quad \bar{x} \in \bar{x}x^{-1}y\mathcal{O}_\ell$$

and $\operatorname{nrd}(\bar{x}x^{-1}y) = \operatorname{nrd}(y)$, so it suffices to prove $\bar{x}x^{-1}y \in \mathcal{O}_\ell$. Observe that $\operatorname{nrd}(y)\operatorname{trd}(x)$ has valuation $m + \operatorname{val}(\operatorname{trd}(x))$, and so its quotient by $\operatorname{nrd}(x)$ is in $\mathbb{Z}_\ell$ by Eq. (7). Since $y^{-1}x \in \mathcal{O}_\ell$, we have

$$\operatorname{trd}(x)(x^{-1}y) = \frac{\operatorname{trd}(x)}{\operatorname{nrd}(y^{-1}x)} \overline{y^{-1}x} \in \mathcal{O}_\ell,$$

and therefore

$$\bar{x}x^{-1}y = \operatorname{trd}(x)(x^{-1}y) - y \in \mathcal{O}_\ell.$$

This proves that $\bar{x}x^{-1}y\mathcal{O}_\ell \in \mathfrak{N}_{\ell^m}(\bar{x})$ as desired. Since $\bar{x}$ is also m-norm-reduced, the same construction defines an inverse map $\mathfrak{N}_{\ell^m}(\bar{x}) \to \mathfrak{N}_{\ell^m}(x)$, proving that the two sets have the same cardinality.

Finally we consider the case $\operatorname{val}(\operatorname{trd}(x)) = m$. Set $n := \operatorname{val}(\operatorname{nrd}(x))$. If $n \leq 2m$ then x is m-norm-reduced, and we may continue the argument as above. Otherwise $n > 2m$ and so we have $\operatorname{val}(\operatorname{trd}(x)) \neq n - m$. By Lemma 3.9, there is an $(n-m)$-norm-reduced element $x' \in \mathcal{O}_\ell$ with $x - x' \in \ell^{n-m}\mathcal{O}_\ell$, and so we may use the above argument to prove that $N_{\ell^{n-m}}(x) = N_{\ell^{n-m}}(\bar{x})$. This implies $N_{\ell^m}(x) = N_{\ell^m}(\bar{x})$ by Lemma 3.8(2). $\square$

We are now ready to prove the commutativity of the Brandt algebra $\mathfrak{B}(\mathcal{O})$.

Proof of Theorem 3.1. By Lemma 3.7, it suffices to prove that $B_{\ell^k} B_{\ell^m} = B_{\ell^m} B_{\ell^k}$, for all primes ℓ and all $m, k \geq 0$. Setting $n := m + k$, this is equivalent to prove that $B_{\ell^{n-m}} B_{\ell^m} = B_{\ell^m} B_{\ell^{n-m}}$, for all primes ℓ and for all $0 \leq m \leq n$.

Fix $1 \leq i, j \leq h$, and let $J \in \mathfrak{b}_{\ell^n}(i, j)$. Write $(I_j)_\ell = \gamma \mathcal{O}_\ell$ for some $\gamma \in \mathcal{A}_\ell^\times$, so that $\gamma^{-1} J_\ell = x\mathcal{O}_\ell$ for some $x \in \mathcal{O}_\ell$ with $\operatorname{val}(\operatorname{nrd}(x)) = n$. We can define a map

$$\mathfrak{N}_{\ell^m}(J, I_j) \to \mathfrak{N}_{\ell^m}(x, \mathcal{O}_\ell)$$

by $J' \mapsto \gamma^{-1}J'_\ell$. The map is well-defined because the set on the left varies over locally principal right $\mathcal{O}$-ideals, so $\gamma^{-1}J'_\ell$ is a principal right $\mathcal{O}_\ell$-ideal of the correct norm, with $J \subset J'$ implying $x \in \gamma^{-1}J'_\ell$. The constraint $\frac{\operatorname{nrd}(J')}{\operatorname{nrd}(I_j)} = \ell^m$ forces $J'_{\ell'} = (I_j)_{\ell'}$ for all primes $\ell' \neq \ell$, so by the local-global dictionary, this map is in fact a bijection. We therefore have

$$|\mathfrak{N}_{\ell^m}(J, I_j)| = N_{\ell^m}(x). \tag{8}$$

Replacing m with $n - m$, the same argument together with Lemma 3.8(2) and Proposition 3.10 shows that

$$|\mathfrak{N}_{\ell^{n-m}}(J, I_j)| = N_{\ell^{n-m}}(x) = N_{\ell^m}(\bar{x}) = N_{\ell^m}(x).$$

This proves that $|\mathfrak{N}_{\ell^m}(J, I_j)| = |\mathfrak{N}_{\ell^{n-m}}(J, I_j)|$ for all $J \in \mathfrak{b}_{\ell^n}(i, j)$, so that by Proposition 3.5 we have $(B_{\ell^m}B_{\ell^{n-m}})(i, j) = (B_{\ell^{n-m}}B_{\ell^m})(i, j)$. □

4. Submaximal orders

Let $\mathcal{A}$ be a quaternion algebra over $\mathbb{Q}$, and let $\mathfrak{O} \subset \mathcal{A}$ be a maximal order. Our strategy to prove Theorem 7.2 is to show that $\mathfrak{O}$ has more elements of a given norm than any of its proper suborders. It will be useful to consider the maximal elements among the set of all proper suborders, and so we introduce the following convenient terminology.

Definition 4.1. An order $\mathcal{O} \subsetneq \mathfrak{O}$ is a *submaximal order of* $\mathfrak{O}$ (or, simply, $\mathcal{O}$ is *submaximal in* $\mathfrak{O}$) if it is maximal with respect to inclusion among the proper suborders of $\mathfrak{O}$.

4.1. Classification.

Lemma 4.2. *Let $\mathfrak{O}$ be a maximal order in $\mathcal{A}$, and let $\mathcal{O} \subsetneq \mathfrak{O}$ be a submaximal order. Then one of the following holds:*

(*a*) $[\mathfrak{O} : \mathcal{O}] = r$ *for a split prime r of $\mathcal{A}$, and $\mathcal{O}$ is Eichler.*
(*b*) $[\mathfrak{O} : \mathcal{O}] = p$ *for a ramified prime p of $\mathcal{A}$, and $\mathcal{O} = \mathbb{Z} + \mathfrak{P}$, with $\mathfrak{P}$ the unique maximal two-sided ideal of norm p in $\mathfrak{O}$.*
(*c*) $[\mathfrak{O} : \mathcal{O}] = r^2$ *for a split prime r of $\mathcal{A}$, and $\mathcal{O} = R + r\mathfrak{O}$ for some imaginary quadratic order $R \subseteq \mathfrak{O}$ with $R/rR \simeq \mathbb{F}_{r^2}$.*
(*d*) $[\mathfrak{O} : \mathcal{O}] = p^2$ *for a ramified prime p of $\mathcal{A}$, and $\mathcal{O} = R + p\mathfrak{O}$ for some imaginary quadratic order $R \subseteq \mathfrak{O}$ with $R/pR \simeq \mathbb{F}_{p^2}$.*

Remark 4.3. Cases (*c*) and (*d*) can be combined into a single case by ignoring the splitting behavior of the prime in $\mathcal{A}$, but the two cases behave differently enough in practice that it will be useful to separate them. The condition $R/\ell R \simeq \mathbb{F}_{\ell^2}$ is equivalent to the condition that ℓ is inert in the integral closure of R and does not divide the conductor of R; for odd ℓ, this is equivalent to the condition that the discriminant of R is a quadratic non-residue modulo ℓ.

Proof. By the local-global principle for lattices (e.g., see [Voi21, Theorem 9.1.1] together with [Voi21, Lemma 10.4.3]), there exists a prime ℓ such that $\mathcal{O}$ is maximal at every prime except ℓ, so ℓ is the unique prime dividing $[\mathfrak{O} : \mathcal{O}]$. Let $\pi : \mathfrak{O} \to \mathfrak{O}/\ell\mathfrak{O}$ be reduction mod ℓ. By Lemma 2.2, $\mathcal{O}$ cannot surject onto $\mathfrak{O}/\ell\mathfrak{O}$ under π, so $\mathcal{O}$ lands in some proper subring $A \subsetneq \mathfrak{O}/\ell\mathfrak{O}$. For any such A, $\pi^{-1}(A)$ will be a suborder of $\mathfrak{O}$ containing $\mathcal{O}$, hence equal to it. Thus it suffices to classify maximal proper subrings A of the 4-dimensional $\mathbb{F}_\ell$-algebra $Q := \mathfrak{O}/\ell\mathfrak{O}$. Note that $A \neq \mathbb{F}_\ell$, because we have proper subrings of the form $\mathbb{F}_\ell[\alpha]$ for any $\alpha \in Q \smallsetminus \mathbb{F}_\ell$.

First suppose that A contains an element α with irreducible minimal polynomial over $\mathbb{F}_\ell$, i.e. it contains a copy of $\mathbb{F}_{\ell^2}$. Then A is a vector space over $\mathbb{F}_{\ell^2}$; since $A \neq Q$, A must equal $\mathbb{F}_\ell[\alpha]$ by dimension considerations. A preimage of α in $\mathfrak{O}$ generates a quadratic order R in which ℓ is inert, and $\pi^{-1}(A) = R + \ell\mathfrak{O}$. This gives cases (*c*) and (*d*).

Now we assume every element of $A \smallsetminus \mathbb{F}_\ell$ has reducible minimal polynomial, and divide into cases depending on whether ℓ is split or ramified in $\mathcal{A}$.

- If $r := \ell$ is a split prime of $\mathcal{A}$, then $Q \simeq M_2(\mathbb{F}_r)$, and we can consider the natural action of Q on $\mathbb{F}_r^2$. Our working assumption implies that A does not contain any nontrivial field extension of $\mathbb{F}_r$, so by [Rac74, Theorem 1], A is the stabilizer of a 1-dimensional subspace of $\mathbb{F}_r^2$. Picking a basis of $\mathbb{F}_r^2$ so that the first basis vector is in this stabilized subspace, we see that A is isomorphic to the set of upper triangular matrices in $M_2(\mathbb{F}_r)$. Therefore $\pi^{-1}(A) \otimes \mathbb{Z}_r \simeq \left(\begin{smallmatrix} \mathbb{Z}_r & \mathbb{Z}_r \\ r\mathbb{Z}_r & \mathbb{Z}_r \end{smallmatrix}\right)$, and so $\mathcal{O}$ is Eichler with index r in $\mathfrak{O}$ ([Voi21, Proposition 23.4.3]).
- If $p := \ell$ is a ramified prime of $\mathcal{A}$, then $Q \simeq \mathbb{F}_{p^2}[\epsilon]$, with noncommutative multiplication
$$(a + b\epsilon)(c + d\epsilon) = ac + (b\bar{c} + ad)\epsilon.$$
Here $\epsilon^2 = 0$ and $\bar{c} = c^p$ is the Frobenius involution on $\mathbb{F}_{p^2}$. Note that $C := \mathbb{F}_p + \mathbb{F}_{p^2}\epsilon$ is a 3-dimensional subalgebra of Q. Any $\alpha \in Q \smallsetminus C$ is of the form $a + b\epsilon$ with $a \in \mathbb{F}_{p^2} \smallsetminus \mathbb{F}_p$, and hence it satisfies an irreducible quadratic minimal polynomial over $\mathbb{F}_p$ (as one can check that $a + b\epsilon$ and a have the same minimal polynomial in this case). Since A contains no such elements, we have $A \subseteq C$ and hence $A = C$ by maximality. This corresponds under π^{-1} to the order $\mathbb{Z} + \mathfrak{P}$, where $\mathfrak{P} = \pi^{-1}(\mathbb{F}_{p^2}\epsilon)$ is the unique maximal two-sided ideal of $\mathfrak{O}$ with norm p. □

For the rest of this paper, following the notation in Lemma 4.2, we use p to refer to a prime at which $\mathcal{A}$ ramifies, and for a submaximal order $\mathcal{O} \subseteq \mathfrak{O}$ we write r for a prime dividing $[\mathfrak{O} : \mathcal{O}]$ at which $\mathcal{A}$ splits. We continue to use ℓ for arbitrary primes, unless otherwise specified.

Remark 4.4. We give a few indications of how submaximal orders fit within existing classifications of quaternion orders. Let $\mathcal{O} \subseteq \mathfrak{O}$ be submaximal. Then $\mathcal{O}$ is Bass ([Voi21, Definition 24.5.1]), which follows from the proof of Lemma 4.5 below. Since $\mathcal{A}$ is defined over $\mathbb{Q}$, this implies by [CSV21] that $\mathcal{O}$ is basic ([Voi21, Definition 24.5.6]). The order $\mathcal{O}$ is Eichler if and only if $[\mathfrak{O} : \mathcal{O}] = r$ is a split prime in $\mathcal{A}$ (case (a)), in which case it is also hereditary ([Voi21, Definition 21.1.1]).

If $[\mathfrak{O} : \mathcal{O}]$ is equal to r, p, or p^2 (cases (a), (b), and (d)) then $\mathcal{O}$ is a "special order" in the terminology of Hijikata, Pizer, and Shemanske [HPS89, Definition 6.1]. But since special orders are Eichler away from the ramified primes of $\mathcal{A}$ by definition, submaximal orders with $[\mathfrak{O} : \mathcal{O}] = r^2$ (case (c)) are not special.

One useful property shared by all submaximal orders is that their level as a lattice is equal to their reduced discriminant.

Lemma 4.5. *For any submaximal order $\mathcal{O}$ in a maximal order $\mathfrak{O}$, the level of $\mathcal{O}$ as a lattice (with respect to the trace form on $\mathcal{A}$) is equal to* $\operatorname{disc}(\mathcal{O})$.

Proof. To prove this we will first show that $\mathcal{O}$ is Gorenstein. An order $\mathcal{O}$ in a quaternion algebra is called **Gorenstein** if its dual lattice $\mathcal{O}^\sharp$ is in fact an invertible right $\mathcal{O}$-ideal. This is a local property. One can show that being Gorenstein is equivalent to $\operatorname{Hom}_{\mathbb{Z}}(\mathcal{O}, \mathbb{Z})$ being a projective $\mathcal{O}$-bimodule, which connects the definition to a more general theory of Gorenstein modules. Eichler orders, and hence maximal orders, are Gorenstein, but a general order is not. For all this, and much more, see [Voi21, §24.2].

In order to show a submaximal order $\mathcal{O}$ is Gorenstein, let $\mathcal{O}'$ denote the Gorenstein saturation of $\mathcal{O}$, so that $\mathcal{O} = \mathbb{Z} + f\mathcal{O}'$ for some $f \geq 1$ [Voi21, Proposition 24.1.4]. Then $[\mathcal{O}' : \mathcal{O}] = f^3$. But the index of a submaximal order $\mathcal{O}$ in any superorder is either 1, a prime, or the square of a prime, as one can see by calculating discriminants. Therefore, we must have $f = 1$ and hence $\mathcal{O}' = \mathcal{O}$.

Since $\mathcal{O}$ is Gorenstein, we have $\operatorname{nrd}(\mathcal{O}^\sharp)^{-1} = \operatorname{disc}(\mathcal{O})$ [Voi21, Lemma 24.2.5], and so the result follows by Lemma 2.5. □

4.2. **Genus vectors in the Brandt module.** Let I, J be invertible right $\mathcal{O}$-ideals. Recall that we can equip I and J with their normalized trace forms $\langle\cdot,\cdot\rangle_I$ and $\langle\cdot,\cdot\rangle_J$, (Definition 2.3) and write $J \in \mathrm{Gen}(I)$ if the completions of these lattices are isometric at every prime.

Let $I_1, \ldots, I_h$ be representatives of the invertible right $\mathcal{O}$-ideal classes, and $v_1, \ldots, v_h$ the corresponding Brandt module standard basis vectors. For each i, we define the **genus vector** $\gamma_i \in \mathfrak{M}(\mathcal{O})$ by

$$\gamma_i := \sum_{\substack{1\le j\le h \\ I_j \in \mathrm{Gen}(I_i)}} \frac{1}{2w_j} v_j,$$

and let $\mathfrak{M}_{\mathrm{G}}(\mathcal{O}) \subseteq \mathfrak{M}(\mathcal{O})$ denote the subspace spanned by the genus vectors γ_i for $1 \le i \le h$. By construction, $\gamma_i = \gamma_j$ if $I_j \in \mathrm{Gen}(I_i)$, so the number of distinct genus vectors is typically smaller than h; in fact we will see below that for many orders (including all maximal orders) there is a unique genus vector. Let

$$e := \sum_{1\le j\le h} \frac{1}{2w_j} v_j \tag{9}$$

be the sum of the (distinct) genus vectors.

Lemma 4.6. *Let $\mathcal{O} \subseteq \mathcal{A}$ be a quaternion order. The vector e defined by Eq. (9) is an eigenvector of B_n for all $n \ge 0$.*

Proof. The image of B_0 is the span of e, so e is an eigenvector of B_0. For any $n \ge 1$ and $1 \le i \le h$, the coefficient of v_i in $B_n e$ is

$$\sum_{1\le j\le h} \frac{B_n(i,j)}{2w_j} = \sum_{1\le j\le h} \frac{B_n(j,i)}{2w_i}$$

by self-adjointness of B_n (see Lemma 2.9). By Lemma 3.3 this equals $\frac{1}{2w_i}$ times a constant that does not depend on i, so $B_n e$ is a multiple of e. □

We now focus on maximal and submaximal orders, and find that the genus subspace is either 1- or 2-dimensional.

4.2.1. 1*-dimensional genus subspace.*

Lemma 4.7. *Let $\mathcal{O} \subseteq \mathcal{A}$ be a maximal or submaximal order. Let $\mathfrak{O}$ be a maximal order containing $\mathcal{O}$, and suppose that $[\mathfrak{O} : \mathcal{O}]$ is* not *equal to an odd ramified prime of $\mathcal{A}$. Then $\mathfrak{M}_{\mathrm{G}}(\mathcal{O})$ is* 1*-dimensional, spanned by e.*

Proof. If $\mathcal{O}$ is submaximal then we use the classification result in Lemma 4.2: we are in cases (a), (c) and (d) of the lemma, and also in case (b) if $p = 2$.

We first show that for every prime ℓ, every element of $\mathbb{Z}_\ell^\times$ is in the image of $\mathrm{nrd}\colon \mathcal{O}_\ell \to \mathbb{Z}_\ell$. If $\mathcal{O}_\ell$ is maximal, or if $[\mathfrak{O} : \mathcal{O}] = \ell^2$, then $\mathcal{O}_\ell$ contains an embedding of $\mathbb{Z}_{\ell^2}$ (the ring of integers of the unramified quadratic extension of $\mathbb{Q}_\ell$), and already $\mathrm{nrd}\colon \mathbb{Z}_{\ell^2}^\times \to \mathbb{Z}_\ell^\times$ is surjective. If instead $\ell = r = [\mathfrak{O} : \mathcal{O}]$ is split in $\mathcal{A}$, then $\mathcal{O}_\ell$ contains an embedding of $\mathbb{Z}_\ell \times \mathbb{Z}_\ell$, and $\mathrm{nrd} : \mathbb{Z}_\ell^\times \times \mathbb{Z}_\ell^\times \to \mathbb{Z}_\ell^\times$ $((x,y) \mapsto xy)$ is surjective. Finally, if $\ell = p = [\mathfrak{O} : \mathcal{O}] = 2$ is ramified in $\mathcal{A}$, then $\mathfrak{O}_2$ contains an embedding of $\mathbb{Z}_4$, so $\mathcal{O}_2$ contains an embedding of $2\mathbb{Z}_4$, and $\mathrm{nrd} : (1 + 2\mathbb{Z}_4) \to (1 + 2\mathbb{Z}_2) = \mathbb{Z}_2^\times$ is surjective [CF10, §I.7 Proposition 3].

Now, we prove that any two invertible right $\mathcal{O}$-ideals I, I' with normalized trace forms are in the same genus. Fix a prime ℓ. Since I, I' are locally principal, there exists $\alpha \in \mathcal{A}_\ell^\times$ with $I'_\ell = \alpha I_\ell$. This implies $\mathrm{nrd}(I')$ and $\mathrm{nrd}(\alpha)\mathrm{nrd}(I)$ have the same ℓ-adic valuation. Since we have proven that every element of $\mathbb{Z}_\ell^\times$ is in the image of the norm map, there exists $u \in \mathcal{O}_\ell$ with $\mathrm{nrd}(u) = \frac{\mathrm{nrd}(I')}{\mathrm{nrd}(\alpha)\mathrm{nrd}(I)}$. Then for any $x, y \in I_\ell$ we have

$$\langle \alpha x u, \alpha y u\rangle_{I'} = \frac{\mathrm{trd}((\alpha x u)(\bar{u}\bar{y}\bar{\alpha}))}{\mathrm{nrd}(I')} = \frac{\mathrm{nrd}(u)\mathrm{nrd}(\alpha)\mathrm{trd}(x\bar{y})}{\mathrm{nrd}(I')} = \frac{\mathrm{trd}(x\bar{y})}{\mathrm{nrd}(I)} = \langle x, y\rangle_I.$$

So $x \mapsto \alpha x u$ is an isometry from I_ℓ to I'_ℓ. $\square$

4.2.2. 2-*dimensional genus subspace.* Now suppose instead that $p = [\mathfrak{O} : \mathcal{O}]$ is an odd ramified prime for $\mathcal{A}$. This case is covered in detail in [Piz80b, Section 5]. Pizer's setup assumes that the only ramified prime of $\mathcal{A}$ is p, but the summary below holds more generally for arbitrary quaternion algebras over $\mathbb{Q}$. We also note that Pizer works with left $\mathcal{O}$-ideals while we work with right $\mathcal{O}$-ideals.

For each invertible right $\mathcal{O}$-ideal I, the normalized quadratic form $\frac{1}{2}\langle x, x\rangle_I = \frac{\mathrm{nrd}(x)}{\mathrm{nrd}(I)}$ on I either always outputs quadratic residues mod p (and I has **positive character**) or always outputs quadratic non-residues mod p (and I has **negative character**), where we ignore the outputs in $p\mathbb{Z}_p$ [Piz80b, Proposition 5.1]. Ideals have the same character if and only if they are in the same genus: this can be seen by a modification of the proof of Lemma 4.7 (see the proof of [Piz80b, Lemma 5.20] and [Piz80b, Remark 5.21] for more details). In particular this implies that the character is the same for all ideals in the same ideal class.

Now recall the representatives $I_1 = \mathcal{O}, \dots, I_h$ of the right ideal classes of $\mathcal{O}$. By [Piz80b, Theorem 5.9], the class number h of $\mathcal{O}$ is even, and we may label these representatives so that $I_1, \dots, I_{h/2}$ have positive character, $I_{h/2+1}, \dots, I_h$ have negative character, and for each $1 \le i \le h/2$, I_i and $I_{i+h/2}$ have isomorphic left orders (so that $w_i = w_{i+h/2}$). Then we set

$$e' = \sum_{i=1}^{h/2} \frac{1}{2w_i} v_i - \sum_{i=h/2+1}^{h} \frac{1}{2w_i} v_i. \tag{10}$$

That is, e is the sum of the two genus vectors, and e' is the difference of the two genus vectors. The above observations allow us to conclude the following result.

Lemma 4.8. *Let $\mathcal{O}$ be submaximal in a maximal order $\mathfrak{O}$. Assume that* $[\mathfrak{O} : \mathcal{O}]$ is *equal to an odd ramified prime of $\mathcal{A}$. Then $\mathfrak{M}_{\mathrm{G}}(\mathcal{O})$ is 2-dimensional, spanned by e and e'.*

Since $w_i = w_{i+h/2}$ for all $1 \le i \le h/2$, e' is orthogonal to e. In the next lemma we show that e', like e, is also an eigenvector for any Brandt matrix B_n, $n \ge 0$. The computation of the eigenvalues of e and e' is the primary focus of Section 6.2.

Lemma 4.9. *Let $\mathcal{O}$ be submaximal in a maximal order $\mathfrak{O}$, and assume that $[\mathfrak{O} : \mathcal{O}]$ is equal to an odd ramified prime of $\mathcal{A}$. Then e' as given by* Eq. (10) *is an eigenvector of B_n for all $n \ge 0$.*

Proof. The image of B_0 is the span of e, but since B_0 is self-adjoint and $\langle e, e'\rangle = 0$ we have $\langle e, B_0 e'\rangle = 0$. Therefore we must have $B_0 e' = 0$.

We now assume $n \ge 1$. Following the proof of [Piz80b, Theorem 5.15] for $s = 0$, with respect to the standard basis $\{v_1, \dots, v_h\}$, we can write

$$B_n = \begin{pmatrix} C_n & D_n \\ D_n & C_n \end{pmatrix}$$

for some matrices C_n, D_n of size $h/2 \times h/2$. We first show that C_n is self-adjoint and has constant column sums (cf. [Piz80b, Lemma 5.23], parts (i) and (ii)). For self-adjointness, letting D' be the diagonal matrix with $w_1, \dots, w_{h/2}$ along the main diagonal and $D = \left(\begin{smallmatrix} D' & 0 \\ 0 & D' \end{smallmatrix}\right)$, we have $DB_n = B_n^t D$ by Lemma 2.9 and hence $D'C_n = C_n^t D'$. To show $\sum_{i=1}^{h/2} B_n(i, j)$ is a constant for all $1 \le j \le h/2$, note that this quantity is simply the number of invertible right $\mathcal{O}$-ideals $J \subseteq_n I_j$ with positive character, and the conditions defining such J's are local, so the proof follows as in Lemma 3.3.

With respect to the basis $\{v_1, \dots, v_h\}$, let $u = \left(\frac{1}{2w_1}, \dots, \frac{1}{2w_{h/2}}\right)^t$, so that $e = \left(\begin{smallmatrix} u \\ u \end{smallmatrix}\right)$, $e' = \left(\begin{smallmatrix} u \\ -u \end{smallmatrix}\right)$. Using the properties of C_n shown above, the proof of Lemma 4.6 shows that u

is an eigenvector for C_n. Now since e is an eigenvector for B_n, we find that $(C_n + D_n)u$ is a scalar multiple of u, so u is also an eigenvector of D_n. This implies $B_n e' = \begin{pmatrix} (C_n - D_n)u \\ (D_n - C_n)u \end{pmatrix}$ is a scalar multiple of e' as desired. □

4.3. **Brandt algebra relations.** Let $\mathcal{O}$ be a maximal or a submaximal order in the quaternion algebra $\mathcal{A}$. For the proof of Theorem 7.2 we will need to consider the action of the Brandt matrices B_{ℓ^k} on $\mathfrak{M}(\mathcal{O})$ for various primes ℓ and arbitrary exponents $k \geq 1$. We divide into cases depending on the power of the prime that divides the index of $\mathcal{O}$ in a maximal order $\mathfrak{O}$.

For our purposes we will not need to study ramified primes of $\mathcal{A}$ that divide the index, so by Lemma 4.2 we only need to consider three cases: $\ell \nmid [\mathfrak{O} : \mathcal{O}]$ (Lemma 4.10 and Lemma 4.11), ℓ equals a split prime r with $[\mathfrak{O} : \mathcal{O}] = r$ (Proposition 4.13), or ℓ equals a split prime r with $[\mathfrak{O} : \mathcal{O}] = r^2$ (Proposition 4.22). One of the main results we provide for each of these orders is a recurrence relation satisfied by the Brandt matrices B_{ℓ^k} for $k \geq 0$: these are Lemma 4.10, Proposition 4.13, and Corollary 4.20, respectively. As this section is rather long and technical, we recommend readers skip ahead to Section 5 on a first read-through and only refer back to this section as necessary.

4.3.1. $\ell \nmid [\mathfrak{O} : \mathcal{O}]$. In this case $\mathcal{O}_\ell$ is maximal, and we have the following two classical results, depending on whether ℓ is split or ramified in $\mathcal{A}$. (We note that $\ell \mid \operatorname{disc}(\mathcal{O}) \iff \ell \mid \operatorname{disc}(\mathfrak{O})$.)

Lemma 4.10. *Let $\mathcal{O} \subseteq \mathcal{A}$ be a quaternion order and $\ell \nmid \operatorname{disc}(\mathcal{O})$ a prime. Then for all $k \geq 1$ we have*

$$B_{\ell^{k+1}} = B_\ell B_{\ell^k} - \ell B_{\ell^{k-1}}.$$

Proof. See [Eic73, Chapter II Theorem 2] or [Voi21, Proposition 41.3.6]. □

Lemma 4.11. *Let $\mathcal{O} \subseteq \mathcal{A}$ be a quaternion order with $\mathcal{O}_p$ maximal, where p is a ramified prime for $\mathcal{A}$. Then for all $k \geq 1$ we have $B_{p^k} = B_p^k$, and $B_p^2 = 1$.*

Proof. In this setting $\mathcal{O}_p$ is the valuation ring of the division algebra $\mathcal{A}_p$, so for each $k \geq 0$ there is a unique ideal of norm p^k in $\mathcal{O}_p$, which is contained in all the ideals of norm $p^{k'}$ for $k' \leq k$. Using the local-global dictionary, this implies that if I is an invertible right $\mathcal{O}$-ideal, then there is a unique ideal of relative norm p^k in I for every $k \geq 1$.

This implies that for all $J \in \mathfrak{b}_{p^k}(i,j)$ we have $|\mathfrak{N}_p(J, I_j)| = 1$. By Proposition 3.5 this implies $B_{p^{k-1}} B_p = B_{p^k}$, which gives $B_p^k = B_{p^k}$ by induction. We compute $B_{p^2} = 1$ by noting that the only $\mathcal{O}_p$-ideal of norm p^2 is $p\mathcal{O}_p$, so for any $1 \leq j \leq h$ the unique ideal of relative norm p^2 in I_j is pI_j, which is in the same ideal class as I_j. Thus $|\mathfrak{b}_{p^2}(i,j)|$ is 1 if $i = j$ and is 0 otherwise. □

4.3.2. $[\mathfrak{O} : \mathcal{O}] = r$. Let $\mathfrak{O}$ be a maximal order containing $\mathcal{O}$ and suppose $[\mathfrak{O} : \mathcal{O}] = r$ for a split prime r of $\mathcal{A}$. The main result in this section is Proposition 4.13, a recurrence relation on the Brandt matrices B_{r^k}.

By the classification of submaximal orders Lemma 4.2, $\mathcal{O}$ is an Eichler order of index r in $\mathfrak{O}$. Let $\mathfrak{r}$ be the unique invertible two-sided ideal of $\mathcal{O}$ with norm r [Voi21, 23.4.19]. We can define an operator

$$W_\mathfrak{r} : \mathfrak{M}(\mathcal{O}) \to \mathfrak{M}(\mathcal{O}) \tag{11}$$

on standard basis vectors induced by $[I_i] \mapsto [I_i \mathfrak{r}]$.

Lemma 4.12. *$W_\mathfrak{r}$ is an involution that is self-adjoint with respect to the inner product on the Brandt module $\mathfrak{M}(\mathcal{O})$ and commutes with every Brandt matrix $B_n \in \mathfrak{B}(\mathcal{O})$.*

Proof. Since $\mathfrak{r}^2 = r\mathcal{O}$, $W_\mathfrak{r}$ is an involution. Thus, for $1 \le i, j \le h$, we have $W_\mathfrak{r} v_i = v_j$ if and only if $v_i = W_\mathfrak{r} v_j$, and so if these equalities do not hold we have

$$\langle v_i, W_\mathfrak{r} v_j\rangle = 0 = \langle W_\mathfrak{r} v_i, v_j\rangle.$$

If the equalities do hold, then $[I_j] = [I_i\mathfrak{r}]$, which implies that I_i and I_j have isomorphic left orders and so

$$\langle v_i, W_\mathfrak{r} v_j\rangle = \langle v_i, v_i\rangle = w_i = w_j = \langle v_j, v_j\rangle = \langle W_\mathfrak{r} v_i, v_j\rangle.$$

This proves self-adjointness.

Letting σ be the permutation of $\{1, \ldots, h\}$ induced by $[I_{\sigma(i)}] := [I_i\mathfrak{r}]$, the following holds for any $n \ge 1$:

$$\begin{aligned}(B_n W_\mathfrak{r})(i,j) &= B_n(i, \sigma(j))\\ &= \{J \subseteq_n I_j\mathfrak{r} : [J] = [I_i]\}\\ &= \{J \subseteq_n I_j : [J] = [I_i\mathfrak{r}]\}\\ &= B_n(\sigma(i), j) = (W_\mathfrak{r} B_n)(i,j),\end{aligned}$$

where in the third line we used the bijection $J \mapsto J\mathfrak{r}^{-1}$.

Finally, since $W_\mathfrak{r}$ is self-adjoint and has constant column sum (equal to 1), the argument from Lemma 3.4 shows that it commutes with B_0. $\square$

As a result, we may pick a basis for $\mathfrak{M}(\mathcal{O})$ consisting of simultaneous eigenvectors for all Brandt operators as well as for $W_\mathfrak{r}$. The main result we will prove in this section is a recurrence relation for the Brandt operators $B_{r^n} \in \mathfrak{B}(\mathcal{O})$.

Proposition 4.13. *Let $\mathcal{O} \subseteq \mathfrak{O}$ be submaximal of index r, where r is a split prime of $\mathcal{A}$, and let $W_\mathfrak{r}$ be as in Lemma 4.12. Then for all $n \ge 1$,*

$$B_{r^{n+1}} = B_{r^n}(B_r - rW_\mathfrak{r}) - rB_{r^{n-1}}.$$

Proof. In light of Proposition 3.5, we may express $(B_{r^n}B_r)(i,j)$ as a sum over $J \in \mathfrak{b}_{r^{n+1}}(i,j)$ of the values $|\mathfrak{N}_r(J, I_j)|$, and by Eq. (8), we can write $|\mathfrak{N}_r(J, I_j)| = N_r(x)$, where $(I_j)_r = \gamma\mathcal{O}_r$ and $\gamma^{-1}J_r = x\mathcal{O}_r$; since $J \subseteq_{r^{n+1}} I_j$ we have $x \in \mathcal{O}_r$ and $\operatorname{val}(\operatorname{nrd}(x)) = n+1$. So we will partition $\mathfrak{b}_{r^{n+1}}(i,j)$ into subsets depending on the value of $N_r(x)$, and use this to find an alternate expression for $(B_{r^n}B_r)(i,j)$.

Lemma 4.14. *Let $x \in \mathcal{O}_r$ and suppose* $\operatorname{val}(\operatorname{nrd}(x)) \ge 2$. *Then*

$$N_r(x) = \begin{cases} 2r+1 & \text{for } x \in r\mathcal{O}_r,\\ r+1 & \text{for } x \in \mathfrak{r}_r \smallsetminus r\mathcal{O}_r,\\ 1 & \text{for } x \in \mathcal{O}_r \smallsetminus \mathfrak{r}_r.\end{cases}$$

Proof. Since $\mathcal{O}$ is Eichler of index r, it is isomorphic to the subring of $M_2(\mathbb{Z}_r)$ consisting of matrices with bottom-left entry divisible by r [Voi21, Proposition 23.4.3]. For ease of notation we assume equality. In this notation we have

$$\mathcal{O}_r = \begin{pmatrix}\mathbb{Z}_r & \mathbb{Z}_r\\ r\mathbb{Z}_r & \mathbb{Z}_r\end{pmatrix} \qquad \text{and} \qquad \mathfrak{r}_r = \begin{pmatrix} r\mathbb{Z}_r & \mathbb{Z}_r\\ r\mathbb{Z}_r & r\mathbb{Z}_r\end{pmatrix}.$$

We define the set S to consist of the following $2r+1$ elements of $\mathcal{O}_r$:

$$\alpha_i := \begin{pmatrix} r & i\\ 0 & 1\end{pmatrix}, \quad \beta_i := \begin{pmatrix}1 & 0\\ ir & r\end{pmatrix} \quad \text{for } i = 0, \ldots, r-1, \qquad \varpi := \begin{pmatrix}0 & 1\\ r & 0\end{pmatrix}.$$

These elements generate each integral invertible right $\mathcal{O}_r$-ideal of norm r exactly once, with ϖ generating the two-sided ideal $\mathfrak{r}_r$.

Now let $x = \left(\begin{smallmatrix} a & b \\ rc & d \end{smallmatrix}\right) \in \mathcal{O}_r$. To determine the number of integral invertible right $\mathcal{O}_r$-ideals J of norm r containing x, it suffices to determine, for each $\mu \in S$, whether or not $\mu^{-1}x \in \mathcal{O}_r$. To this end, we compute $\mu^{-1}x$ for each $\mu \in S$:

$$\alpha_i^{-1}x = \begin{pmatrix} \frac{a}{r} - ci & \frac{b-di}{r} \\ cr & d \end{pmatrix}, \qquad \beta_i^{-1}x = \begin{pmatrix} a & b \\ c - ai & \frac{d}{r} - bi \end{pmatrix}, \qquad \varpi^{-1}x = \begin{pmatrix} c & \frac{d}{r} \\ a & b \end{pmatrix}.$$

If a and d are both units, then $x \in \mathcal{O}_r^\times$, which does not occur because of the assumption $\operatorname{val}(\operatorname{nrd}(x)) \geq 2$. If d is a unit (so $r \mid a$), then there is exactly one $\mu \in S$ for which $\mu^{-1}x \in \mathcal{O}_r$, namely $\mu = \alpha_i$ with $i \equiv bd^{-1} \bmod r$. Likewise, if a is a unit (so $r \mid d$), then we can only take $\mu = \beta_i$ with $i \equiv ca^{-1} \bmod r$.

Now consider the case that neither a nor d is a unit; this is exactly the case that $x \in \mathfrak{r}_r$. In this setting, if b and c are both units, then $x \in \varpi\mathcal{O}_r^\times$, which does not occur by $\operatorname{val}(\operatorname{nrd}(x)) \geq 2$. If b is a unit (so $r \mid c$), then there are $r+1$ options: we can take $\mu = \varpi$ or $\mu = \beta_i$ for any i. If c is a unit (so $r \mid b$), then again there are $r+1$ options: we can take $\mu = \varpi$ or $\mu = \alpha_i$ for any i. If a, b, c, d are all divisible by r (so $x \in r\mathcal{O}_r$), then there are $2r+1$ options, because $\mu^{-1}x \in \mathcal{O}_r$ for all $\mu \in S$. □

Now that we have these counts for each case, we need to count how many $J \in \mathfrak{b}_{r^{n+1}}(i,j)$ land in each of these cases. We have

$$x \in r\mathcal{O}_r \quad \Leftrightarrow \quad J \subset rI_j \quad \Leftrightarrow \quad \frac{1}{r}J \subset I_j \quad \Leftrightarrow \quad \frac{1}{r}J \in \mathfrak{b}_{r^{n-1}}(i,j),$$

so the number of J with $N_r(x) = 2r+1$ is equal to $B_{r^{n-1}}(i,j)$. Likewise,

$$x \in \mathfrak{r}_r \quad \Leftrightarrow \quad J\mathfrak{r}^{-1} \subset I_j \quad \Leftrightarrow \quad J\mathfrak{r}^{-1} \in \mathfrak{b}_{r^n}(\sigma(i),j),$$

where σ is the involution on $\{1,\ldots,h\}$ induced by $[I_{\sigma(i)}] = [I_i\mathfrak{r}]$. That is, there is a bijection between ideals $J \in \mathfrak{b}_{r^{n+1}}(i,j)$ such that $x \in \mathfrak{r}_r$ and ideals in I_j of relative norm r^n and with ideal class $[I\mathfrak{r}]$, and these are counted by $(W_\mathfrak{r}B_{r^n})(i,j)$. So

$$\begin{aligned}(B_{r^n}B_r)(i,j) &= \sum_{\substack{J \in \mathfrak{b}_{r^{n+1}}(i,j) \\ N_r(x) = 2r+1}} N_r(x) + \sum_{\substack{J \in \mathfrak{b}_{r^{n+1}}(i,j) \\ N_r(x) = r+1}} N_r(x) + \sum_{\substack{J \in \mathfrak{b}_{r^{n+1}}(i,j) \\ N_r(x) = 1}} N_r(x) \\ &= (2r+1)\cdot B_{r^{n-1}}(i,j) \\ &\qquad + (r+1)\cdot(W_\mathfrak{r}B_{r^n}(i,j) - B_{r^{n-1}}(i,j)) \\ &\qquad + 1\cdot(B_{r^{n+1}}(i,j) - W_\mathfrak{r}B_{r^n}(i,j)) \\ &= rB_{r^{n-1}}(i,j) + rW_\mathfrak{r}B_{r^n}(i,j) + B_{r^{n+1}}(i,j).\end{aligned}$$

This implies the statement of Proposition 4.13. □

4.3.3. $[\mathfrak{O} : \mathcal{O}] = r^2$. Let $\mathfrak{O}$ be a maximal order containing $\mathcal{O}$ and suppose that $[\mathfrak{O} : \mathcal{O}] = r^2$ for r a split prime of $\mathcal{A}$. In particular, $\mathfrak{O}_\ell = \mathcal{O}_\ell$ for all primes $\ell \neq r$. The main result of this section is Proposition 4.22, which computes the eigenvalues of eigenvectors for Brandt matrices in $\mathfrak{B}(\mathcal{O})$ by relating them to eigenvectors for Brandt matrices in $\mathfrak{B}(\mathfrak{O})$.

We begin by setting some notation. Fix representative invertible $\mathcal{O}$-ideals $I_1 = \mathcal{O}$, $I_2, \ldots, I_h$, and let $v_1, \ldots, v_h$ be the corresponding standard basis for $\mathfrak{M}(\mathcal{O})$. Also fix representative invertible $\mathfrak{O}$-ideals $I_1' = \mathfrak{O}, I_2', \ldots, I_{h'}'$, and let $w_1, \ldots, w_{h'}$ be the corresponding standard basis for $\mathfrak{M}(\mathfrak{O})$. We use B_n and $\widetilde{B}_n$ to denote Brandt matrices in $\mathfrak{B}(\mathcal{O})$ and $\mathfrak{B}(\mathfrak{O})$, respectively. Following the notation of Section 3.2, for $1 \leq i,j \leq h$ let

$$\mathfrak{b}_n(i,j) = \{J \subseteq_n I_j : [J] = [I_i]\}$$

as J varies over right $\mathcal{O}$-ideals, and for $1 \leq i,j \leq h'$ let

$$\widetilde{\mathfrak{b}}_n(i,j) = \{J \subseteq_n I_j' : [J] = [I_i']\}$$

as J varies over right $\mathfrak{O}$-ideals.

We define a map ρ from invertible right $\mathcal{O}$-ideals to invertible right $\mathfrak{O}$-ideals, given by ideal extension:

$$\rho(I) := I\mathfrak{O}.$$

Note that this is well-defined because invertible ideals are locally principal. This map will allow us to explicitly relate counts of $\mathcal{O}$-ideals to counts of $\mathfrak{O}$-ideals. See [Voi21, 26.6.1] for a discussion of the map ρ.

For $i \in \{1, \dots, h\}$, we define $\widetilde{i} \in \{1, \dots, h'\}$ by $[I'_{\widetilde{i}}] = [\rho(I_i)]$. Using the extension map ρ we may define an operator on Brandt modules:

$$\begin{aligned} E : \mathfrak{M}(\mathcal{O}) &\to \mathfrak{M}(\mathfrak{O}) \\ v_i &\mapsto w_{\widetilde{i}}. \end{aligned}$$

The map E will allow us to relate Brandt eigenvectors for $\mathfrak{B}(\mathcal{O})$ to Brandt eigenvectors for $\mathfrak{B}(\mathfrak{O})$. We begin by considering Brandt matrices B_n for n coprime to r.

Lemma 4.15. *For all n coprime to r,*

$$EB_n = \widetilde{B}_n E.$$

Proof. Fix $1 \leq i' \leq h'$ and $1 \leq j \leq h$. The quantity $(EB_n)(i', j)$ counts the set

$$\bigcup_{1 \leq i \leq h \,:\, \widetilde{i} = i'} \mathfrak{b}_n(i, j) = \left\{ J \subseteq_n I_j : [J\mathfrak{O}] = [I'_{i'}] \right\}$$

and the quantity $(\widetilde{B}_n E)(i', j)$ counts the set

$$\widetilde{\mathfrak{b}}_n(i', \widetilde{j}) = \left\{ J \subseteq_n I'_{\widetilde{j}} : [J] = [I'_{i'}] \right\}.$$

So it suffices to construct a bijection between these two sets. Recall that $[I'_{\widetilde{j}}] = [I_j\mathfrak{O}]$. As the size of $\widetilde{\mathfrak{b}}_n(i', \widetilde{j})$ does not depend on the choice of representative $I'_{\widetilde{j}}$ in its equivalence class, we may assume without loss of generality that $I'_{\widetilde{j}} = I_j\mathfrak{O}$.

We now consider the map ρ with restricted domain and codomain:

$$\begin{aligned} \rho : \bigcup_{i \,:\, \widetilde{i} = i'} \mathfrak{b}_n(i, j) &\to \widetilde{\mathfrak{b}}_n(i', \widetilde{j}), \\ J &\mapsto J\mathfrak{O}. \end{aligned} \tag{12}$$

For any J in the domain, its image is an invertible $\mathfrak{O}$-ideal contained in $\rho(I_j)$ with relative norm n and with $[\rho(J)] = [\rho(I_i)] = [I'_{i'}]$, showing that the map is well-defined. On the other hand, any I' in the codomain uniquely determines an ideal J in the domain: since n is coprime to r we must have $J_r = (I_j)_r$, and since $\mathfrak{O}_\ell = \mathcal{O}_\ell$ for all $\ell \neq r$ we must have $J_\ell = I'_\ell$. We can conclude by the local-global dictionary that Eq. (12) is a bijection, proving that $EB_n = \widetilde{B}_n E$. $\square$

We now consider Brandt matrices B_{r^k} for $k \geq 1$. In this case the map ρ of Eq. (12) is neither injective nor surjective. In order to quantify the failure of injectivity, we define a "splitting" operator

$$\begin{aligned} S : \mathfrak{M}(\mathfrak{O}) &\to \mathfrak{M}(\mathcal{O}) \\ w_{j'} &\mapsto \sum_{i=1}^{h} c_{i,j'} v_i, \qquad c_{i,j'} = \left| \left\{ J \in \rho^{-1}(I'_{j'}) : [J] = [I_i] \right\} \right|. \end{aligned}$$

Note that $c_{i,j'} = 0$ unless $\tilde{i} = j'$.

Lemma 4.16. *The operator $E \circ S$ acts on $\mathfrak{M}(\mathfrak{O})$ via scalar multiplication by $r^2 - r$.*

Proof. Let $I := I'_{j'}$. The operator $(E \circ S)$ multiplies $w_{j'}$ by $|\rho^{-1}(I)|$. We can compute this quantity by the local-global correspondence: a given $J \in \rho^{-1}(I)$ is determined at each prime ℓ by an $\mathcal{O}_\ell$-ideal J_ℓ with $J_\ell \mathfrak{O}_\ell = I_\ell$. For all $\ell \neq r$ we have $\mathcal{O}_\ell = \mathfrak{O}_\ell$ and so $J_\ell = I_\ell$. Now suppose we write $I_r = \alpha \mathfrak{O}_r$ for some $\alpha \in \mathcal{A}_r^\times$. Then for every $u \in \mathfrak{O}_r^\times$, $J_r := \alpha u \mathcal{O}_r$ defines an $\mathcal{O}_r$-ideal with $J_r \mathfrak{O}_r = I_r$. All such $\mathcal{O}_r$-ideals are obtained in this way, and we have $\alpha u \mathcal{O}_r = \alpha u' \mathcal{O}_r$ if and only if $u^{-1}u' \in \mathcal{O}_r^\times$. Thus $|\rho^{-1}(I)| = [\mathfrak{O}_r^\times : \mathcal{O}_r^\times]$. Using [Voi21, Lemma 26.6.7] we find that this value equals $r^2 - r$, because r is split in $\mathcal{A}$ and $\mathcal{O}$ has Eichler symbol -1 at r. ☐

In order to quantify the failure of surjectivity, we will show that for any $J \in \mathfrak{b}_{r^k}(i,j)$, $J\mathfrak{O}$ is contained not only in $I_j\mathfrak{O}$, but in fact in $rI_j\mathfrak{O}$.

Lemma 4.17. *If $J \subseteq I$ are invertible right $\mathcal{O}$-ideals with $J_r \neq I_r$, then $J \subseteq rI\mathfrak{O}$.*

Proof. It suffices to work locally at r. Let $I_r = \gamma \mathcal{O}_r$ for some $\gamma \in \mathcal{A}_r^\times$, so that $\gamma^{-1}J_r$ is a right $\mathcal{O}_r$-ideal strictly contained in $\mathcal{O}_r$. Since $\mathcal{O}_r / r\mathfrak{O}_r \simeq \mathbb{F}_{r^2}$ is a field, and $x \in \mathcal{O}_r$ is a unit if and only if $\mathrm{nrd}(x) \not\equiv 0 \bmod r$, we can conclude that $\mathcal{O}_r$ is the disjoint union of $\mathcal{O}_r^\times$ and $r\mathfrak{O}_r$. Since $\gamma^{-1}J_r$ does not contain a unit, it must be contained in $r\mathfrak{O}_r$. Hence $J_r \subseteq \gamma r \mathfrak{O}_r = rI_r\mathfrak{O}_r$. ☐

Remark 4.18. This can also be seen more explicitly. For convenience suppose r is odd. Up to isomorphism we may write $\mathfrak{O}_r = M_2(\mathbb{Z}_r)$, and $\mathcal{O}_r$ is the ring of matrices of the form $\left(\begin{smallmatrix} a & ub \\ b+rc & a+rd \end{smallmatrix}\right)$ for $a,b,c,d \in \mathbb{Z}_r$ and some fixed non-quadratic residue $u \in \mathbb{Z}_r^\times$. Following the proof above, we can then write $\gamma^{-1}J_r$ as $\alpha\mathcal{O}_r$ for some $\alpha \in \mathcal{O}_r \smallsetminus \mathcal{O}_r^\times$. The norm of α then has the form

$$(a^2 - ub^2) + r(ad - ubc) \in r\mathbb{Z}_r$$

for some $a,b,c,d \in \mathbb{Z}_r$, and since u is not a square in $\mathbb{Z}_r$, this forces $a,b \in r\mathbb{Z}_r$. Thus $\alpha \in r\mathfrak{O}_r$.

Proposition 4.19. *We have $B_r = 0$, and for $n \geq 2$ we have*

$$B_{r^n} = S\widetilde{B}_{r^{n-2}}E.$$

Proof. For $n \geq 1$, $1 \leq i,j \leq h$, and any $J \in \mathfrak{b}_{r^n}(i,j)$, we have $J_r \neq (I_j)_r$ and therefore J is contained in $rI_j\mathfrak{O}$ by Lemma 4.17. Then $\mathrm{nrd}(J)$ is a multiple of $r^2\mathrm{nrd}(I_j)$, so $B_r(i,j) = 0$.

Now suppose $n \geq 2$. For any $1 \leq i,j \leq h$ we define a variant of Eq. (12). As in the proof of Lemma 4.15, by changing representatives for the $\mathfrak{O}$-ideal classes, we may assume without loss of generality that $\rho(I_j) = I'_{\tilde{j}}$ so we have a modification of the ρ map

$$\frac{1}{r}\rho : \bigcup_{k\,:\,\tilde{k}=\tilde{i}} \mathfrak{b}_{r^n}(k,j) \to \widetilde{\mathfrak{b}}_{r^{n-2}}(\tilde{i},\tilde{j}), \qquad J \mapsto \tfrac{1}{r}\rho(J). \tag{13}$$

This is well-defined by Lemma 4.17. Now let I' be an element of the codomain, so that $I' \subseteq I_j\mathfrak{O}$. For any $J \in r\rho^{-1}(I')$ we have $\frac{1}{r}J \subseteq \frac{1}{r}J\mathfrak{O} = I'$, so that

$$J \subseteq rI' \subseteq I_j(r\mathfrak{O}) \subseteq I_j\mathcal{O} = I_j.$$

Since $[\rho(J)] = [I'] = [I'_{\tilde{i}}]$, we can conclude that J is in the domain of this map. Therefore every element of $r\rho^{-1}(I')$ is in the domain of Eq. (13).

By splitting $\mathfrak{b}_{r^n}(i,j)$ along fibres of the map Eq. (13), we have

$$B_{r^n}(i,j) = \sum_{I' \in \widetilde{\mathfrak{b}}_{r^{n-2}}(\tilde{i},\tilde{j})} \left| r\rho^{-1}(I') \cap \mathfrak{b}_{r^n}(i,j) \right|.$$

Now we compute the size of each term in the sum. Since we have checked that every $J \in r\rho^{-1}(I')$ is in $\mathfrak{b}_{r^n}(k,j)$ for some k, the only thing we have to check is the ideal class condition: that is, we must count the number of $J \in \rho^{-1}(I')$ with $[J] = [I_i]$. But this value depends only on the ideal class of I'. So replacing each I' with $I'_{\tilde{i}}$, we have

$$B_{r^n}(i,j) = c_{i,\tilde{i}}\left(\widetilde{B}_{r^{n-2}}(\widetilde{i},\widetilde{j})\right),$$

where $c_{i,\tilde{i}}$ is the number of $J \in \rho^{-1}(I'_{\tilde{i}})$ with $[J] = [I_i]$. Equivalently, $c_{i,\tilde{i}}$ is the component of v_i in $Sw_{\tilde{i}}$, so

$$B_{r^n}(i,j) = (S\widetilde{B}_{r^{n-2}})(i,\widetilde{j}) = (S\widetilde{B}_{r^{n-2}}E)(i,j). \qquad \square$$

As a consequence of this formula, together with Lemma 4.16 and Lemma 4.10, it is straightforward to derive the following description of the algebra generated by the Brandt matrices $B_{r^k} \in \mathfrak{B}(\mathcal{O})$.

Corollary 4.20. *The matrices B_{r^k} for all $k \geq 0$ are contained in the subalgebra of $\mathfrak{B}(\mathcal{O})$ generated by commuting matrices B_{r^2} and B_{r^3}: we have $B_r = 0$ and*

$$B_{r^{k+1}} = \frac{1}{r^2 - r} B_{r^k} B_{r^3} - r B_{r^{k-1}}$$

for all $k \geq 2$.

Remark 4.21. There is a natural generalization of Proposition 4.19. Suppose $\mathcal{O}$ is an order and ℓ is a prime such that $\mathcal{O}_\ell$ is Bass but not Eichler. Let $\widetilde{\mathcal{O}}$ be the order such that $\widetilde{\mathcal{O}}_\ell$ is the radical idealizer of $\mathcal{O}_\ell$, and $\widetilde{\mathcal{O}}_{\ell'} = \mathcal{O}_{\ell'}$ for $\ell' \neq \ell$. Let $\mathfrak{r}$ be the right $\mathcal{O}$-ideal where $\mathfrak{r}_\ell = \mathrm{rad}(\mathcal{O}_\ell)$ is the Jacobson radical of $\mathcal{O}_\ell$, and $\mathfrak{r}_{\ell'} = \mathcal{O}_{\ell'}$ for $\ell' \neq \ell$. Then $\mathfrak{r}$ is a two-sided invertible $\widetilde{\mathcal{O}}$-ideal, and using a similar argument we can show $B_{\ell^n} = 0$ for $1 \leq n < e$ and $B_{\ell^n} = SW_\mathfrak{r}\widetilde{B}_{\ell^{n-e}}E$ for $n \geq e$, where e is the valuation of $\mathrm{nrd}(\mathfrak{r})$ and $W_\mathfrak{r}$ is the operator on $\mathfrak{M}(\widetilde{\mathcal{O}})$ induced by $[I] \mapsto [I\mathfrak{r}]$.

We end this section by consolidating the above results into a statement that relates eigenvalues of Brandt matrices for $\mathcal{O}$ to eigenvalues of Brandt matrices for $\mathfrak{O}$. Given any order $\mathcal{O}$ in $\mathcal{A}$, if $u \in \mathfrak{M}(\mathcal{O})$ is an eigenvector for every Brandt matrix in $\mathfrak{B}(\mathcal{O})$, we call u a **Brandt eigenvector** and let $\lambda_n(u)$ denote the eigenvalue of u under $B_n \in \mathfrak{B}(\mathcal{O})$.

Proposition 4.22. *Let $\mathcal{O}$ be a submaximal order contained in a maximal order $\mathfrak{O}$ with $[\mathfrak{O} : \mathcal{O}] = r^2$ for a split prime r of $\mathcal{A}$. If $u \in \mathfrak{M}(\mathcal{O})$ is a Brandt eigenvector, then $\lambda_r(u) = 0$ and one of the following holds.*

(a) $\lambda_{r^k}(u) = 0$ *for all* $k \geq 2$.
(b) *There exists $w \in \mathfrak{M}(\mathfrak{O})$, an eigenvector for $\widetilde{B}_n \in \mathfrak{B}(\mathfrak{O})$ for all $n \geq 1$, satisfying $\lambda_{r^k}(u) = (r^2 - r)\lambda_{r^{k-2}}(w)$ for all $k \geq 2$ and $\lambda_n(u) = \lambda_n(w)$ for n coprime to r.*

Proof. If $Eu = 0$ then $B_{r^k}u = 0$ for all $k \geq 1$ by Proposition 4.19, so $\lambda_{r^k}(u) = 0$ for all $k \geq 1$. So from now on suppose $w := Eu \neq 0$. Then from Lemma 4.15 we have $\widetilde{B}_n Eu = \lambda_n(u)Eu$ for n coprime to r, and by applying E to both sides of Proposition 4.19 we have $B_r(u) = 0$ and for $k \geq 2$ we have

$$ES\widetilde{B}_{r^{k-2}}(Eu) = \lambda_{r^k}(u)(Eu).$$

We have $ES = r^2 - r$ by Lemma 4.16. This proves that Eu is an eigenvector for $\widetilde{B}_n$ for any $n \geq 1$, with $\lambda_r(u) = 0$ and $\lambda_{r^k}(u) = (r^2 - r)\lambda_{r^{k-2}}(Eu)$ for $k \geq 2$. $\square$

Note that in fact the w from Proposition 4.22(b) is an eigenvector for the full Brandt algebra $\mathfrak{B}(\mathfrak{O})$, a fact which can be derived from the proofs of Lemma 6.8(c) (in the case u is a multiple of the genus vector of $\mathcal{O}$) and Lemma 6.15 (in the case u is cuspidal).

5. Background: modular forms

5.1. **Modular forms and Hecke operators.** There are many excellent general references for modular forms; we mention [DS05] and [Ogg69] that the reader wishing for more details regarding the fundamentals may consult.

The group $\mathrm{SL}_2(\mathbb{Z})$ acts on the upper half plane $\mathbb{H} = \{z \in \mathbb{C} : \Im(z) > 0\}$ by fractional linear transformations as $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) z := \frac{az+b}{cz+d}$. Via $z \mapsto \left(\begin{smallmatrix} z \\ 1 \end{smallmatrix}\right)$, we can view this as derived from the action of $\mathrm{SL}_2(\mathbb{C})$ on $\mathbb{P}^1(\mathbb{C})$. This interpretation allows us to extend the action of $\mathrm{SL}_2(\mathbb{Z})$ to $\mathbb{H}^* = \mathbb{H} \sqcup \mathbb{P}^1(\mathbb{Q})$. One refers to $(1\ 0)^t$ as ∞.

A **congruence subgroup** of $\mathrm{SL}_2(\mathbb{Z})$ is a subgroup $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$ such that Γ contains the kernel of the reduction-by-N map $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ for some integer N. For $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$ a congruence subgroup, the **modular curve** $X(\Gamma)$ is the quotient $\Gamma \backslash \mathbb{H}^*$. The **cusps** of a congruence subgroup Γ are the Γ-orbits in $\mathbb{P}^1(\mathbb{Q})$.

Fixing an integer $N \geq 1$, we will be particularly interested in the congruence subgroup of level N which is the preimage of the Borel subgroup of $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$, denoted

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \pmod N \right\}.$$

A **modular form** of **weight** $k \in \mathbb{Z}_{\geq 0}$ for a congruence subgroup Γ is a holomorphic function $f : \mathbb{H} \to \mathbb{C}$, which is bounded in neighborhoods of the cusps of Γ and satisfies the transformation law $f(\gamma z) = (cz+d)^k f(z)$ for all $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$. The set of modular forms of weight k for $\Gamma_0(N)$ is a finite dimensional vector space over $\mathbb{C}$ denoted $\mathcal{M}_k(\Gamma_0(N))$. Every modular form $f(z)$ for $\Gamma_0(N)$ has a Fourier expansion at ∞

$$f(q) = \sum_{n=0}^{\infty} a_n q^n, \qquad q = q(z) = e^{2i\pi z} \quad , \quad a_n \in \mathbb{C}. \tag{14}$$

This is often referred to as its q**-expansion**, and we will often write modular forms in terms of their q-expansions. We will say $f \in \mathcal{M}_k(\Gamma_0(N))$ is **normalized** if $a_1 = 1$. Since there exist nonzero modular forms with $a_1 = 0$, it is not true in general that every element of $\mathcal{M}_k(\Gamma_0(N))$ is a scalar multiple of a normalized form.

A modular form $f \in \mathcal{M}_k(\Gamma_0(N))$ is a **cusp form** if the suitably defined q-series expansion of $f(\gamma z)$ has zero constant coefficient for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ (equivalently, f vanishes at the cusps of $\Gamma_0(N)$). We denote by $\mathcal{S}_k(\Gamma_0(N))$ the subspace of cusp forms of $\mathcal{M}_k(\Gamma_0(N))$.

Fixing a positive integer N and an integer $k \geq 0$, there is a special family of linear operators $\{T_n\}_{n \geq 1}$ acting on $\mathcal{M}_k(\Gamma_0(N))$ known as the **Hecke operators**. For an integer n coprime to N, the nth Hecke operator acts on the q-expansion of a modular form $f(q) = \sum_{m=0}^{\infty} a_m q^m \in \mathcal{M}_k(\Gamma_0(N))$ as

$$(T_n f)(q) = \sum_{m=0}^{\infty} \sum_{d \mid \gcd(m,n)} d^{k-1} a_{\frac{mn}{d^2}} q^m \tag{15}$$

(cf. [DS05, Proposition 5.3.1]).

Definition 5.1. Fix a positive integer N. We define the **anemic Hecke algebra** $\mathbb{T}_N^{\mathrm{an}}$ to be the $\mathbb{Z}$-algebra generated by the Hecke operators T_n on $\mathcal{M}_k(\Gamma_0(N))$ for all $n \geq 1$ with $\gcd(n, N) = 1$. We say $f \in \mathcal{M}_k(\Gamma_0(N))$ is a $\mathbb{T}_N^{\mathrm{an}}$**-eigenvector** if f is an eigenvector for all $T \in \mathbb{T}_N^{\mathrm{an}}$.

When the level N is clear from context we may write $\mathbb{T}^{\mathrm{an}}$ instead of $\mathbb{T}_N^{\mathrm{an}}$. There is also a larger **Hecke algebra** $\mathbb{T} = \mathbb{T}_N$ of operators on $\mathcal{M}_k(\Gamma_0(N))$ generated as a $\mathbb{Z}$-algebra by suitably defined Hecke operators T_n for all $n \geq 1$; see [DS05, Section 5.3] for a full definition. A $\mathbb{T}$-eigenvector, that is a modular form $f \in \mathcal{M}_k(\Gamma_0(N))$ that is an eigenvector for all $T \in \mathbb{T}$, is often called a **Hecke eigenform**.

Remark 5.2. Hecke operators may also be defined on other spaces of modular forms, but we have restricted our presentation to the case relevant to this paper. For example, [DS05] discusses Hecke operators for $\Gamma_1(N)$, and the results we need can be obtained by noting that $\mathcal{M}_2(\Gamma_0(N))$ is the set of modular forms for $\Gamma_1(N)$ with trivial character (see [DS05, Exercise 4.3.3(a)]). For instance, note that the diamond operators $\langle d\rangle$ for $(d, N) = 1$ (see [DS05, §5.2]) act as the identity on $\mathcal{M}_2(\Gamma_0(N))$.

The action of $\mathbb{T}$ on $\mathcal{M}_k(\Gamma_0(N))$ preserves the cuspidal subspace $\mathcal{S}_k(\Gamma_0(N))$. If n_1, n_2 are a pair of coprime integers, then the Hecke operators T_{n_1} and T_{n_2} satisfy $T_{n_1}T_{n_2} = T_{n_1n_2}$. If ℓ is a prime coprime to N, then the Hecke operators T_{ℓ^n} satisfy the recursion relation $T_{\ell^{n+1}} = T_\ell T_{\ell^n} - \ell^{k-1}T_{\ell^{n-1}}$, where T_1 is the identity operator. In particular, $\mathbb{T}^{\mathrm{an}}$ consists of pairwise commuting linear operators (cf. [DS05, §5.3]). For a congruence subgroup Γ, we can make $\mathcal{S}_k(\Gamma)$ into an inner product space with the **Petersson inner product** $\langle \bullet, \bullet\rangle : \mathcal{S}_k(\Gamma) \times \mathcal{S}_k(\Gamma) \to \mathbb{C}$, defined by

$$\langle f, g\rangle = \frac{1}{V_\Gamma}\int_{X(\Gamma)} f(z)\overline{g(z)}\left(\Im(z)\right)^k d\mu(z),$$

where $d\mu(z) = \frac{dxdy}{y^2}$ is the hyperbolic measure on the upper half plane and V_Γ denotes the volume of the modular curve $X(\Gamma)$ with respect to this measure.

If n is coprime to N, then the Hecke operator T_n, viewed as a linear operator on $\mathcal{S}_k(\Gamma_0(N))$, is self-adjoint with respect to the Petersson inner product (cf. [DS05, Theorem 5.5.3] and Remark 5.2). In particular, this implies that $\mathcal{S}_k(\Gamma_0(N))$ has a basis of $\mathbb{T}^{\mathrm{an}}$-eigenvectors with real eigenvalues.

For positive integers $M \mid N$ there is an inclusion $\Gamma_0(N) \hookrightarrow \Gamma_0(M)$ giving an inclusion of modular forms $\mathcal{M}_k(\Gamma_0(M)) \hookrightarrow \mathcal{M}_k(\Gamma_0(N))$. Moreover, if $f \in \mathcal{M}_k(\Gamma_0(M))$ and $N = Mm$, then $f(z) \mapsto f(mz)$ is also an inclusion $\mathcal{M}_k(\Gamma_0(M)) \to \mathcal{M}_k(\Gamma_0(N))$; in terms of the q-expansion, we write this as $f(q) \mapsto f(q^m)$. For a fixed integer N, the **old subspace** $\mathcal{S}_k^{\mathrm{old}}(\Gamma_0(N))$ of $\mathcal{S}_k(\Gamma_0(N))$ is the subspace generated by

$$\bigcup_{\substack{M\mid N,\\ M\neq N}}\ \bigcup_{m\mid \frac{N}{M}} \{f(q^m) \ :\ f \in \mathcal{S}_k(\Gamma_0(M))\},$$

and the **new subspace** $\mathcal{S}_k^{\mathrm{new}}(\Gamma_0(N))$ of $\mathcal{S}_k(\Gamma_0(N))$ is, per definition, the orthogonal complement of $\mathcal{S}_k^{\mathrm{old}}(\Gamma_0(N))$ in $\mathcal{S}_k(\Gamma_0(N))$ with respect to the Petersson inner product. Both $\mathcal{S}_k^{\mathrm{new}}(\Gamma_0(N))$ and $\mathcal{S}_k^{\mathrm{old}}(\Gamma_0(N))$ are stable under the action of the full Hecke algebra $\mathbb{T}$ (cf. [DS05, Proposition 5.6.2]).

A modular form $f \in \mathcal{S}_k(\Gamma_0(N))$ is called a **newform** if it is a normalized $\mathbb{T}^{\mathrm{an}}$-eigenvector in the new subspace $\mathcal{S}_k^{\mathrm{new}}(\Gamma_0(N))$. In fact $\mathcal{S}_k^{\mathrm{new}}(\Gamma_0(N))$ has a basis of newforms, which are all eigenvectors for the full Hecke algebra $\mathbb{T}$ with all eigenvalues in $\mathbb{R}$ [AL70, Theorem 5].

If $f(q) = \sum a_n q^n$ is a newform in $\mathcal{S}_2(\Gamma_0(N))$, the *Deligne bound* (credited in this case to Eichler and Shimura) gives a precise upper bound on the sizes of coefficients: $|a_\ell| \le 2\sqrt{\ell}$ for all primes ℓ (including $\ell \mid N$). Since a newform f of level N has an Euler product

$$L(s, f) = \prod_{p\nmid N}\left(1 - a_p p^{-s} + p^{k-1-2s}\right)^{-1}\prod_{p\mid N}\left(1 - a_p p^{-s}\right)^{-1},$$

it follows that this bound can be generalized to all coefficients.

Lemma 5.3 (Deligne bound). *If $f(q) = \sum a_n q^n$ is a newform of weight 2 (and any level), then*

$$|a_n| \le \sigma_0(n)\sqrt{n},$$

where $\sigma_0(n)$ is the number of positive divisors of n.

At certain points in the paper we study cusp forms that are $\mathbb{T}^{\mathrm{an}}$-eigenvectors, but which do not necessarily lie in the new subspace. We can use the following result to decompose these forms.

Lemma 5.4. *Let $g \in \mathcal{S}_2(\Gamma_0(N))$ be a $\mathbb{T}^{\mathrm{an}}$-eigenvector. Then for some $M \mid N$, there exists a single newform $f \in \mathcal{S}_2^{\mathrm{new}}(\Gamma_0(M))$ and coefficients $c_d \in \mathbb{C}$ such that*

$$g(q) = \sum_{d \mid \frac{N}{M}} c_d f(q^d).$$

Proof. By [AL70, Theorem 5], $\mathcal{S}_2(\Gamma_0(N))$ has a basis consisting of cusp forms $f(q^d)$, where M varies over divisors of N, f varies over newforms in $\mathcal{S}_2^{\mathrm{new}}(\Gamma_0(M))$, and d varies over divisors of $\frac{N}{M}$. We use this to write

$$g(q) = \sum_{i=1}^{k} c_i f_i(q^{d_i})$$

for some $k \geq 1$, $c_i \in \mathbb{C} \smallsetminus \{0\}$, $M_i \mid N$, $f_i \in \mathcal{S}_2^{\mathrm{new}}(\Gamma_0(M_i))$, and $d_i \mid \frac{N}{M_i}$. If for some $1 \leq i, j \leq k$ we have $f_i \neq f_j$, then again by [AL70, Theorem 5], there exist infinitely many primes ℓ (and in particular, at least one prime ℓ coprime to N) for which f_i and f_j have distinct eigenvalues under T_ℓ. But g is an eigenvector under T_ℓ, so all f_i must have the same eigenvalue under T_ℓ, a contradiction. Hence all f_i are equal to a single newform f. □

5.2. **Theta functions.** Let Λ be an even integral lattice with respect to a bilinear form $\langle \cdot, \cdot \rangle$ (see Section 2.1). The **theta function** $\theta_\Lambda(q)$ of Λ is the formal q-series

$$\theta_\Lambda(q) = \sum_{x \in \Lambda} q^{\frac{1}{2}\langle x, x \rangle} = \sum_{n=0}^{\infty} a_n(\Lambda) q^n, \quad \text{where} \quad a_n(\Lambda) = |\{x \in \Lambda \ : \ \tfrac{1}{2}\langle x, x \rangle = n\}|.$$

It is well known that the theta function of Λ is a modular form of weight $\frac{\mathrm{rank}(\Lambda)}{2}$ and level equal to the level of the lattice (Definition 2.1), with a certain character; see e.g. [Ogg69, Theorem 20].

We will be specifically focused on the case where $\Lambda = \mathcal{O}$ is an order in a definite quaternion algebra $\mathcal{A}$ equipped with the standard bilinear form (trace form) on $\mathcal{A}$, so that $a_n(\mathcal{O})$ is the number of elements $x \in \mathcal{O}$ with $\mathrm{nrd}(x) = n$. In this case, the theta function of $\mathcal{O}$ is a modular form of weight 2 and level $N = \mathrm{nrd}(\mathcal{O}^\sharp)^{-1}$ by Lemma 2.5.

Recall from Section 2.5 that $\mathfrak{M}(\mathcal{O})$ denotes the Brandt module of $\mathcal{O}$. The following two results are well known throughout the literature (see for instance [Koh01, §3.3], or [Piz80a, Theorem 2.14] and [Piz80a, Proposition 2.23]), though they are typically stated for a more restrictive class of orders. Since we are working in a more general setting, we include proofs for completeness, though these proofs largely follow those appearing in the literature.

Fix the standard basis $\{v_1, \dots, v_h\}$ for $\mathfrak{M}(\mathcal{O})$ corresponding to the enumeration $\{[I_1] = [\mathcal{O}], [I_2], \dots, [I_h]\}$ of the right ideal classes of $\mathcal{O}$ as in Section 2.5.

Lemma 5.5. *Let $\mathcal{O}$ be an order in a definite rational quaternion algebra, and suppose $N := \mathrm{nrd}(\mathcal{O}^\sharp)^{-1}$. There exists a symmetric bilinear pairing*

$$\Theta : \mathfrak{M}(\mathcal{O}) \times \mathfrak{M}(\mathcal{O}) \to \mathcal{M}_2(\Gamma_0(N))$$
$$\Theta(v, w) := \sum_{n=0}^{\infty} \langle v, B_n w \rangle q^n.$$

In particular, $2\Theta(v_i, v_j)$ is the theta function of the lattice $I_j I_i^{-1}$ with normalized trace form (Definition 2.3)*, and $2\Theta(v_1, v_1) = \theta_{\mathcal{O}}$.*

Proof. We begin by considering Θ only as a map valued in the formal power series ring $\mathbb{R}[\![q]\!]$. The inner product on $\mathfrak{M}(\mathcal{O})$ is symmetric and bilinear, and B_n is self-adjoint (Lemma 2.9), so Θ is symmetric and bilinear. To prove that the image lies in $\mathcal{M}_2(\Gamma_0(N))$, it suffices to prove that $\Theta(v_i, v_j) \in \mathcal{M}_2(\Gamma_0(N))$ for all $1 \le i, j \le h$.

For arbitrary i, j, we have $\langle v_i, B_n v_j\rangle = w_i B_n(i, j)$, so that by Eq. (4), $2\Theta(v_i, v_j)$ is the theta function of the lattice $I_j I_i^{-1}$ with normalized trace form; this proves the final sentence in the statement of the lemma, since we took $I_1 = \mathcal{O}$. Observe that the level of an integral quadratic form can be computed locally (for instance, because the dual lattice can be defined locally, and integrality of a quadratic form is a local property). Locally at each prime ℓ, we can write $(I_i)_\ell = \alpha\mathcal{O}_\ell$ and $(I_j)_\ell = \beta\mathcal{O}_\ell$ for some $\alpha, \beta \in \mathcal{A}_\ell^\times$, so that $(I_j I_i^{-1})_\ell = \beta\mathcal{O}_\ell\alpha^{-1}$. Now let $u := \frac{\operatorname{nrd}(\alpha)}{\operatorname{nrd}(I_i)}\frac{\operatorname{nrd}(I_j)}{\operatorname{nrd}(\beta)}$, which is in $\mathbb{Z}_\ell^\times$. Then the map $x \mapsto \beta x\alpha^{-1}$ is an isometry from $\mathcal{O}_\ell$ to the lattice $(I_j I_i^{-1})_\ell$ equipped with the bilinear form $(x, y) \mapsto u\langle x, y\rangle_{I_j I_i^{-1}}$. Multiplying the bilinear form of a $\mathbb{Z}_\ell$-lattice by a unit does not affect the level. So taking all ℓ together, this proves that $I_j I_i^{-1}$ with the normalized trace form has the same level as $\mathcal{O}$.

Since the level of $\mathcal{O}$ is equal to $N = \operatorname{nrd}(\mathcal{O}^\sharp)^{-1}$ by Lemma 2.5, we have $\Theta(v_i, v_j) \in \mathcal{M}_2(\Gamma_1(N))$ with character $d \mapsto \left(\frac{\Delta}{d}\right)$ by [Ogg69, Theorems 20, 20+] (see also [Voi21, 40.4.5]), where $\Delta = \det(\langle b_i, b_j\rangle_{I_j I_i^{-1}})_{1\le i,j\le 4}$ for a basis (b_1, b_2, b_3, b_4) of $I_j I_i^{-1}$. Since Δ is a perfect square by Lemma 2.4, the character is trivial and $\Theta(v_i, v_j) \in \mathcal{M}_2(\Gamma_0(N))$. $\square$

We next prove a compatibility between Brandt operators B_n and Hecke operators T_n for n coprime to N. This compatibility does not hold more generally: even if u is an eigenvector for all Brandt matrices in $\mathfrak{B}(\mathcal{O})$, $\Theta(u, u)$ may fail to be an eigenvector for Hecke operators T_ℓ with $\ell \mid N$. See Lemma 6.14 for a counterexample.

Lemma 5.6. *For all $v, w \in \mathfrak{M}(\mathcal{O})$ and all $n \ge 1$ with $\gcd(n, N) = 1$, we have*

$$T_n\Theta(v, w) = \Theta(v, B_n w).$$

Before we begin the proof we set up some notation. Since $\mathfrak{B}(\mathcal{O})$ is a commutative algebra (Theorem 3.1) of operators that are self-adjoint with respect to the inner product on $\mathfrak{M}(\mathcal{O})$, there exists a simultaneous eigenbasis $u_1, \ldots, u_h \in \mathfrak{M}(\mathcal{O})$ consisting of vectors that are pairwise orthogonal with respect to the inner product. We can normalize this basis so that $\langle u_i, u_i\rangle = 1$ for all i; we will refer to this basis as a **normalized Brandt eigenbasis**, and each element of the normalized Brandt eigenbasis will be called a **Brandt eigenvector**. In the case of a submaximal $\mathcal{O}$ in a maximal $\mathfrak{O}$ with $[\mathfrak{O} : \mathcal{O}] = r$, we will further impose the condition that each u_i in a normalized Brandt eigenbasis is an eigenvector of the operator $W_{\mathfrak{r}}$ defined by Eq. (11); this can be ensured by Lemma 4.12. Observe that for any $i \ne j$ we have

$$\Theta(u_i, u_j) = \sum_{n=0}^{\infty}\langle u_i, B_n u_j\rangle q^n = 0, \tag{16}$$

because for all $n \ge 0$, $B_n u_j$ is a scalar multiple of u_j and hence orthogonal to u_i.

Proof of Lemma 5.6. Let $u_1, \ldots, u_h$ be a Brandt eigenbasis for $\mathfrak{M}(\mathcal{O})$. For all $n \ge 0$ define the linear functional λ_n by $B_n u_i = \lambda_n(u_i)u_i$ for each i. Let $v = a_1u_1 + \cdots + a_hu_h$ and $w = b_1u_1 + \cdots + b_hu_h$. By bilinearity together with Eq. (16), we have

$$\Theta(v, w) = \sum_{i=1}^{h} a_ib_i\Theta(u_i, u_i), \qquad \Theta(v, B_nw) = \sum_{i=1}^{h} a_ib_i\lambda_n(u_i)\Theta(u_i, u_i).$$

So it suffices to prove the following: for any Brandt eigenvector $u \in \mathfrak{M}(\mathcal{O})$, and any n coprime to N, if we set

$$f_u := \frac{\Theta(u, u)}{\langle u, u\rangle} = \sum_{n=0}^{\infty}\lambda_n(u)q^n,$$

then $T_n f_u = \lambda_n(u) f_u$. Indeed, since $\langle u_i, u_i \rangle = 1$ for each i, we have that $\Theta(u_i, u_i) = f_{u_i}$ and hence by linearity,

$$T_n \Theta(v, w) = \sum_{i=1}^{h} a_i b_i T_n f_{u_i} = \sum_{i=1}^{h} a_i b_i \lambda_n(u_i) f_{u_i} = \Theta(v, B_n w).$$

First consider the case $n = \ell$ for some prime $\ell \nmid N$. As seen above, $f_u \in \mathcal{M}_2(\Gamma_0(N))$, so by Eq. (15), T_ℓ acts on f_u as

$$T_\ell f_u = \sum_{m=0}^{\infty} \left(\lambda_{\ell m}(u) + \ell \lambda_{\frac{m}{\ell}}(u) \right) q^m. \tag{17}$$

Here for convenience we set $\lambda_c(u) = 0$ for any c that is not a non-negative integer: that is, if $\ell \nmid m$ then the coefficient of q^m in Eq. (17) above is $\lambda_{\ell m}(u)$. Now for each $m \geq 1$ write $m = \ell^k m'$ with m' coprime to ℓ. Since $B_{\ell^k m'} = B_{\ell^k} B_{m'}$ by Lemma 3.6 we have the equality $\lambda_{\ell^k m'}(u) = \lambda_{\ell^k}(u) \lambda_{m'}(u)$ of Brandt eigenvalues (and a similar decomposition for $\lambda_{\frac{m}{\ell}}(u)$ if $\ell \mid m$), so the coefficient of q^m in Eq. (17) equals

$$\lambda_{m'}(u) \left(\lambda_{\ell^{k+1}}(u) + \ell \lambda_{\ell^{k-1}}(u) \right).$$

Now by an immediate simplification if $k = 0$, and by Lemma 4.10 if $k \geq 1$, this in turn equals

$$\lambda_{m'}(u) \lambda_{\ell^k}(u) \lambda_\ell(u) = \lambda_\ell(u) \lambda_m(u).$$

Hence $T_\ell f_u = \lambda_\ell(u) f_u$.

Now for the Hecke operator T_{ℓ^k} with $\ell \nmid N$ and $k \geq 1$, we induct on k. Using the relation $T_\ell T_{\ell^k} = T_{\ell^{k+1}} + \ell T_{\ell^{k-1}}$ and the inductive assumption, we obtain

$$\begin{aligned} T_{\ell^{k+1}} f_u &= (T_\ell T_{\ell^k} - \ell T_{\ell^{k-1}}) f_u \\ &= \left(\lambda_\ell(u) \lambda_{\ell^k}(u) - \ell \lambda_{\ell^{k-1}}(u) \right) f_u. \end{aligned}$$

By the Brandt matrix relation $B_\ell B_{\ell^k} = B_{\ell^{k+1}} + \ell B_{\ell^{k-1}}$, we can rewrite this as $\lambda_{\ell^{k+1}}(u) f_u$, completing the induction. Finally, for the general case for $n = \prod_i \ell_i^{k_i}$ with $\gcd(n, N) = 1$, we use multiplicativity of Hecke operators and of Brandt matrices. □

5.3. **The cuspidal subspace.** Recall from Section 4.2 that the Brandt module of an order $\mathcal{O} \subseteq \mathcal{A}$ contains a *genus subspace* $\mathfrak{M}_{\mathrm{G}}(\mathcal{O})$. Let $\mathfrak{M}_{\mathrm{G}}(\mathcal{O})^\perp$ denote the subspace of $\mathfrak{M}(\mathcal{O})$ of vectors orthogonal to $\mathfrak{M}_{\mathrm{G}}(\mathcal{O})$ with respect to the Brandt module inner product. We call $\mathfrak{M}_{\mathrm{G}}(\mathcal{O})^\perp$ the **cuspidal subspace** of $\mathfrak{M}(\mathcal{O})$, and refer to elements of $\mathfrak{M}_{\mathrm{G}}(\mathcal{O})^\perp$ as being cuspidal, motivated by the following lemma. As with Lemmas 5.5 and 5.6, this lemma also occurs in the literature for a more restricted class of orders (cf. [Piz80a, Theorem 2.21], [Piz80b, Proposition 5.27], and [Koh01, §3.3]), and we provide a proof of the more general case for completeness, following the proofs in the literature.

Lemma 5.7. *The map Θ from Lemma 5.5 sends $\mathfrak{M}(\mathcal{O}) \times \mathfrak{M}_{\mathrm{G}}(\mathcal{O})^\perp$ and $\mathfrak{M}_{\mathrm{G}}(\mathcal{O})^\perp \times \mathfrak{M}(\mathcal{O})$ into the space $\mathcal{S}_2(\Gamma_0(N))$ of cusp forms in $\mathcal{M}_2(\Gamma_0(N))$.*

Proof. By symmetry of Θ it suffices to prove $\mathfrak{M}(\mathcal{O}) \times \mathfrak{M}_{\mathrm{G}}(\mathcal{O})^\perp$ is sent to cusp forms. A straightforward computation shows that $\mathfrak{M}_{\mathrm{G}}(\mathcal{O})^\perp$ is spanned by vectors of the form $v_i - v_j$ for $1 \leq i, j \leq h$ satisfying $I_j \in \mathrm{Gen}(I_i)$ (each of these vectors is orthogonal to every element of $\mathfrak{M}_{\mathrm{G}}(\mathcal{O})$ by construction, and a counting argument shows that together with $\mathfrak{M}_{\mathrm{G}}(\mathcal{O})$ they span the full Brandt module). So by bilinearity of Θ, it suffices to prove that $\Theta(v_k, v_i) - \Theta(v_k, v_j)$ is a cusp form for all $1 \leq i, j, k \leq h$ with $I_j \in \mathrm{Gen}(I_i)$.

Since I_i and I_j (with their associated normalized trace forms) are in the same genus, $I_i I_k^{-1}$ and $I_j I_k^{-1}$ are also in the same genus. To see this, let ℓ be an arbitrary prime, and let $\phi : (I_i)_\ell \to (I_j)_\ell$ be an isometry. Write $(I_k)_\ell = \gamma\mathcal{O}_\ell$ for some $\gamma \in \mathcal{A}_\ell^\times$. Then for any $x, y \in (I_i I_k^{-1})_\ell$, we have $x\gamma, y\gamma \in (I_i)_\ell$, so

$$\begin{aligned}\frac{\operatorname{trd}((\phi(x\gamma)\gamma^{-1})\overline{(\phi(y\gamma)\gamma^{-1})})}{\operatorname{nrd}(I_j I_k^{-1})} &= \frac{\operatorname{nrd}(\gamma)^{-1}}{\operatorname{nrd}(I_k)^{-1}} \frac{\operatorname{trd}(\phi(x\gamma)\overline{\phi(y\gamma)})}{\operatorname{nrd}(I_j)} \\ &= \frac{\operatorname{nrd}(\gamma)^{-1}}{\operatorname{nrd}(I_k)^{-1}} \frac{\operatorname{trd}((x\gamma)\overline{(y\gamma)})}{\operatorname{nrd}(I_i)} \\ &= \frac{\operatorname{trd}(x\overline{y})}{\operatorname{nrd}(I_i I_k^{-1})}.\end{aligned}$$

Thus $x \mapsto \phi(x\gamma)\gamma^{-1}$ defines an isometry $(I_i I_k^{-1})_\ell \to (I_j I_k^{-1})_\ell$, with bilinear forms respectively induced by the normalized trace forms on $I_i I_k^{-1}$ and $I_j I_k^{-1}$. It is a classical result that if two lattices are in the same genus then the difference of their theta functions is a cusp form (see e.g. [Sie35, p. 577]). So we can conclude that $\Theta(v_k, v_i) - \Theta(v_k, v_j)$, and hence every modular form in the image of $\mathfrak{M}(\mathcal{O}) \times \mathfrak{M}_{\mathrm{G}}(\mathcal{O})^\perp$ under Θ, is a cusp form. $\square$

The space of genus vectors for a submaximal order $\mathcal{O}$ was determined in Section 4.2: we found that $\mathfrak{M}_{\mathrm{G}}(\mathcal{O})$ is typically 1-dimensional (spanned by the vector e), but in certain special cases it is 2-dimensional (spanned by e and e'). As a consequence of Lemma 5.7, we have that $\Theta(u, u) \in \mathcal{S}_2(\Gamma_0(N))$ for any Brandt eigenvector u orthogonal to e (and to e' if defined).

6. Decomposing theta functions of quaternion orders

For the remainder of the paper, $\mathcal{A}$ is a *definite* quaternion algebra over $\mathbb{Q}$. As in the discussion immediately following Lemma 5.6, we define a **normalized Brandt eigenbasis** of $\mathfrak{M}(\mathcal{O})$ to be a basis $u_1, \dots, u_h$ which is orthonormal with respect to the standard inner product on $\mathfrak{M}(\mathcal{O})$, and such that each u_i is an eigenvector for every operator in $\mathfrak{B}(\mathcal{O})$; such a basis exists by Theorem 3.1. If u is a Brandt eigenvector, we write $\lambda_n(u)$ to denote the eigenvalue of u under the action of the Brandt matrix B_n.

6.1. Theta function decomposition.

Since Theorem 3.1 guarantees the existence of a normalized Brandt eigenbasis, the following result is a more precise version of Theorem 1.6. We will use this result to decompose $\theta_{\mathcal{O}}$ into individual modular forms that are easier to analyze, from which we can obtain explicit bounds on the number of elements of each norm in a quaternion order.

Theorem 6.1. *Let $\mathcal{O}$ be an order in a definite quaternion algebra $\mathcal{A}$ over $\mathbb{Q}$, $N := \operatorname{nrd}(\mathcal{O}^\sharp)^{-1}$, and let h be the class number of $\mathcal{O}$. Let $u_1, \dots, u_h$ be a normalized Brandt eigenbasis of $\mathfrak{M}(\mathcal{O})$, and set $f_i := \Theta(u_i, u_i)$ for each $i = 1, \dots, h$. Then each f_i is a normalized $\mathbb{T}^{\mathrm{an}}$-eigenvector, and*

$$\theta_{\mathcal{O}}(q) = c_1 f_1(q) + \cdots + c_h f_h(q)$$

for some non-negative real numbers $c_1, \dots, c_h$ with $c_1 + \cdots + c_h = |\mathcal{O}^\times|$.

Proof. By Lemma 5.6, we can conclude that $f_i := \Theta(u_i, u_i)$ is a $\mathbb{T}^{\mathrm{an}}$-eigenvector for all $i = 1, \dots, h$. The coefficient of q in f_i is $\langle u_i, u_i \rangle = 1$ because B_1 is the identity matrix, so f_i is normalized. We are left to show that we can write

$$\theta_{\mathcal{O}} = c_1 f_1 + \cdots + c_h f_h$$

for non-negative real coefficients $c_1, \dots, c_h$ satisfying $c_1 + \cdots + c_h = |\mathcal{O}^\times|$.

Let $v_1, \dots, v_h$ be the standard basis of $\mathfrak{M}(\mathcal{O})$, ordered so that $I_1 = \mathcal{O}$. By Lemma 5.5, we have $2\Theta(v_1, v_1) = \theta_{\mathcal{O}}$. Let $(c_{ij})_{i,j} \in \mathrm{GL}_h(\mathbb{R})$ be a change of basis matrix so that

$$v_i = c_{i1} u_1 + \cdots + c_{ih} u_h$$

for all $i = 1, \dots, h$. By Eq. (16) we have $\Theta(u_i, u_j) = 0$ for all $i \neq j$, so together with bilinearity and symmetry of Θ we have

$$\theta_{\mathcal{O}} = 2\Theta(v_1, v_1) = 2\sum_{i=1}^{h} c_{1i}^2 \Theta(u_i, u_i) = 2\sum_{i=1}^{h} c_{1i}^2 f_i, \tag{18}$$

so $c_i := 2c_{1i}^2 \geq 0$. We find $|\mathcal{O}^\times| = \sum_i c_i$ by comparing the coefficient of q on each side. $\square$

Remark 6.2. Even though the conclusion of Theorem 6.1 only involves Hecke operators T_n with n coprime to N, the full commutativity of $\mathfrak{B}(\mathcal{O})$ is still essential to the proof. If $u_1, \dots, u_h$ forms an orthogonal basis of eigenvectors of B_n for all n coprime to N, but are not eigenvectors for all B_n, then Eq. (16) may not necessarily hold.

We will apply Theorem 6.1 to maximal and submaximal orders $\mathcal{O}$. For this, it will be helpful to have a more refined statement, obtained by separating the Brandt eigenvectors into two types: those in the genus subspace $\mathfrak{M}_{\mathrm{G}}(\mathcal{O})$, and those orthogonal to the genus subspace.

Corollary 6.3. *Let $\mathcal{O}$ be a maximal or submaximal order in a definite quaternion algebra $\mathcal{A}$ over $\mathbb{Q}$, and $N := \operatorname{disc}(\mathcal{O})$. There exist $g \in \mathcal{M}_2(\Gamma_0(N))$ and $\mathfrak{B}(\mathcal{O})$-eigenvectors $u_1, \dots, u_m \in \mathfrak{M}(\mathcal{O})$ for some $m \geq 0$, satisfying the following properties.*

(*a*) *The modular form g satisfies*

$$g(q) = \begin{cases} d_1 \frac{\Theta(e,e)}{\langle e,e\rangle} + d_2 \frac{\Theta(e',e')}{\langle e',e'\rangle} & \text{if } [\mathfrak{O} : \mathcal{O}] = p \neq 2 \text{ and } p \mid \operatorname{disc}(\mathcal{A}), \\ d_1 \frac{\Theta(e,e)}{\langle e,e\rangle} & \text{otherwise,} \end{cases}$$

for some $d_1, d_2 \geq 0$, where e, e' are as defined in Eq. (9) *and* Eq. (10).

(*b*) *For $1 \leq i \leq m$, $f_i := \Theta(u_i, u_i)$ is a normalized $\mathbb{T}^{\mathrm{an}}$-eigenvector in $\mathcal{S}_2(\Gamma_0(N))$.*

(*c*) *There exist $c_1, \dots, c_m \geq 0$ with $c_1 + \cdots + c_m \leq |\mathcal{O}^\times|$ such that*

$$\theta_{\mathcal{O}}(q) - g(q) = c_1 f_1(q) + \cdots + c_m f_m(q). \tag{19}$$

Proof. We use Lemma 4.5 to obtain $N = \operatorname{nrd}(\mathcal{O}^\sharp)^{-1} = \operatorname{disc}(\mathcal{O})$. If $[\mathfrak{O} : \mathcal{O}]$ is not equal to an odd ramified prime of $\mathcal{A}$, then we have from Lemma 4.7 that $\mathfrak{M}(\mathcal{O})$ has only one (linearly independent) genus vector, namely e. In the other case, by Lemma 4.8, it has exactly two (linearly independent) genus vectors, namely e and e'.

Since e and e' are $\mathfrak{B}(\mathcal{O})$-eigenvectors (see Lemma 4.6 and Lemma 4.9), we can extend these to a Brandt eigenbasis for $\mathfrak{M}(\mathcal{O})$. Lemma 5.7 shows that every Brandt eigenbasis vector outside the genus subspace is cuspidal. We normalize the Brandt eigenbasis and apply Theorem 6.1; the desired result is obtained by re-labeling. $\square$

Remark 6.4. It is possible to have $m = 0$ in Corollary 6.3, for instance whenever $\mathcal{O}$ is a quaternion order of class number 1. In this case (*b*) is vacuous and (*c*) states that $\theta_{\mathcal{O}} = g$.

The form g should be thought of as the "Eisenstein component" of $\theta_{\mathcal{O}}$. We avoid this terminology as we have not shown in all cases that it belongs to the Eisenstein subspace in $\mathcal{M}_2(\Gamma_0(N))$. In particular, while $\Theta(e,e)$ (and $\Theta(e',e')$ when it exists) is a $\mathbb{T}^{\mathrm{an}}$-eigenvector, it is not necessarily an eigenvector for the full Hecke algebra $\mathbb{T}$. Thus it may have a nonzero cuspidal contribution, provided that the coefficient of q^n for this cusp form is zero for all n coprime to $\operatorname{disc}(\mathcal{O})$.

We compute the coefficients of g in Section 6.2, and bound the coefficients of each cusp form f_i in Section 6.3.

6.2. **Contribution from genus vectors.** As established in Corollary 6.3, the contribution of the genus subspace to the theta function $\theta_{\mathcal{O}}$ is captured entirely by the modular form g. For maximal orders and most submaximal orders, this contribution comes only from the standard genus vector e. The unique exception occurs for submaximal orders of index odd $p \mid \operatorname{disc}(\mathcal{A})$, which have a second genus vector e'.

Our goal in this section is to explicitly compute and bound the coefficients of g, treating the exceptional two-vector case separately in Section 6.2.3. Let $\phi(n) = n\prod_{\ell \mid n}\left(1-\frac{1}{\ell}\right)$ be Euler's totient function.

Proposition 6.5. *Let $\mathcal{O}$ be maximal or submaximal, $\ell \nmid \operatorname{disc}(\mathcal{A})$ prime, and $k \geq 1$. Let $g(q) = \sum g_n(\mathcal{O})q^n$ be as defined in Corollary 6.3. If $\mathcal{O}$ is maximal, then*

$$g_{\ell^k} := g_{\ell^k}(\mathcal{O}) = \frac{24}{\phi(\operatorname{disc}(\mathcal{A}))}\frac{\ell^{k+1}-1}{\ell-1}, \tag{20}$$

and if $\mathcal{O}$ is submaximal, then

$$0 \leq g_{\ell^k} \leq \frac{16}{\phi(\operatorname{disc}(\mathcal{A}))}\frac{\ell^{k+1}-1}{\ell-1}. \tag{21}$$

We prove this in three stages. We first compute d_1, which requires computing the eigenvalue of e under B_0. We then compute the eigenvalue of e under B_{ℓ^k}. Finally we consider the contribution from e' when it exists.

Note that $e = B_0 v_i$ for any standard basis vector v_i; it is a column eigenvector for B_n for all $n \geq 0$.

Remark 6.6. This description assumes we treat the Brandt module as column vectors with $v \mapsto B_n v$. We have $B_n v = v'$ if and only if $(v^t D)B_n = (v')^t D$, where recall D is the diagonal matrix with $D(i,i) = w_i$ as above. Thus the genus *row* vector is the constant vector $e^t D = \frac{1}{2}(1,1,\ldots,1)$, which up to scaling is the convention in [Voi21].

We now consider the coefficients of $\frac{\Theta(e,e)}{\langle e,e\rangle} = \sum_{n=0}^{\infty}\lambda_n(e)q^n$, where $\lambda_n(e)$ denotes the eigenvalue of e under the action of B_n. Since B_1 is the identity matrix, we have $\lambda_1(e) = 1$.

6.2.1. *Constant term.* A direct computation gives us $\lambda_0(e) = \sum \frac{1}{2w_i}$. We can evaluate this using the Eichler mass formula [Voi21, Main Theorem 25.3.19]:

$$\lambda_0(e) = \sum_{i=1}^{h}\frac{1}{2w_i} = \frac{\operatorname{disc}(\mathcal{O})}{24}\prod_{\ell \mid \operatorname{disc}(\mathcal{O})}\lambda(\mathcal{O},\ell),$$

where for $\ell \mid \operatorname{disc}(\mathcal{O})$ we have

$$\lambda(\mathcal{O},\ell) = \begin{cases} 1+\frac{1}{\ell} & \text{if } \mathcal{O}_\ell \text{ is residually split,}\\ 1-\frac{1}{\ell} & \text{if } \mathcal{O}_\ell \text{ is residually inert,}\\ 1-\frac{1}{\ell^2} & \text{if } \mathcal{O}_\ell \text{ is residually ramified}\end{cases}$$

(see [Voi21, Definition 24.3.2]). If $\mathcal{O}$ is maximal and $\ell \mid \operatorname{disc}(\mathcal{O})$ then $\mathcal{A}_\ell$ is a division algebra and $\mathcal{O}_\ell$ is residually inert. Now consider the classification of submaximal orders in Lemma 4.2, with r being a split prime of $\mathcal{A}$ and p a ramified prime of $\mathcal{A}$. If $\mathcal{O}$ is an Eichler order of index r then $\mathcal{O}_r$ is residually split; if $\mathcal{O}$ has index p then $\mathcal{O}_p$ is residually ramified; in every other case, $\ell \mid \operatorname{disc}(\mathcal{O})$ implies $\mathcal{O}_\ell$ is residually inert. We therefore have the following for any maximal or submaximal order $\mathcal{O} \subseteq \mathfrak{O}$:

$$\lambda_0(e) = \frac{\phi(\operatorname{disc}(\mathcal{A}))}{24}\cdot\begin{cases} 1 & \text{if } [\mathfrak{O}:\mathcal{O}] = 1,\\ r+1 & \text{if } [\mathfrak{O}:\mathcal{O}] = r,\\ p+1 & \text{if } [\mathfrak{O}:\mathcal{O}] = p,\\ r(r-1) & \text{if } [\mathfrak{O}:\mathcal{O}] = r^2,\\ p^2 & \text{if } [\mathfrak{O}:\mathcal{O}] = p^2.\end{cases} \tag{22}$$

6.2.2. *Nonconstant terms.* Now suppose $n > 1$ is coprime to $\operatorname{disc}(\mathcal{A})$. For each submaximal $\mathcal{O}$ we will determine $\lambda_n(e)$.

Lemma 6.7. *For $n \geq 1$,*

$$\lambda_n(e) = |\{J \subseteq_n I_j\}|, \tag{23}$$

and the quantity does not depend on j. Hence, $\lambda_{mn}(e) = \lambda_m(e)\lambda_n(e)$ provided that m, n are coprime, and the value of $\lambda_{\ell^k}(e)$ depends only on the isomorphism type of the completion $\mathcal{O}_\ell$.

Proof. Since we have $e^t B_n^t = \lambda_n(e)e^t$, we obtain $e^t B_n^t D = \lambda_n(e)(e^t D)$, where as before D is the diagonal matrix with $D(i,i) = w_i$, the matrix giving the inner product, and $e^t D = \frac{1}{2}(1, 1, \ldots, 1)$ is the genus row vector. Since B_n is self-adjoint by Lemma 2.9, we have that $DB_n = B_n^t D$, and so it follows that $(e^t D)B_n = \lambda_n(e)(e^t D)$. This implies that $\lambda_n(e)$ is equal to the sum of entries in a column of B_n, and so Eq. (23) follows from Lemma 3.3. The other assertions follow easily from Eq. (23) together with the local-global dictionary. $\square$

By the multiplicativity of $\lambda_n(e)$ in Lemma 6.7, it suffices to determine $\lambda_n(e)$ for prime powers. This is carried out in the following lemma for powers of a prime ℓ in the cases of maximal and submaximal orders.

Lemma 6.8. *Let $\mathcal{O}$ be maximal or submaximal, ℓ and r primes not dividing* $\operatorname{disc}(\mathcal{A})$*, and $k \geq 1$.*

(*a*) *If $\ell \nmid \operatorname{disc}(\mathcal{A})$ and $\mathcal{O}_\ell$ is maximal, then $\lambda_{\ell^k}(e) = \frac{\ell^{k+1}-1}{\ell-1}$.*
(*b*) *If $[\mathfrak{O} : \mathcal{O}] = r$, then $\lambda_{r^k}(e) = 2\frac{r^{k+1}-1}{r-1} - 1$.*
(*c*) *If $[\mathfrak{O} : \mathcal{O}] = r^2$, then $\lambda_{r^k}(e) = r^k - r$.*

Proof. Part (*a*) is classical. Indeed, since $\lambda_1(e) = 1$ and $\lambda_\ell(e) = \ell + 1$ for all primes $\ell \nmid \operatorname{disc}(\mathcal{A})$, using the relation $\lambda_{\ell^{k+1}}(e) = \lambda_{\ell^k}(e)\lambda_\ell(e) - \ell\lambda_{\ell^{k-1}}(e)$ (cf. Lemma 4.10), we can obtain by induction that $\lambda_{\ell^k}(e) = \frac{\ell^{k+1}-1}{\ell-1}$.

For (*b*): note that this case is the one studied in Section 4.3.2, and so Lemma 4.14 applies. Indeed, the proof of this lemma shows that there are $2r + 1$ right ideals of norm r in $\mathcal{O}$. This implies that $\lambda_r(e) = 2r + 1$ using Lemma 6.7. We also observe that $W_\mathfrak{r} e = e$ by definition of $W_\mathfrak{r}$ (see Section 4.3.2 for its definition). We obtain the relation $\lambda_{r^{k+1}}(e) = (r+1)\lambda_{r^k}(e) - r\lambda_{r^{k-1}}(e)$ by the recurrence relation in Proposition 4.13, and so the assertion follows.

For (*c*): by the definition of E, since e has strictly positive entries (in the standard basis), Ee also has strictly positive entries. Hence the Fourier expansion of $\Theta(Ee, Ee)$ has nonzero constant term, and Ee is not in the cuspidal subspace of $\mathfrak{M}(\mathfrak{O})$. Now let $\widetilde{e} \in \mathfrak{M}(\mathfrak{O})$ be the genus vector for $\mathfrak{O}$, and take $n \geq 1$ such that $\lambda_n(\widetilde{e})$ is distinct from $\lambda_n(u)$ for all cuspidal Brandt eigenvectors $u \in \mathfrak{M}(\mathfrak{O})$; such n exists, for example, by applying part (*a*) and taking k sufficiently large so that $\lambda_{\ell^k}(\widetilde{e})$ is larger than the Deligne bound. By Proposition 4.22, Ee is an eigenvector for $\widetilde{B}_n$. Since Ee is not in the cuspidal subspace, and $\widetilde{e}$ has distinct eigenvalues from all cuspidal eigenvectors, this forces Ee to be a multiple of $\widetilde{e}$. Hence $\lambda_{r^k}(Ee) = \frac{r^{k+1}-1}{r-1}$ by part (*a*), and applying Proposition 4.22 again we find that

$$\lambda_{r^k}(e) = (r^2 - r)\frac{r^{k-1} - 1}{r - 1} = r^k - r. \qquad \square$$

We have now determined $\lambda_n(e)$ for all n coprime to $\operatorname{disc}(\mathcal{A})$. As discussed above, certain cases admit a second genus vector e'. We compute the corresponding eigenvalues $\lambda_n(e')$ in the following subsection.

6.2.3. *Two-dimensional genus space.* Let $\mathcal{O}$ be submaximal of index $p > 2$ ramified in $\mathcal{A}$. This is the special case of two genus vectors e and e' as was mentioned above (see Lemma 4.8). The proof of [Piz80b, Theorem 5.34(b)] shows that $\lambda_n(e') = \left(\frac{n}{p}\right)\lambda_n(e)$ for all $n \geq 0$, where $\left(\frac{n}{p}\right)$ is the Legendre symbol. In particular, $\lambda_0(e') = 0$. We show that $\Theta(e,e)$ and $\Theta(e',e')$ contribute equally to $\theta_{\mathcal{O}}$. More precisely, $d_1 = d_2$ and $\langle e,e\rangle = \langle e',e'\rangle$ in $g(q) = d_1\frac{\Theta(e,e)}{\langle e,e\rangle} + d_2\frac{\Theta(e',e')}{\langle e',e'\rangle}$ (see Corollary 6.3).

Lemma 6.9. *The coefficients of the normalized non-cusp forms in $\theta_{\mathcal{O}}$ are equal: $d_1 = d_2$.*

Proof. Let v_1 be the first vector in the standard basis of $\mathfrak{M}(\mathcal{O})$, ordered so that $I_1 = \mathcal{O}$. For a Brandt eigenbasis $u_1 = e, u_2 = e', u_3, \dots, u_h$, we write

$$v_1 = a_1 e + a_2 e' + \sum_{i=3}^{h} a_i u_i,$$

and we have $a_1 = \frac{\langle v_1,e\rangle}{\langle e,e\rangle}$ and $a_2 = \frac{\langle v_1,e'\rangle}{\langle e',e'\rangle}$ by orthogonality. As in Eq. (18), we write using the decomposition of the theta function,

$$\theta_{\mathcal{O}} = 2\Theta(v_1,v_1) = 2a_1^2\Theta(e,e) + 2a_2^2\Theta(e',e') + \cdots + 2a_h^2\Theta(u_h,u_h),$$

and so

$$d_1 = 2a_1^2\langle e,e\rangle = \frac{2\langle v_1,e\rangle^2}{\langle e,e\rangle} \quad \text{and} \quad d_2 = 2a_2^2\langle e',e'\rangle = \frac{2\langle v_1,e'\rangle^2}{\langle e',e'\rangle}.$$

By direct computation of the inner products, the negative signs in e' are squared, implying that $\langle e,e\rangle = \langle e',e'\rangle$. Furthermore, since $\langle v_1,e\rangle = v_1^t De$, where D is the diagonal matrix with $D(i,i) = w_i$, we obtain that

$$\langle v_1,e\rangle = w_1\left(\frac{1}{2w_1}\right) = \frac{1}{2}.$$

Similarly, $\langle v_1,e'\rangle = w_1\left(\frac{1}{2w_1}\right) = \frac{1}{2}$. We now see that $\langle e,e\rangle = \langle e',e'\rangle$ and $\langle v_1,e\rangle = \langle v_1,e'\rangle$. Therefore, $d_1 = d_2$, which proves the assertion. $\square$

6.2.4. *Proof of* Proposition 6.5. We now compute the non-cuspidal part's contribution for maximal orders, and bound the contribution for all submaximal orders. Recall first that the q-expansions of the cusp forms have zero constant term. Recall also that there is only one case, namely $[\mathfrak{O} : \mathcal{O}] = p \neq 2$, for which the genus space is 2-dimensional.

First suppose $\mathcal{O}$ is maximal. Comparing constant terms in Eq. (19), we have $1 = d_1\lambda_0(e)$, so that $d_1 = \frac{24}{\phi(\mathrm{disc}(\mathcal{A}))}$ by Eq. (22). Then $g_{\ell^k} = d_1\lambda_{\ell^k}(e)$ is determined by Lemma 6.8(a), and Eq. (20) holds as desired.

Suppose now that $\mathcal{O}$ is submaximal, so that it is one of the four types of orders listed in Lemma 4.2. We first consider the case that ℓ divides $\mathrm{disc}(\mathcal{O})$. Since $\ell \nmid \mathrm{disc}(\mathcal{A})$, this implies $\ell = r$ is a split prime of $\mathcal{A}$ and either $[\mathfrak{O} : \mathcal{O}] = r$ or r^2. By Lemma 6.8 we therefore have $\lambda_{r^k}(e) = 2\frac{r^{k+1}-1}{r-1} - 1$ or $\lambda_{r^k}(e) = r^k - r$, respectively. Observing that

$$\frac{r^k - r}{r(r-1)} \leq \frac{2\frac{r^{k+1}-1}{r-1} - 1}{r+1} \leq \frac{2}{r+1}\cdot\frac{r^{k+1}-1}{r-1}$$

for all primes r and all $k \geq 0$, we can conclude in either case (using the corresponding value of $\lambda_0(e)$ from Eq. (22)) that

$$g_{r^k} = \frac{\lambda_{r^k}(e)}{\lambda_0(e)} \leq \frac{24}{\phi(\mathrm{disc}(\mathcal{A}))}\cdot\frac{2}{r+1}\cdot\frac{r^{k+1}-1}{r-1},$$

giving the desired bound in Eq. (21) since $r \geq 2$.

Now suppose $\ell \nmid \mathrm{disc}(\mathcal{O})$, so $\mathcal{O}_\ell$ is maximal. We first consider the case that $[\mathfrak{O} : \mathcal{O}]$ is not an odd ramified prime of $\mathcal{A}$, so $\mathfrak{M}_{\mathrm{G}}(\mathcal{O})$ is one-dimensional. From Eq. (22) we see that $\lambda_0(e) \geq \frac{\phi(\mathrm{disc}(\mathcal{A}))}{24} \cdot 2$ for any submaximal order $\mathcal{O}$. It follows from Lemma 6.8(a) that

$$g_{\ell^k} = \frac{\lambda_{\ell^k}(e)}{\lambda_0(e)} \leq \frac{12}{\phi(\mathrm{disc}(\mathcal{A}))} \cdot \frac{\ell^{k+1}-1}{\ell-1},$$

which is strictly smaller than the desired bound in Eq. (21).

Finally suppose $[\mathfrak{O} : \mathcal{O}] = p \neq 2$ is ramified in $\mathcal{A}$, so that $\mathfrak{M}_{\mathrm{G}}(\mathcal{O})$ is two-dimensional. By Lemma 6.9, we have that $g(q) = d_1\left(\frac{\Theta(e,e)}{\langle e,e\rangle} + \frac{\Theta(e',e')}{\langle e',e'\rangle}\right)$, and so comparing constant terms in Eq. (19), we again have $1 = d_1\lambda_0(e)$. Recall that $\lambda_n(e') = \left(\frac{n}{p}\right)\lambda_n(e)$. Thus

$$g_{\ell^k} \;=\; \frac{\lambda_{\ell^k}(e) + \lambda_{\ell^k}(e')}{\lambda_0(e)} \;=\; \frac{\lambda_{\ell^k}(e)}{\lambda_0(e)}\left(1 + \left(\frac{\ell^k}{p}\right)\right) \leq \frac{2\lambda_{\ell^k}(e)}{\lambda_0(e)}.$$

Using Eq. (22) to compute $\lambda_0(e) = \frac{\phi(\mathrm{disc}(\mathcal{A}))}{24}(p+1)$ and Lemma 6.8(a) to compute $\lambda_{\ell^k}(e)$, we conclude

$$g_{\ell^k} \leq \frac{2}{p+1} \cdot \frac{24}{\phi(\mathrm{disc}(\mathcal{A}))} \cdot \frac{\ell^{k+1}-1}{\ell-1}.$$

This implies Eq. (21), because $p \geq 3$. Since we have seen that the desired bounds in Proposition 6.5 hold in all cases, the proof is complete.

Remark 6.10. If we assume $\ell \geq 3$, then the same proof shows that

$$g_{\ell^k} \leq \frac{12}{\phi(\mathrm{disc}(\mathcal{A}))} \cdot \frac{\ell^{k+1}-1}{\ell-1}$$

for any submaximal order $\mathcal{O}$. If $2 \nmid \mathrm{disc}(\mathcal{A})$ then this stronger bound is sharp: for any prime $\ell \geq 3$ and $k \geq 0$, equality is obtained for submaximal orders of index 4.

6.3. **Cuspidal contribution.** Now suppose $u \in \mathfrak{M}(\mathcal{O})$ is a simultaneous eigenvector of $\mathfrak{B}(\mathcal{O})$ (and also of $W_{\mathfrak{r}}$ if $\mathcal{O}$ has index r), normalized so that $\langle u,u\rangle = 1$. Assume that $\Theta(u,u) \in \mathcal{S}_2(\Gamma_0(N))$ is a cusp form, where the level is $N = \mathrm{disc}(\mathcal{O})$ (Lemma 4.5).

Following the proof of Lemma 5.6, $\Theta(u,u)$ is a $\mathbb{T}^{\mathrm{an}}$-eigenvector. While it does *not* generally belong to the new subspace $\mathcal{S}_2^{\mathrm{new}}(\Gamma_0(N))$, the Deligne bound (Lemma 5.3) still applies to its n-th Fourier coefficient $\lambda_n(u)$ whenever n is coprime to N.

The complication arises when our prime ℓ divides the level. Indeed, since $N = [\mathfrak{O} : \mathcal{O}]\,\mathrm{disc}(\mathcal{A})$ and the prime ℓ may equal the index r, there are situations where ℓ divides N, preventing the direct use of the Deligne bound. To bypass this, we establish slightly weaker bounds on $\lambda_{r^k}(u)$ by relating $\Theta(u,u)$ to suitable newforms, and applying the Deligne bound to the coefficients of those newforms instead.

Proposition 6.11. *Let $\mathcal{O}$ be maximal or submaximal, and let $u \in \mathfrak{M}(\mathcal{O})$ be a cuspidal Brandt eigenvector. Let $\ell \nmid \mathrm{disc}(\mathcal{A})$ be prime. If $\mathcal{O}$ has index r in a maximal order, assume further that u is an eigenvector of $W_{\mathfrak{r}}$. Then for all $k \geq 1$,*

$$|\lambda_{\ell^k}(u)| \leq 2k\ell^{(k+2)/2}.$$

As in Section 4.3, we break into cases depending on the power of ℓ dividing $\mathrm{disc}(\mathcal{O})$. That is, either $\mathcal{O}_\ell$ is maximal, $\mathcal{O}$ is submaximal of index r and $\ell = r$, or $\mathcal{O}$ is submaximal of index r^2 and $\ell = r$. In each case we relate the coefficients of $\Theta(u,u)$ to the coefficients of a newform f of appropriate level.

Lemma 6.12. *Let $\mathcal{O}$ be a maximal or submaximal order, $\ell \nmid \mathrm{disc}(\mathcal{O})$ a prime, and let $u \in \mathfrak{M}(\mathcal{O})$ be a cuspidal Brandt eigenvector. Then there exists a newform*

$$f(q) = \sum b_n q^n \in \mathcal{S}_2(\Gamma_0(M))$$

for some $M \mid \mathrm{disc}(\mathcal{O})$ such that $\lambda_{\ell^k}(u) = b_{\ell^k}$ for all $k \geq 0$.

Proof. The condition $\ell \nmid \operatorname{disc}(\mathcal{O})$ implies $\mathcal{A}$ is split at ℓ and $\mathcal{O}_\ell$ is maximal. By Lemma 5.4, there is a newform f of some level $M \mid N$, $N := \operatorname{disc}(\mathcal{O})$, with

$$\Theta(u,u)(q) = \sum_{d \mid \frac{N}{M}} c_d f(q^d)$$

for some coefficients $c_d \in \mathbb{C}$, with $c_1 = 1$ by comparing the coefficient of q on each side. Since $\ell \nmid \operatorname{disc}(\mathcal{O})$ by assumption, every nontrivial divisor of $\frac{N}{M}$ has a prime factor distinct from ℓ. So only the $d = 1$ term contributes to the coefficient of q^{ℓ^k}. □

Remark 6.13. By a result of Ponomarev, we can strengthen Lemma 6.12 if $\mathcal{O}$ is maximal: for any cuspidal Brandt eigenvector $u \in \mathfrak{M}(\mathcal{O})$, $\Theta(u,u)$ is a newform in $\mathcal{S}_2(\Gamma_0(\operatorname{disc}(\mathcal{A})))$ [Pon09, Theorem 4].

Lemma 6.14. *Let $\mathcal{O}$ be submaximal of index r, and $u \in \mathfrak{M}(\mathcal{O})$ a cuspidal Brandt eigenvector which is also an eigenvector for $W_{\mathfrak{r}}$. Then either $|\lambda_{r^k}(u)| = 1$ for all $k \geq 0$, or there exists a newform*

$$f(q) = \sum b_n q^n \in \mathcal{S}_2(\Gamma_0(\operatorname{disc}(\mathcal{A})))$$

and $\delta \in \{-1, 1\}$ such that

$$\Theta(u,u)(q) = f(q) + r\delta f(q^r).$$

Proof. Let $N = \operatorname{disc}(\mathcal{O})$. We write $a_n := \lambda_n(u)$ for convenience. By Lemma 5.6 and Lemma 5.4, there is a newform $f \in \mathcal{S}_2(\Gamma_0(M))$ for some $M \mid N$ such that

$$\Theta(u,u)(q) = \sum_{d \mid \frac{N}{M}} c_d f(q^d)$$

for some coefficients $c_d \in \mathbb{C}$. Write $f(q) = \sum_{n=1}^{\infty} b_n q^n$ with $b_1 = 1$. Comparing coefficients on both sides,

$$a_n = \sum_{d \mid \gcd\left(n, \frac{N}{M}\right)} c_d b_{n/d}.$$

Observe that N is squarefree. We first prove that $\operatorname{disc}(\mathcal{A})$ must divide M. If not, suppose p divides $\operatorname{disc}(\mathcal{A})$ but not M. Then since the level of f is coprime to p, its coefficients obey the standard Hecke relations $b_{p^{k+1}} = b_{p^k} b_p - p b_{p^{k-1}}$. On the other hand we have $a_p = \pm 1$ since $B_p^2 = 1$ (Lemma 4.11). Solving the equations $a_p = b_p + c_p$ and

$$1 = a_{p^2} = b_{p^2} + c_p b_p = b_p^2 - p + c_p b_p,$$

we find $|b_p| = p + 1$, which contradicts the Deligne bound. Hence we have either $N = M$ or $N = rM$.

First consider the case $N = M$, so $\Theta(u,u)$ is a newform of level $\operatorname{disc}(\mathcal{O})$. Then $a_{r^k} = b_{r^k}$ for all $k \geq 0$. Since r divides the level of the newform f, the coefficients satisfy $a_{r^k} = a_r^k$ for all $k \geq 0$ [AL70, Theorem 3(ii)′]. Applying the recurrence relation of Proposition 4.13 to a_{r^k} we find

$$a_r^k = a_r^{k-1}(a_r - r\delta) - r a_r^{k-2},$$

where δ is the eigenvalue of u under the involution $W_{\mathfrak{r}}$. Taking $k = 2$ we find $\delta a_r = -1$, so that $|a_r| = 1$. Hence $|a_{r^k}| = 1$ for all k.

Now consider the case $N = rM$, so $\Theta(u,u)(q) = f(q) + cf(q^r)$ for some $f \in \mathcal{S}_2(\Gamma_0(M))$ and $c \in \mathbb{C}$. This results in the equation

$$a_{r^k} = b_{r^k} + c b_{r^{k-1}}$$

for all $k \geq 0$ (using the convention that $b_x = 0$ if x is not a positive integer). Now since r is coprime to the level of f, the standard Hecke operator recurrence relation applies to the coefficients b_{r^k}, while the a_{r^k} coefficients are instead subject to the recurrence relation of Proposition 4.13.

We took our eigenbasis so that u is an eigenvector for $W_{\mathfrak{r}}$ as well, so as a consequence, letting $\delta = \pm 1$ be the eigenvalue of u under $W_{\mathfrak{r}}$, we have for all $k \geq 2$,

$$\begin{aligned} b_{r^k} + c b_{r^{k-1}} &= (b_{r^{k-1}} b_r - r b_{r^{k-2}}) + c(b_{r^{k-2}} b_r - r b_{r^{k-3}}), \\ a_{r^k} &= a_{r^{k-1}}(a_r - r\delta) - r a_{r^{k-2}} \\ &= (b_{r^{k-1}} + c b_{r^{k-2}})(b_r + c - r\delta) - r(b_{r^{k-2}} + c b_{r^{k-3}}), \end{aligned}$$

and the equality of these expressions simplifies to

$$(c - r\delta)(b_{r^{k-1}} + c b_{r^{k-2}}) = 0.$$

If $b_{r^{k-1}} + c b_{r^{k-2}} = 0$ for all $k \geq 2$ then $b_r = -c$ and $b_{r^2} = c^2$, contradicting $b_{r^2} = b_r^2 - r$. Thus $c = r\delta$, so we can write

$$\Theta(u,u)(q) = f(q) + r\delta f(q^r). \qquad \square$$

Lemma 6.15. *Let $\mathcal{O}$ be submaximal of index r^2, and let $u \in \mathfrak{M}(\mathcal{O})$ be a cuspidal Brandt eigenvector. Then either $\lambda_{r^k}(u) = 0$ for all $k \geq 1$, or there exists a newform*

$$f(q) = \sum b_n q^n \in \mathcal{S}_2(\Gamma_0(\mathrm{disc}(\mathcal{A})))$$

such that $\lambda_r(u) = 0$ and $\lambda_{r^k}(u) = (r^2 - r) b_{r^{k-2}}$ for all $k \geq 2$.

Proof. If $\lambda_{r^k}(u) = 0$ for all $k \geq 1$ then we are done, so suppose this is not the case. Then by Proposition 4.22, we have $\lambda_r(u) = 0$, and there exists $w \in \mathfrak{M}(\mathfrak{O})$, an eigenvector for $\widetilde{B}_n \in \mathfrak{B}(\mathfrak{O})$ for all $n \geq 1$, satisfying $\lambda_{r^k}(u) = (r^2 - r)\lambda_{r^{k-2}}(w)$ for $k \geq 2$. Now since u is cuspidal we have $B_0 u = 0$, so the sum of entries of u is 0. By definition of E, the sum of entries of Eu equals the sum of entries of u, so we can conclude $\widetilde{B}_0(Eu) = 0$. Thus Eu is in fact a cuspidal eigenvector for the full Brant algebra $\mathfrak{B}(\mathfrak{O})$. Therefore there is a newform $f(q) = \sum b_n q^n$ such that $\lambda_{r^k}(w) = b_{r^k}$ by Lemma 6.12. Hence $\lambda_{r^k}(u) = (r^2 - r) b_{r^{k-2}}$ for all $k \geq 2$ as desired. $\square$

Proof of Proposition 6.11. If $\mathcal{O}_\ell$ is maximal then there is a newform $f(q) = \sum b_n q^n$ such that $\lambda_{\ell^k}(u) = b_{\ell^k}$ by Lemma 6.12, so $|\lambda_{\ell^k}(u)| \leq (k+1)\ell^{k/2}$ by the Deligne bound.

If $\mathcal{O}$ is submaximal of index r and $\ell = r$, then either $|\lambda_{r^k}(u)| = 1$ for all k, or there is a newform $f(q) = \sum b_n q^n$ such that $\lambda_{r^k}(u) = b_{r^k} \pm r b_{r^{k-1}}$ by Lemma 6.14, so

$$\begin{aligned} |\lambda_{r^k}(u)| &\leq |b_{r^k}| + r|b_{r^{k-1}}| \\ &\leq (k+1) r^{k/2} + r k r^{(k-1)/2} \\ &= (k + 1 + k\sqrt{r}) r^{k/2}. \end{aligned}$$

If $\mathcal{O}$ is submaximal of index r^2 and $\ell = r$, then either $\lambda_{r^k}(u) = 0$ for all k, or there is a newform $f(q) = \sum b_n q^n$ such that $\lambda_r(u) = 0$ and $\lambda_{r^k}(u) = (r^2 - r) b_{r^{k-2}}$ by Lemma 6.15. In the latter case, with the Deligne bound $|b_{r^{k-2}}| \leq (k-1) r^{(k-2)/2}$, we therefore have

$$|\lambda_{r^k}(u)| \leq (r^2 - r)(k-1) r^{(k-2)/2} \leq k r^{(k+2)/2}$$

for $k \geq 2$ (and the bound for $k = 1$ is immediate). $\square$

7. Counting elements of prescribed norm

Let $\mathcal{A}$ be a definite quaternion algebra over $\mathbb{Q}$. For an order $\mathcal{O} \subset \mathcal{A}$, we let $a_n(\mathcal{O})$ denote the number of elements of norm n in $\mathcal{O}$, so that $\theta_{\mathcal{O}}(q) = \sum_{n=0}^{\infty} a_n(\mathcal{O}) q^n$.

Proposition 7.1. *Let $\mathcal{O} \subset \mathcal{A}$ be an order, $\ell \nmid \mathrm{disc}(\mathcal{A})$ a prime, and $k \geq 1$ be an integer. If $\mathcal{O}$ is maximal then*

$$a_{\ell^k}(\mathcal{O}) \geq \frac{24}{\phi(\mathrm{disc}(\mathcal{A}))} \frac{\ell^{k+1} - 1}{\ell - 1} - 12k\ell^{(k+2)/2},$$

and if $\mathcal{O}$ is not maximal then

$$a_{\ell^k}(\mathcal{O}) \le \frac{16}{\phi(\operatorname{disc}(\mathcal{A}))}\frac{\ell^{k+1}-1}{\ell-1} + 12k\ell^{(k+2)/2}.$$

Proof. Once we prove the upper bound on $a_{\ell^k}(\mathcal{O})$ for submaximal orders, the same upper bound for all non-maximal orders will follow, because any non-maximal order $\mathcal{O}$ is contained in a submaximal order $\mathcal{O}'$ and hence $a_{\ell^k}(\mathcal{O}) \le a_{\ell^k}(\mathcal{O}')$. We therefore assume, without loss of generality, that $\mathcal{O}$ is either maximal or submaximal.

We first consider the case that $|\mathcal{O}^\times| > 6$. Up to isomorphism, there are precisely three such quaternion orders [Voi21, Theorem 11.5.14]: these are the quaternion orders with discriminants 2 (the Hurwitz order), 3, and 4 (the Lipschitz order). The lattice structures of these orders are well studied (they are D_4, the orthogonal direct sum $A_2 \oplus A_2$, and $\mathbb{Z}^4$, respectively), and in particular there are closed formulas for their theta functions: for any quaternion order of discriminant N with $(N-1) \mid 12$, one has

$$|\mathcal{O}^\times| = \frac{24}{N-1} \qquad \text{and} \qquad \theta_{\mathcal{O}}(q) = 1 + |\mathcal{O}^\times| \sum_{m=1}^{\infty}\left(\sum_{N\nmid d\mid m} d\right) q^m$$

(for instance, using [Voi21, Chapter 26]), so in particular we have

$$a_{\ell^k}(\mathcal{O}) = \begin{cases} \dfrac{24}{\phi(\operatorname{disc}(\mathcal{A}))}\dfrac{\ell^{k+1}-1}{\ell-1}, & \operatorname{disc}\mathcal{O} = 2 \text{ or } 3 \text{ (maximal)} \\ \dfrac{8}{\phi(\operatorname{disc}(\mathcal{A}))}\dfrac{\ell^{k+1}-1}{\ell-1}, & \operatorname{disc}\mathcal{O} = 4 \text{ (submaximal).} \end{cases}$$

In all these cases the desired inequalities hold.

From now on we assume $|\mathcal{O}^\times| \le 6$. Recall that in the case of a submaximal $\mathcal{O}$ of index r in a maximal order, we assume that each u_i is an eigenvector of the operator $W_{\mathfrak{r}}$ defined by Eq. (11) (this is possible by Lemma 4.12). First we use Eq. (19) to write

$$\theta_{\mathcal{O}}(q) - g(q) = c_1 f_1(q) + \cdots + c_m f_m(q)$$

with $g(q)$ the non-cuspidal part, $c_1, \ldots, c_m$ non-negative real numbers, $c_1 + \cdots + c_m \le |\mathcal{O}^\times| \le 6$, and $f_1, \ldots, f_m$ normalized cusp forms that are $\mathbb{T}^{\mathrm{an}}$-eigenvectors. Write $f_i(q) = \sum a_{i,n} q^n$ for each i, and $g(q) = \sum g_n(\mathcal{O}) q^n$. By Proposition 6.11, we have $|a_{i,\ell^k}| \le 2k\ell^{(k+2)/2}$ for all i and all $k \ge 1$. So comparing the coefficient of q^{ℓ^k} on both sides, we obtain

$$\begin{aligned} \left|a_{\ell^k}(\mathcal{O}) - g_{\ell^k}(\mathcal{O})\right| &\le c_1|a_{1,\ell^k}| + \cdots + c_m|a_{m,\ell^k}| \\ &\le (c_1 + \cdots + c_m) 2k\ell^{(k+2)/2} \\ &\le 12k\ell^{(k+2)/2}, \end{aligned}$$

and therefore

$$g_{\ell^k}(\mathcal{O}) - 12k\ell^{(k+2)/2} \le a_{\ell^k}(\mathcal{O}) \le g_{\ell^k}(\mathcal{O}) + 12k\ell^{(k+2)/2}.$$

We now apply Proposition 6.5 to obtain the exact value of $g_{\ell^k}(\mathcal{O})$ if $\mathcal{O}$ is maximal, and an upper bound on $g_{\ell^k}(\mathcal{O})$ if $\mathcal{O}$ is submaximal. $\square$

Theorem 7.2. *Let $\mathfrak{O}$ be a maximal order in a definite quaternion algebra $\mathcal{A}$ over $\mathbb{Q}$, and let $\ell \nmid \operatorname{disc}(\mathcal{A})$ be a prime. For any $k \ge 1$ satisfying*

$$\ell^{(k-2)/2} > 3\phi(\operatorname{disc}(\mathcal{A}))k, \tag{24}$$

the elements of norm ℓ^k in $\mathfrak{O}$ generate $\mathfrak{O}$ as a $\mathbb{Z}$-algebra.

Proof. Let $\mathcal{O}$ be any proper suborder of $\mathfrak{O}$, and apply Proposition 7.1:

$$\begin{aligned} a_{\ell^k}(\mathfrak{O}) - a_{\ell^k}(\mathcal{O}) &\geq \frac{8}{\phi(\mathrm{disc}(\mathcal{A}))}\frac{\ell^{k+1}-1}{\ell-1} - 24k\ell^{(k+2)/2} \\ &\geq \frac{8}{\phi(\mathrm{disc}(\mathcal{A}))}\ell^k - 24k\ell^{(k+2)/2}, \end{aligned}$$

which is strictly positive by Eq. (24). This proves that the elements of norm ℓ^k in $\mathfrak{O}$ do not lie in any proper suborder of $\mathfrak{O}$, so the $\mathbb{Z}$-algebra they generate must be the entirety of $\mathfrak{O}$. □

If we take $\mathcal{A} = \mathcal{A}_p$ to be ramified at exactly p and ∞, then we have $\phi(\mathrm{disc}(\mathcal{A})) = p-1$, so Theorem 1.1 follows.

8. Bounding the number of generators

Suppose we wish to find a set of generators for a maximal order $\mathfrak{O}$ with norms constrained to lie in some set S of positive integers. As a running example we may take $S = \{\ell^k : k \geq 0\}$ for some prime $\ell \nmid \mathrm{disc}(\mathcal{A})$. Theorem 7.2 tells us that we can replace the set of allowable norms S with the singleton $\{\ell^k\}$ for any sufficiently large value of k. On the other hand, the corresponding set of generators for $\mathfrak{O}$ may a priori be quite large.

Proposition 8.1. *Let $\mathfrak{O}$ be a maximal order in a definite quaternion algebra $\mathcal{A}$ over $\mathbb{Q}$. Suppose that $\mathfrak{O}$ is generated as a $\mathbb{Z}$-algebra by a finite set C_0, and let n be the maximum norm attained by any element of C_0. Then $\mathfrak{O} = \mathbb{Z}[C]$ for a subset $C \subseteq C_0$ with*

$$|C| \leq 2\log_2(n) - \log_2(\mathrm{disc}(\mathcal{A})) + 4.$$

Proof. Since $\mathfrak{O} = \mathbb{Z}[C_0]$, there must exist non-rational $\alpha, \beta \in C_0$ that do not lie in the same subfield of $\mathcal{A}$. Let $\Lambda_0 = \mathbb{Z}[\alpha, \beta]$. Since $\alpha, \beta \in \mathfrak{O}$, one can check that $\Lambda_0 = \mathrm{Span}_{\mathbb{Z}}(1, \alpha, \beta, \alpha\beta)$, a full rank lattice. Given $\{b_1, b_2, b_3, b_4\} = \{1, \alpha, \beta, \alpha\beta\}$, a $\mathbb{Z}$-basis of Λ_0, recall that the reduced discriminant of Λ_0 is the square root of the determinant of the Gram matrix $\langle b_i, b_j\rangle = \mathrm{trd}(b_i\bar{b_j})$. Since the trace form is positive definite, Hadamard's inequality gives

$$\det\big(\langle b_i, b_j\rangle\big) \leq \prod_{i=1}^{4}\langle b_i, b_i\rangle,$$

where

$$\prod_{i=1}^{4}\langle b_i, b_i\rangle = \prod_{i=1}^{4}\mathrm{trd}(b_i\bar{b}_i) = \prod_{i=1}^{4}2\mathrm{nrd}(b_i) = 2^4\mathrm{nrd}^2(\alpha)\mathrm{nrd}^2(\beta) \leq (2n)^4.$$

It follows that

$$\mathrm{disc}(\Lambda_0) = \sqrt{\det\big(\langle b_i, b_j\rangle\big)} \leq 4n^2.$$

Therefore,

$$[\mathfrak{O} : \Lambda_0] = \frac{\mathrm{disc}(\Lambda_0)}{\mathrm{disc}(\mathfrak{O})} = \frac{\mathrm{disc}(\Lambda_0)}{\mathrm{disc}(\mathcal{A})} \leq \frac{4n^2}{\mathrm{disc}(\mathcal{A})}.$$

Now for $i \geq 0$, if $\Lambda_i \neq \mathfrak{O}$, it is possible to pick $\gamma_{i+1} \in C_0 \smallsetminus \Lambda_i$ and let $\Lambda_{i+1} := \mathbb{Z}[\alpha, \beta, \gamma_1, \ldots, \gamma_{i+1}]$. Then Λ_i is strictly contained in Λ_{i+1}. Once we reach a value of m with $\Lambda_m = \mathfrak{O}$, we take $C = \{\alpha, \beta, \gamma_1, \ldots, \gamma_m\}$. Since

$$[\mathfrak{O} : \Lambda_{i+1}] \leq \frac{1}{2}[\mathfrak{O} : \Lambda_i] \qquad \text{for } i = 0, \ldots, m-1 \text{ and} \qquad [\mathfrak{O} : \Lambda_m] = 1,$$

we must have

$$m \leq \log_2([\mathfrak{O} : \Lambda_0]) \leq \log_2(4n^2/\,\mathrm{disc}(\mathcal{A})).$$

□

As an example, let p, ℓ be primes with $p > \frac{1}{3}\ell^{30} + 1$. If $k = \lfloor \frac{5}{2} \log_\ell(3p-3) \rfloor + 1$, so we are in the conditions of Theorem 1.1, then $\mathfrak{O}$ is generated by the set of elements of norm ℓ^k. So we may conclude that there is a set of at most

$$5 \log_\ell(3p-3) \log_2(\ell) + 2\log_2(\ell) - \log_2(p) + 4 \leq \log_2(p^4\ell^2) + 12$$

elements of norm ℓ^k that generates $\mathfrak{O}$. While the number of elements of this generating set is polynomial in the bit length of p, it should be possible to find an even smaller generating set. The following result tells us that we can indeed find a generating set with a very small number of elements, provided that we remove the constraint that the norms must be small.

Theorem 8.2. *Let $\mathfrak{O}$ be a maximal order in a definite quaternion algebra $\mathcal{A}$ over $\mathbb{Q}$. Let S be any infinite set consisting of positive integers coprime to* $\operatorname{disc}(\mathfrak{O})$. *Then $\mathfrak{O} = \mathbb{Z}[C]$ for a subset $C \subseteq \mathfrak{O}$ with $|C| \leq 3$ and $\operatorname{nrd}(x) \in S$ for all $x \in C$.*

More precisely, we will show that there exist elements $\alpha, \beta, \gamma \in \mathfrak{O}$, each with norm in S, such that $\mathfrak{O} = \operatorname{Span}_{\mathbb{Z}}(1, \alpha, \beta, \gamma, \alpha\beta, \alpha\gamma, \beta\gamma)$.

Remark 8.3. Here are a few remarks on the optimality of Theorem 8.2. First note that this result does not hold for arbitrary quaternion orders, even if we exclude the bound on $|C|$. For instance, if $\mathcal{O}$ is an Eichler order of index 2 in a maximal order, then $\mathcal{O}$ is not generated by any number of elements of odd norm [GL25b, 2.3.4].

We also note that Theorem 8.2 does not hold under the constraint $|C| \leq 2$, even if we remove all constraints on the norms of elements of C. Indeed, if $\mathfrak{O}$ is generated as a $\mathbb{Z}$-algebra by two elements $a_1, a_2 \in \mathfrak{O}$, then by [Voi21, Definition 15.4.4] we have

$$\operatorname{disc}(\mathcal{A}) = \operatorname{disc}(\mathfrak{O}) = |\operatorname{trd}((a_1a_2 - a_2a_1)\overline{a_1a_2})| = |\operatorname{nrd}(a_1a_2 - a_2a_1)|,$$

so that $\mathfrak{O}$ contains an element $a_1a_2 - a_2a_1$ of trace 0 and norm $\operatorname{disc}(\mathcal{A})$. Not every maximal order contains such an element: for instance, if $\operatorname{disc}(\mathcal{A}) = p$, then $\mathfrak{O}$ has an element of trace 0 and norm p if and only if it is isomorphic to the endomorphism ring of an elliptic curve over $\mathbb{F}_p$. One can use [HKTV25, Proposition 3.1] to show a partial converse: if $\mathfrak{O}$ is isomorphic to the endomorphism ring of an elliptic curve over $\mathbb{F}_p$, then it is generated as a $\mathbb{Z}$-algebra by two elements.

The key ingredient in the proof of Theorem 8.2 is the notion of theta functions of lattice cosets: given a lattice $\Lambda \subseteq \mathcal{A}$, an element $t \in \Lambda$, and a positive integer m, we may define the theta function of $t + m\Lambda$. If $t \notin m\Lambda$ then this is an *inhomogeneous theta function.* We use the following local-global principle for representability of integers by inhomogeneous theta functions, which is a variant of [GL25a, Theorem 6.6.1].

Theorem 8.4. *Let $\mathcal{A}$ be a definite quaternion algebra over $\mathbb{Q}$ and $\mathcal{O} \subseteq \mathcal{A}$ an order. Let $d \geq 2$ be a squarefree positive integer, and $t \in \mathcal{O}$ with $t \notin \ell\mathcal{O}$ for all primes $\ell \mid d$. For any sufficiently large positive integer μ coprime to* $\operatorname{disc}(\mathcal{O})$, *if there exists $x_\ell \in t + (d\mathcal{O} \otimes \mathbb{Z}_\ell)$ with $\operatorname{nrd}(x_\ell) = \mu$ for all primes ℓ, then there exists $x \in t + d\mathcal{O}$ with $\operatorname{nrd}(x) = \mu$.*

In fact Theorem 8.4 still holds even if we drop the squarefree constraint on d, but the necessary modifications to the proof take some extra work which is not needed for our applications here.

Proof. If d is prime then this is a special case of [GL25a, Theorem 6.6.1], with $L = \mathbb{Q}$, $n = 4$, $\Lambda = \mathcal{O}$, $\mathfrak{p} = d\mathbb{Z}$, and $M = t + d\mathcal{O}$. We indicate here how the proof changes for arbitrary squarefree d. A rough sketch of the proof is to decompose the theta function of M into a cuspidal part and a non-cuspidal part, and, as in Proposition 6.5, to show that the non-cuspidal part grows sufficiently rapidly as μ increases (faster than the cuspidal part). The coefficient of q^μ in the non-cuspidal part of the theta function of M is labeled $c_\mu(\theta_{\mathfrak{g}(M)})$ in [GL25a]. The only part of the proof of [GL25a, Theorem 6.6.1] that does not go through verbatim is the claim that the ratio of $c_\mu(\theta_{\mathfrak{g}(M)})$ to $c_\mu(\theta_{\mathfrak{g}(\Lambda)})$ is bounded below by a positive constant that does not depend on μ; for this, we must modify the proof of [GL25a, Lemma 6.5.6].

Applying (6.21) and (6.22) of loc. cit. to both M and Λ we obtain

$$c_\mu(\theta_{\mathfrak{g}(M)}) \prod_\ell c_{\mu,\ell}(E_\Lambda; 0) = \frac{m(\mathfrak{g}(\Lambda))}{m(\mathfrak{g}(M))} c_\mu(\theta_{\mathfrak{g}(\Lambda)}) \prod_\ell c_{\mu,\ell}(E_M; 0),$$

where $c_{\mu,\ell}(E_M; 0)$ and $c_{\mu,\ell}(E_\Lambda; 0)$ are certain local densities associated to Eisenstein series, defined in [Shi99, p. 71], and $m(\mathfrak{g}(\Lambda))$ and $m(\mathfrak{g}(M))$ are masses of the genera of Λ and M respectively. Since $M_\ell = \Lambda_\ell$ for all primes $\ell \nmid d$, we may restrict the product on each side to $\ell \mid d$. The proof of [GL25a, Lemma 6.5.6] shows that $c_{\mu,\ell}(E_\Lambda; 0) \neq 0$. So it suffices to prove, for each prime $\ell \mid d$, that $c_{\mu,\ell}(E_M; 0)/c_{\mu,\ell}(E_\Lambda; 0)$ is bounded below by a positive constant that does not depend on μ.

Now let $M' = t + \ell\Lambda$. The proof of [GL25a, Lemma 6.5.6] shows that the ratio $c_{\mu,\ell}(E_{M'}; 0)/c_{\mu,\ell}(E_\Lambda; 0)$ is bounded below by a positive constant that does not depend on μ. Since d is squarefree we have $M_\ell = M'_\ell$ and therefore $c_{\mu,\ell}(E_M; 0) = c_{\mu,\ell}(E_{M'}; 0)$, proving the desired result. ☐

We will apply this local-global principle in the following way.

Corollary 8.5. *Let $\mathfrak{O}$ be a maximal order in a definite quaternion algebra over $\mathbb{Q}$. Let S be an infinite set of positive integers coprime to $\operatorname{disc}(\mathfrak{O})$, and $m \geq 2$ a positive integer. For all $\ell \mid m$ and $s \in S$, let $x_{\ell,s} \in \mathfrak{O}_\ell \smallsetminus \ell\mathfrak{O}_\ell$ be an element satisfying $\operatorname{nrd}(x_{\ell,s}) = s$. Then there exists $x \in \mathfrak{O}$ and $s \in S$ such that $\operatorname{nrd}(x) = s$ and $x \equiv x_{\ell,s} \bmod \ell\mathfrak{O}_\ell$ for all $\ell \mid m$.*

Proof. Let m' be the squarefree product of primes dividing m. For each $s \in S$, there exists $t_s \in \mathfrak{O}$ such that $t_s \equiv x_{\ell,s} \bmod \ell\mathfrak{O}_\ell$ for all $\ell \mid m$. Since S is infinite but $\mathfrak{O}/m'\mathfrak{O}$ is finite, by the pigeonhole principle there exists an infinite subset $S' \subseteq S$ such that the cosets $t_s + m'\mathfrak{O}$ are equal for all $s \in S'$; pick $t \in \mathfrak{O}$ with $t \equiv t_s \bmod m'\mathfrak{O}$ for all $s \in S'$.

Now consider any $s \in S'$. If $\ell \mid m$ then

$$t + m'\mathfrak{O}_\ell = t_s + m'\mathfrak{O}_\ell = t_s + \ell\mathfrak{O}_\ell = x_{\ell,s} + \ell\mathfrak{O}_\ell, \tag{25}$$

which contains an element $x_{\ell,s}$ of norm s. Since $x_{\ell,s} \notin \ell\mathfrak{O}_\ell$ we have $t \notin \ell\mathfrak{O}$. If $\ell \nmid m$ then $t + m'\mathfrak{O}_\ell = \mathfrak{O}_\ell$, which contains an element of norm s because $\operatorname{nrd} : \mathfrak{O}_\ell \to \mathbb{Z}_\ell$ is surjective for maximal orders. So we may apply Theorem 8.4 to conclude that for sufficiently large $s \in S'$, there exists $x \in t + m'\mathfrak{O}$ with $\operatorname{nrd}(x) = s$; for such x we have $x \equiv x_{\ell,s} \bmod \ell\mathfrak{O}_\ell$ by Eq. (25). ☐

Throughout the argument below, if A is a $\mathbb{Z}$-algebra and $x \in A$, and a prime $\ell \in \mathbb{Z}$ is given, we write $\widetilde{x}$ for the image of x in $A/\ell A$.

Lemma 8.6. *Let ℓ be prime and $n \in \mathbb{Z}_\ell^\times$. There exists $a \in \mathbb{Z}_{\ell^2}$ with $N_{\mathbb{Q}_{\ell^2}/\mathbb{Q}_\ell}(a) = n$ and $\widetilde{a} \in \mathbb{F}_{\ell^2} \smallsetminus \mathbb{F}_\ell$.*

Proof. Since the norm map $\mathbb{F}_{\ell^2}^\times \to \mathbb{F}_\ell^\times$ is surjective with fibres of size $\ell + 1$, with at most two elements of each fibre in $\mathbb{F}_\ell^\times$, there exists $\widetilde{a} \in \mathbb{F}_{\ell^2}^\times \smallsetminus \mathbb{F}_\ell^\times$ with $N_{\mathbb{F}_{\ell^2}/\mathbb{F}_\ell}(\widetilde{a}) = n$ in $\mathbb{F}_\ell$. The minimal polynomial of $\widetilde{a}$ then has two distinct roots over $\mathbb{F}_{\ell^2}$, so we can Hensel lift to $a \in \mathbb{Z}_{\ell^2}^\times$ with $N_{\mathbb{Q}_{\ell^2}/\mathbb{Q}_\ell}(a) = n$. ☐

Lemma 8.7. *Let ℓ be a prime, $\mathcal{A}_\ell$ a quaternion algebra over $\mathbb{Q}_\ell$, and $\mathfrak{O}_\ell \subseteq \mathcal{A}_\ell$ a maximal order. Let $n \in \mathbb{Z}_\ell$, and if $\mathcal{A}_\ell$ is a division algebra suppose further that $n \in \mathbb{Z}_\ell^\times$. There exists $x \in \mathfrak{O}_\ell$ with $\operatorname{nrd}(x) = n$ and $\operatorname{disc}(x) \in \mathbb{Z}_\ell^\times$.*

Proof. If $n \in \mathbb{Z}_\ell^\times$, take a as in Lemma 8.6 and x to be the image of a under an embedding $\mathbb{Z}_{\ell^2} \to \mathfrak{O}_\ell$. Letting σ be the nontrivial automorphism of $\mathbb{Q}_{\ell^2}/\mathbb{Q}_\ell$, we have $\operatorname{disc}(x) = (a - \sigma(a))^2$, which is not divisible by ℓ because $\widetilde{a} \neq \widetilde{\sigma(a)}$ in $\mathbb{F}_{\ell^2}$.

If $n \in \ell\mathbb{Z}_\ell$ (in which case we assume $\mathcal{A}_\ell$ is split), $x = \left(\begin{smallmatrix} 0 & -n \\ 1 & 1 \end{smallmatrix}\right)$ has $\operatorname{nrd}(x) = n$ and $\operatorname{disc}(x) = 1 - 4n \in \mathbb{Z}_\ell^\times$. ☐

Lemma 8.8. *Let ℓ be a prime, $\mathcal{A}_\ell$ a quaternion algebra over $\mathbb{Q}_\ell$, and $\mathfrak{O}_\ell \subseteq \mathcal{A}_\ell$ a maximal order. Let $n \in \mathbb{Z}_\ell$, and if $\mathcal{A}_\ell$ is a division algebra suppose further that $n \in \mathbb{Z}_\ell^\times$. Let $x \in \mathfrak{O}_\ell$ and suppose the reduction $\widetilde{x} \in \mathfrak{O}_\ell/\ell\mathfrak{O}_\ell$ is not in $\mathbb{F}_\ell$. Then there exists $y \in \mathfrak{O}_\ell \smallsetminus \ell\mathfrak{O}_\ell$ with $\mathrm{nrd}(y) = n$ such that $1, x, y, xy$ form a $\mathbb{Z}_\ell$-basis for $\mathfrak{O}_\ell$.*

Proof. It suffices to show that there exists $y \in \mathfrak{O}_\ell$ with $\mathrm{nrd}(y) = n$ and with $\mathbb{F}_\ell[\widetilde{x}, \widetilde{y}] = \mathfrak{O}_\ell/\ell\mathfrak{O}_\ell$. For in this case we necessarily have $\mathbb{Z}_\ell[x, y] = \mathfrak{O}_\ell$ by Lemma 2.2, and a computation shows that any element of $\mathbb{Z}_\ell[x, y]$ is contained in $\mathrm{Span}_{\mathbb{Z}_\ell}(1, x, y, xy)$. This condition also forces $\widetilde{y} \neq 0$ so $y \notin \ell\mathfrak{O}_\ell$.

First suppose $\mathcal{A}_\ell$ is split and $n \in \mathbb{Z}_\ell^\times$. Let $\widetilde{x}$ be the reduction of x in $\mathfrak{O}_\ell/\ell\mathfrak{O}_\ell$. If the minimal polynomial of $\widetilde{x}$ over $\mathbb{F}_\ell$ is irreducible, set $y := \left(\begin{smallmatrix} n & 1 \\ 0 & 1 \end{smallmatrix}\right) \in \mathfrak{O}_\ell$. If instead the minimal polynomial of $\widetilde{x}$ over $\mathbb{F}_\ell$ is reducible, consider the element a from Lemma 8.6 and let y be the image of a under an embedding $\mathbb{Z}_{\ell^2} \to \mathfrak{O}_\ell$, so that $\widetilde{y}$ has irreducible minimal polynomial. In either case, $\mathbb{F}_\ell[\widetilde{x}, \widetilde{y}]$ contains both a copy of $\mathbb{F}_{\ell^2}$ and an element with reducible minimal polynomial that is not in $\mathbb{F}_\ell$, and hence not in $\mathbb{F}_{\ell^2}$. This implies that $\mathbb{F}_\ell[\widetilde{x}, \widetilde{y}] = \mathfrak{O}_\ell/\ell\mathfrak{O}_\ell$ because $\mathbb{F}_\ell[\widetilde{x}, \widetilde{y}]$ and $\mathfrak{O}_\ell/\ell\mathfrak{O}_\ell$ are both vector spaces over $\mathbb{F}_{\ell^2}$, the former of dimension greater than 1 and the latter of dimension 2.

Now let $\ell = p$ with $\mathcal{A}_p$ a division algebra, so $n \in \mathbb{Z}_p^\times$. In this case $\mathfrak{O}_p$ has the form

$$\mathfrak{O}_p = \left\{ \begin{pmatrix} u & pv \\ \sigma(v) & \sigma(u) \end{pmatrix} : u, v \in \mathbb{Z}_{p^2} \right\},$$

where $\mathbb{Z}_{p^2}$ is the ring of integers of the unramified quadratic extension $\mathbb{Q}_{p^2}/\mathbb{Q}_p$ and σ is the non-trivial element of $\mathrm{Gal}(\mathbb{Q}_{p^2}/\mathbb{Q}_p)$. If the reduction $\widetilde{x}$ is not in the diagonal subring of $\mathfrak{O}_p/p\mathfrak{O}_p$, let $y = \left(\begin{smallmatrix} a & 0 \\ 0 & \sigma(a) \end{smallmatrix}\right)$ for a as in Lemma 8.6. If $\widetilde{x}$ is in the diagonal subring (and hence $\mathbb{F}_\ell[\widetilde{x}] \simeq \mathbb{F}_{\ell^2}$ because $\widetilde{x} \notin \mathbb{F}_\ell$), we can apply Lemma 8.6 to get an element $b \in \mathbb{Z}_{p^2}$ with $N_{\mathbb{Q}_{p^2}/\mathbb{Q}_p}(b) = n + p$, and take $y = \left(\begin{smallmatrix} b & p \\ 1 & \sigma(b) \end{smallmatrix}\right)$. Again $\mathbb{F}_\ell[\widetilde{x}, \widetilde{y}]$ contains both an embedding of $\mathbb{F}_{\ell^2}$ and an element that is not in the image of this embedding, so it must equal $\mathfrak{O}_\ell/\ell\mathfrak{O}_\ell$.

Finally let $\mathcal{A}_\ell$ be split and $n \in \ell\mathbb{Z}_\ell$. The element $y_0 = \left(\begin{smallmatrix} -1 & n+1 \\ -1 & 1 \end{smallmatrix}\right) \in M_2(\mathbb{Z}_\ell)$ has $\mathrm{nrd}(y_0) = n$ and characteristic polynomial $x^2 + n$; its reduction $\widetilde{y}_0 \in M_2(\mathbb{F}_\ell)$ has a unique 1-dimensional eigenspace, the span of the column vector $(1, 1)^t$. Since $\widetilde{x} \in M_2(\mathbb{F}_\ell)$ is not a scalar matrix, there exists a nonzero $v \in \mathbb{F}_\ell^2$ which is not an eigenvector for $\widetilde{x}$. Let $\widetilde{w} \in \mathrm{GL}_2(\mathbb{F}_\ell)$ satisfy $\widetilde{w}(1, 1)^t = v$, let $w \in \mathrm{GL}_2(\mathbb{Z}_\ell)$ be a lift of $\widetilde{w}$, and let $y = w y_0 w^{-1} \in M_2(\mathbb{Z}_\ell)$. Then every eigenvector of $\widetilde{y}$ is in the span of v, so $\widetilde{x}$ and $\widetilde{y}$ have no eigenvectors in common. Note also that $\widetilde{y}$ is a nonzero nilpotent element. By [Rac74, Theorem 1], every proper subalgebra of $M_2(\mathbb{F}_\ell)$ is either contained in the image of an embedding of $\mathbb{F}_{\ell^2}$ or has a common eigenvector. As neither of these are true of $\mathbb{F}_\ell[\widetilde{x}, \widetilde{y}]$, we can conclude that $\mathbb{F}_\ell[\widetilde{x}, \widetilde{y}] = \mathfrak{O}_\ell/\ell\mathfrak{O}_\ell$. □

Proof of Theorem 8.2. We begin by taking any $\alpha \in \mathfrak{O} \smallsetminus \mathbb{Z}$ with $\mathrm{nrd}(\alpha) \in S$. Here is one way to construct such an α. Fix any prime ℓ. Then for any $s \in S$ there exists $x_{\ell,s} \in \mathfrak{O}_\ell$ with $\mathrm{nrd}(x_{\ell,s}) = s$ and $\ell \nmid \mathrm{disc}(x_{\ell,s})$ by Lemma 8.7. The discriminant condition forces $x_{\ell,s} \notin \ell\mathfrak{O}_\ell$, so we may apply Corollary 8.5 to obtain $\alpha \in \mathfrak{O}$ and $s_1 \in S$ with $\mathrm{nrd}(\alpha) = s_1$ and $\alpha \equiv x_{\ell,s_1} \bmod \ell\mathfrak{O}_\ell$. Since $\mathrm{disc}(\alpha) \equiv \mathrm{disc}(x_{\ell,s_1}) \not\equiv 0 \bmod \ell\mathbb{Z}_\ell$ we have $\mathrm{disc}(\alpha) \neq 0$ and hence $\alpha \notin \mathbb{Z}$.

Now for all $\ell \mid \mathrm{disc}(\alpha)$ and $s \in S$, use Lemma 8.7 to pick $y_{\ell,s} \in \mathfrak{O}_\ell$ with $\mathrm{nrd}(y_{\ell,s}) = s$ and $\mathrm{disc}(y_{\ell,s}) \in \mathbb{Z}_\ell^\times$. Applying Corollary 8.5 again, we obtain $\beta \in \mathfrak{O} \smallsetminus \mathbb{Z}$ and $s_2 \in S$ with $\mathrm{nrd}(\beta) = s_2$ and $\beta \equiv y_{\ell,s_2} \bmod \ell\mathfrak{O}_\ell$ for all $\ell \mid \mathrm{disc}(\alpha)$. Then $\mathrm{disc}(\beta) \equiv \mathrm{disc}(y_{\ell,s_2}) \not\equiv 0 \bmod \ell\mathbb{Z}_\ell$ for all $\ell \mid \mathrm{disc}(\alpha)$, so $\mathrm{disc}(\beta)$ is coprime to $\mathrm{disc}(\alpha)$.

Finally, take $\Lambda := \mathrm{Span}_{\mathbb{Z}}(1, \alpha, \beta, \alpha\beta)$. Since $\mathbb{Z}[\alpha]$ and $\mathbb{Z}[\beta]$ are quadratic orders with coprime discriminants, Λ is a finite-index sublattice of $\mathfrak{O}$, and we set $m := [\mathfrak{O} : \Lambda]$. If $m = 1$ then we are done, so assume $m \geq 2$. Let ℓ be a prime factor of m and $s \in S$.

Suppose first that $\ell \nmid \mathrm{disc}(\alpha)$. Then $\widetilde{\alpha} \notin \mathbb{F}_\ell$, so by Lemma 8.8, there exists $z_{\ell,s} \in \mathfrak{O}_\ell \smallsetminus \ell\mathfrak{O}_\ell$ with $\mathrm{nrd}(z_{\ell,s}) = s$ and such that $\mathrm{Span}_{\mathbb{Z}_\ell}(1, \alpha, z_{\ell,s}, \alpha z_{\ell,s}) = \mathfrak{O}_\ell$. If instead $\ell \mid \mathrm{disc}(\alpha)$, then $\ell \nmid \mathrm{disc}(\beta)$ by construction of β, so by Lemma 8.8, there exists $z_{\ell,s} \in \mathfrak{O}_\ell \smallsetminus \ell\mathfrak{O}_\ell$ with $\mathrm{nrd}(z_{\ell,s}) = s$ and such that $\mathrm{Span}_{\mathbb{Z}_\ell}(1, \beta, z_{\ell,s}, \beta z_{\ell,s}) = \mathfrak{O}_\ell$. Once again we apply Corollary 8.5 to obtain $\gamma \in \mathfrak{O}$ and $s_3 \in S$ with $\mathrm{nrd}(\gamma) = s_3$, $\gamma \equiv z_{\ell,s_3} \bmod \ell\mathfrak{O}_\ell$ for all $\ell \mid m$. Applying Lemma 2.2, we have the following generating sets for $\mathfrak{O}_\ell$.

- If $\ell \nmid m$ then $\mathfrak{O}_\ell = \mathrm{Span}_{\mathbb{Z}_\ell}(1, \alpha, \beta, \alpha\beta)$.
- If $\ell \mid m$ and $\ell \nmid \mathrm{disc}(\alpha)$ then $\mathfrak{O}_\ell = \mathrm{Span}_{\mathbb{Z}_\ell}(1, \alpha, \gamma, \alpha\gamma)$.
- If $\ell \mid m$ and $\ell \mid \mathrm{disc}(\alpha)$ then $\mathfrak{O}_\ell = \mathrm{Span}_{\mathbb{Z}_\ell}(1, \beta, \gamma, \beta\gamma)$.

Thus the index of $\mathrm{Span}_{\mathbb{Z}}(1, \alpha, \beta, \gamma, \alpha\beta, \alpha\gamma, \beta\gamma)$ in $\mathfrak{O}$ cannot be divisible by any prime, and hence it equals 1. □

Remark 8.9. In Theorem 1.1 we bound the norms of elements in a generating set, but with only very weak bounds on the number of elements in the generating set (Proposition 8.1). By contrast, in Theorem 8.2, we bound the number of elements in a generating set, but have no bound on the norms of the elements. One might ask for a result that balances the two: is it possible to find a generating set consisting of a very small number of elements, each of small norm?

The two results are similar in that they both rely on decomposing a theta function into cuspidal and non-cuspidal contributions, and showing that for suitable n, the coefficient of q^n is dominated by the non-cuspidal contribution. To combine the two approaches, we would need a result analogous to Theorem 6.1, giving an explicit bound on the coefficients of the cusp form components for *cosets* of integral ideals within a maximal order, rather than for suborders. It is as yet unclear whether there is a generalized notion of Brandt module that can produce these inhomogeneous theta functions.

9. Experimental results

In this section we make computational observations demonstrating our main theorems and estimating the tightness of some results. All code and computations were done in `Magma V29.9` [BCP97]. A GitHub repository containing the code used for generating figures and running examples is available at https://GitHub.com/wmahaney/generating_sets_for_maximal_quaternion_orders.git.

9.1. **Estimating minimal exponent for prime power generation.** Theorem 1.1 implies that the endomorphism ring of a supersingular elliptic curve $E/\overline{\mathbb{F}}_p$ can be generated by elements of degree ℓ^k for relatively small values of k. In this section we test this explicitly and see if, in practice, the lower bound on allowable k can be reduced.

First, if E_0 is a uniformly randomly selected supersingular elliptic curve over $\overline{\mathbb{F}}_p$ with p sufficiently large relative to ℓ, and $k \gtrsim 2\log_\ell(p)$, we find that with high probability there is a finite index suborder of $\mathrm{End}(E_0)$ generated as a lattice by some elements of degree ℓ^{k_0} with $0 \le k_0 \le k$. A heuristic argument for this may be given using the bound $d \le 1 + 2\log_\ell\left(\frac{p-1}{\sqrt{6}}\right)$ on the diameter of $G(p,\ell)$ [BS25, Corollary 32], a strengthening of a classical bound due to Pizer [Piz90], together with the expansion properties of $G(p,\ell)$.

Experimentally, we find something stronger for every example we tested:

Experimental Observation 9.1. *Let $k = 1 + 2\lceil \log_\ell\left(\frac{p-1}{\sqrt{6}}\right)\rceil$. For any maximal order $\mathfrak{O}$ in $\mathcal{A}_p$, the elements of norm ℓ^k generate $\mathfrak{O}$ as a $\mathbb{Z}$-algebra.*

The interested reader can check the GitHub repository to see certificates from some random tests and other output data[3] and run tests themselves with the functions.[4]

[3] https://github.com/wmahaney/generating_sets_for_maximal_quaternion_orders/blob/main/data
[4] https://github.com/wmahaney/generating_sets_for_maximal_quaternion_orders/blob/main/does_generate.magma

Given fixed p and ℓ, we can use Theorem 1.1 to compute $k = k(p, \ell)$ such that a maximal order $\mathfrak{O} \subset \mathcal{A}_p$ is guaranteed to be generated as a $\mathbb{Z}$-algebra by the set of all elements in $\mathfrak{O}$ of norm ℓ^k; the experimental results are consistently lower than this theoretical bound. For instance, if $p \approx \ell^{30}$, then Theorem 1.1 gives a lower bound of $k(p, \ell) \approx 75$ while Experimental Observation 9.1 gives a lower bound of $1 + 2\log_\ell\left(\frac{p-1}{\sqrt{6}}\right) \approx 61$.

9.2. **Estimating $|a_{\ell^k}(\mathcal{O}) - g_{\ell^k}(\mathcal{O})|$.** Following the proof of Proposition 7.1 one can show that all maximal and submaximal orders $\mathcal{O} \subset \mathcal{A}_p$ with $p \geq 5$ satisfy

$$|a_{\ell^k}(\mathcal{O}) - g_{\ell^k}(\mathcal{O})| \leq 2k\ell^{\frac{k+2}{2}}|\mathcal{O}^\times|,$$

where $a_{\ell^k}(\mathcal{O})$ is the ℓ^k-coefficient of the theta series $\theta_{\mathcal{O}}$, and $g_{\ell^k}(\mathcal{O})$ is the ℓ^k-coefficient of the modular form g defined in Corollary 6.3(a). Recall that g is the "genus contribution," which in some sense measures the expected size of $a_{\ell^k}(\mathcal{O})$ across all $\mathcal{O}$ in its genus. For $p \geq 5$ we have $|\mathcal{O}^\times| \leq 6$, so this is a stronger bound than Proposition 7.1.

In this section we compute data to estimate how tight the above bound is by calculating the ratio

$$\mathscr{R}_{\ell^k}(\mathcal{O}) = \frac{|a_{\ell^k}(\mathcal{O}) - g_{\ell^k}(\mathcal{O})|}{2k\ell^{\frac{k+2}{2}}|\mathcal{O}^\times|}.$$

We will only present plots and discussion in the case $\ell = 2$, as the plots for $\ell = 3$ look similar; the interested reader can refer to the GitHub to see additional figures or generate their own.

For each prime $p \in [7, 50000]$ and reduced discriminant $N \in \{p, p^2, p^3, pr, pr^2\}$ (we use $r = 5$ here) corresponding to a maximal/submaximal order classification we do the following for each $k \in [1, 15]$:

(1) Randomly sample five orders $\mathcal{O}_i$ of the given classification and compute the ratios $\mathscr{R}_{2^k}(\mathcal{O}_i)$.
(2) Compute the average $C_{2^k}(N)$ of $\mathscr{R}_{2^k}(\mathcal{O}_1), \ldots, \mathscr{R}_{2^k}(\mathcal{O}_5)$.
(3) Plot $(k, p, C_{2^k}(N))$.

To randomize the order chosen in each trial we generated a quaternion order in Magma and then took a simple random walk of length $\lceil \log_2(p) \rceil$ from the generated order. We ran five trials for each p and discriminant classification, using $k_{\min} = 1, k_{\max} = 15$, and a value of $r = 5$ when necessary. The full plots are presented below.

Observationally, the ratio $\mathscr{R}_{2^k}(\mathcal{O})$ never exceeded 0.35 on average, and tends to zero if p is fixed and k tends to infinity (and $r = 5$ when N is not a power of p), indicating that there may be a stronger bound than Proposition 7.1.

References

[AL70] A. O. L. Atkin and J. Lehner. Hecke operators on $\Gamma_0(m)$. *Math. Ann.*, 185:134–160, 1970. doi:10.1007/BF01359701.

[BCNE+19] Efrat Bank, Catalina Camacho-Navarro, Kirsten Eisenträger, Travis Morrison, and Jennifer Park. Cycles in the supersingular ℓ-isogeny graph and corresponding endomorphisms. In Jennifer S. Balakrishnan, Amanda Folsom, Matilde Lalín, and Michelle Manes, editors, *Research Directions in Number Theory*, volume 19 of *Association for Women in Mathematics Series*, pages 41–66. Springer, Cham, 2019. doi:10.1007/978-3-030-19478-9_2.

[BCP97] Wieb Bosma, John Cannon, and Catherine Playoust. The Magma algebra system. I. The user language. *J. Symbolic Comput.*, 24(3-4):235–265, 1997. Computational algebra and number theory (London, 1993). URL: http://dx.doi.org/10.1006/jsco.1996.0125, doi:10.1006/jsco.1996.0125.

[Brz83] J. Brzeziński. On orders in quaternion algebras. *Comm. Algebra*, 11(5):501–522, 1983. doi:10.1080/00927878308822861.

[BS25] Sebastiano Boscardin and Sebastian A. Spindler. Exploring Kaneko's bound: On multi-edges, loops and the diameter of the supersingular ℓ-isogeny graph. Cryptology ePrint Archive, Paper 2025/1361, 2025. URL: https://eprint.iacr.org/2025/1361.

[CF10] J. W. S. Cassels and A. Fröhlich, editors. *Algebraic number theory*. London Mathematical Society, London, 2010. Papers from the conference held at the University of Sussex, Brighton, September 1–17, 1965, Including a list of errata.

[CFG+26] Wouter Castryck, Luca De Feo, Steven D. Galbraith, Péter Kutas, Krijn Reijnders, and Benjamin Wesolowski. The isogeny problems. Cryptology ePrint Archive, Paper 2026/1431, 2026. URL: https://eprint.iacr.org/2026/1431.

[CLG09] Denis X. Charles, Kristin E. Lauter, and Eyal Z. Goren. Cryptographic hash functions from expander graphs. *J. Cryptology*, 22(1):93–113, 2009. doi:10.1007/s00145-007-9002-x.

[CSV21] Sara Chari, Daniel Smertnig, and John Voight. On basic and Bass quaternion orders. *Proc. Amer. Math. Soc. Ser. B*, 8:11–26, 2021. doi:10.1090/bproc/68.

[DS05] Fred Diamond and Jerry Shurman. *A first course in modular forms*, volume 228 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2005.

[EHL+20] Kirsten Eisenträger, Sean Hallgren, Chris Leonardi, Travis Morrison, and Jennifer Park. Computing endomorphism rings of supersingular elliptic curves and connections to path-finding in isogeny graphs. In *ANTS XIV—Proceedings of the Fourteenth Algorithmic Number Theory Symposium*, volume 4 of *Open Book Ser.*, pages 215–232. Math. Sci. Publ., Berkeley, CA, 2020. doi:10.2140/obs.2020.4.215.

[Eic55] Martin Eichler. Zur Zahlentheorie der Quaternionen-Algebren. *J. Reine Angew. Math.*, 195:127–151 (1956), 1955. doi:10.1515/crll.1955.195.127.

[Eic73] M. Eichler. The basis problem for modular forms and the traces of the Hecke operators. In *Modular functions of one variable, I (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972)*, volume Vol. 320 of *Lecture Notes in Math.*, pages 75–151. Springer, Berlin-New York, 1973. doi:10.1007/978-3-540-38509-7_4.

[GL25a] Eyal Z. Goren and Jonathan Love. *Supersingular Elliptic Curves, Quaternion Algebras and Applications to Cryptography*, pages 123–200. SAGE Publications, 2025. doi:10.3233/NICSP250005.

[GL25b] Eyal Z. Goren and Jonathan R. Love. On elements of prescribed norm in maximal orders of a quaternion algebra. *Canadian Journal of Mathematics*, 77(6):1938–1965, 2025. URL: http://dx.doi.org/10.4153/S0008414X24000592, doi:10.4153/s0008414x24000592.

[Gro87] Benedict H. Gross. Heights and the special values of L-series. In *Number Theory (Montreal, Que., 1985)*, volume 7 of *CMS Conf. Proc.*, pages 115–187. Amer. Math. Soc., Providence, RI, 1987.

[HKTV25] Chenfeng He, Gaurish Korpal, Ha T. N. Tran, and Christelle Vincent. Gross lattices of supersingular elliptic curves, 2025. URL: https://arxiv.org/abs/2503.03478, arXiv:2503.03478.

[HPS89] H. Hijikata, A. Pizer, and T. Shemanske. Orders in quaternion algebras. *J. Reine Angew. Math.*, 394:59–106, 1989. doi:10.1515/crll.1989.394.59.

[HW26] Arthur Herlédan Le Merdy and Benjamin Wesolowski. Unconditional foundations for supersingular isogeny-based cryptography. In Benny Applebaum and Huijia (Rachel) Lin, editors, *Theory of Cryptography*, pages 266–297, Cham, 2026. Springer Nature Switzerland. doi:10.1007/978-3-032-12296-4_9.

[KLPT14] David Kohel, Kristin Lauter, Christophe Petit, and Jean-Pierre Tignol. On the quaternion ℓ-isogeny path problem. *LMS Journal of Computation and Mathematics*, 17:418–432, 2014. doi:10.1112/S1461157014000151.

[Koh96] David R. Kohel. *Endomorphism Rings of Elliptic Curves over Finite Fields*. Ph.d. thesis, University of California, Berkeley, 1996. URL: https://www.proquest.com/docview/304241260.

[Koh01] David R. Kohel. Hecke module structure of quaternions. In *Class field theory—its centenary and prospect (Tokyo, 1998)*, volume 30 of *Adv. Stud. Pure Math.*, pages 177–195. Math. Soc. Japan, Tokyo, 2001. doi:10.2969/aspm/03010177.

[Mes86] J.-F. Mestre. La méthode des graphes. Exemples et applications. In *Proceedings of the International Conference on Class Numbers and Fundamental Units of Algebraic Number Fields (Katata, 1986)*, pages 217–242. Nagoya Univ., Nagoya, 1986.

[Ogg69] Andrew Ogg. *Modular forms and Dirichlet series*. W. A. Benjamin, Inc., New York-Amsterdam, 1969.

[Piz80a] Arnold Pizer. An algorithm for computing modular forms on $\Gamma_0(N)$. *J. Algebra*, 64(2):340–390, 1980. doi:10.1016/0021-8693(80)90151-9.

[Piz80b] Arnold Pizer. Theta series and modular forms of level p^2m. *Compositio Mathematica*, 40(2):177–241, 1980. URL: http://eudml.org/doc/89433.

[Piz90] Arnold K. Pizer. Ramanujan graphs and Hecke operators. *Bull. Amer. Math. Soc. (N.S.)*, 23(1):127–137, 1990. URL: https://projecteuclid.org/journalArticle/Download?urlid=bams%2F1183555725.

[Pon09] Paul Ponomarev. Newforms of squarefree level and theta series. *Math. Ann.*, 345(1):185–193, 2009. doi:10.1007/s00208-009-0348-2.

[Rac74] M. L. Racine. On maximal subalgebras. *J. Algebra*, 30:155–180, 1974. doi:10.1016/0021-8693(74)90198-7.

[Shi71] Goro Shimura. *Introduction to the Arithmetic Theory of Automorphic Functions*, volume 11 of *Publications of the Mathematical Society of Japan*. Iwanami Shoten Publishers and Princeton University Press, Tokyo; Princeton, NJ, 1971.

[Shi99] Goro Shimura. The number of representations of an integer by a quadratic form. *Duke Math. J.*, 100(1):59–92, 1999. doi:10.1215/S0012-7094-99-10002-0.

[Sie35] Carl Ludwig Siegel. über die analytische Theorie der quadratischen Formen. *Ann. of Math. (2)*, 36(3):527–606, 1935. doi:10.2307/1968644.

[Voi21] John Voight. *Quaternion algebras*, volume 288 of *Graduate Texts in Mathematics*. Springer, Cham, ©2021. doi:10.1007/978-3-030-56694-4.

[Wes21] Benjamin Wesolowski. The supersingular isogeny path and endomorphism ring problems are equivalent. *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1100–1111, 2021. URL: https://api.semanticscholar.org/CorpusID:235810201.

Kirsten Eisenträger, Department of Mathematics, The Pennsylvania State University, 109 McAllister Building, University Park, PA 16802, USA. eisentra@math.psu.edu

Eyal Z. Goren, Department of Mathematics and Statistics, McGill University, 805 Sherbrooke St. W, Montréal, QC H3A0B9, Canada. eyal.goren@mcgill.ca

Annamaria Iezzi, Univ. Grenoble Alpes, CNRS, Grenoble INP, LJK, 38000 Grenoble, France. annamaria.iezzi@univ-grenoble-alpes.fr

Harun Kir, UMPA, ENS de Lyon, CNRS, UMR 5669, 46 allée d'Italie, 69364 Lyon Cedex 07, France. harun.kir@ens-lyon.fr

Eda Kirimli, University of Birmingham, University Road W, Edgbaston, Birmingham, B15 2TT, UK. e.kirimli@bham.ac.uk

Jonathan R. Love, Department of Mathematics, University of Michigan, 530 Church Street, Ann Arbor, MI 48109, USA. jonlove@umich.edu

William E. Mahaney, Department of Mathematics, Virginia Tech, 225 Stanger Street, Blacksburg, Virginia 24060, USA. mahaney1@vt.edu

Jennifer Park, Department of Mathematics, The Ohio State University, 231 West 18th Avenue, Columbus, OH 43210, USA. park.2720@osu.edu

Maria Sabitova, Department of Mathematics, Queens College, CUNY, 65-30 Kissena Blvd, Flushing, NY 11367, USA. maria.sabitova@qc.cuny.edu

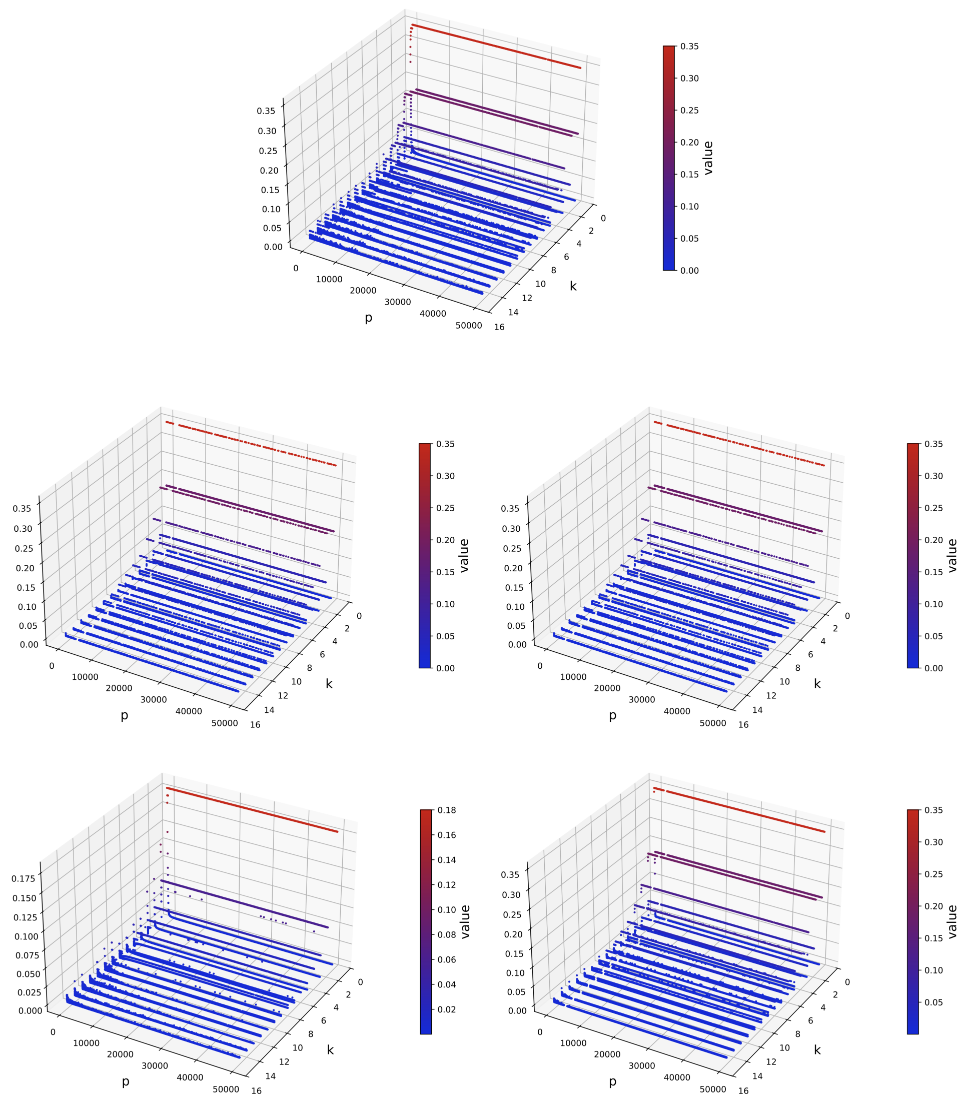

FIGURE 1. Ratio of $|a_{\ell^k}(\mathcal{O}) - g_{\ell^k}(\mathcal{O})|$ to $2k\ell^{\frac{k+2}{2}}|\mathcal{O}^\times|$ for $\ell = 2$ and $k = 1, \ldots, 15$. Primes p range over $[7, 50000]$. From top-to-bottom and left-to-right the order classification is maximal, submaximal of index p, submaximal of index p^2, submaximal of index r, and submaximal of index r^2