# SECTIONS 1–3, 10, 11 AND 14

MIHRAN PAPIKIAN

Abstract. The purpose of these notes is to clarify and give complete proofs of some of the statements in [4, §§1-3, 10, 11 and 14]. These sections comprise the "geometric" portion of the paper leading to the calculation of a certain height pairing, relating this height to special values of $L$-functions of cusp forms, and then comparing the final result with the conjecture of Birch and Swinnerton-Dyer.

## 1. Central simple algebras

Although the only central simple algebras which come up in [4] are the quaternions ramified at two places it will be more instructive to discuss algebras of arbitrary dimension. In fact, many properties of quaternions used in [4] are valid in general and proving some of these properties requires a detour into the theory of Brauer groups and hence naturally leads to the study of general central simple algebras.

1.1. **Definitions and examples.** Let $K$ be a field. By a *$K$-algebra* we mean a (not necessarily commutative) ring $A$ containing $K$ in its center and finite dimensional as a $K$-vector space. A *$K$-subalgebra* of a $K$-algebra $A$ is a subring containing $K$. A *homomorphism of $K$-algebras* $\varphi : A \to B$ is a homomorphism of rings with the property that $\varphi(a) = a$ for all $a \in K$. A $K$-algebra $A$ is simple if it contains no two sided ideals except 0 and $A$. Note that a $K$-homomorphism $\varphi : A \to B$ from a simple $K$-algebra $A$ to an arbitrary $K$-algebra $B$ must be injective as $\ker(\phi)$ is an ideal of $A$ not containing 1. A $K$-algebra is a *division algebra* if every non-zero element $a \in A$ has an inverse, i.e., there exists $b \in A$ such that $ab = 1 = ba$. Clearly division algebras are simple. A $K$-algebra $A$ is *central* if its center is $K$.

*Example* 1.1. The algebra of $n \times n$ matrices $M_n(K)$ over $K$ is both central and simple. Indeed, denote by $e_{i,j}$ the $n \times n$ matrix which has 1 in its $(i, j)$ entry and zeros elsewhere. Clearly $e_{i,j}$, $1 \leq i, j \leq n$, form a basis of $M_n(K)$. Now let $I$ be a two-sided non-zero ideal in $M_n(K)$. There is an element $a \in I$ which has a non-zero entry, say in $(i, j)$ position $a_{i,j}$. Since $e_{i,i}ae_{j,j} = a_{i,j}e_{i,j}$ and $e_{i,j}e_{j,k} = e_{i,k}$ we get $I = M_n(K)IM_n(K) = M_n(K)$. The center of $M_n(K)$ consists of scalar multiples of $\mathrm{diag}(1, 1, \cdots, 1)$. Hence this algebra is also central.

**Theorem 1.2** (Wedderburn Structure Theorem)**.** *Any simple $K$-algebra $A$ is isomorphic to the algebra $M_n(D)$ of $n \times n$ matrices over a division algebra $D$. Moreover, $n$ and $D$ are determined uniquely by $A$. ($D$ is determined uniquely up to an isomorphism.)*

*Proof.* For the proof see either [9, Thm. IV.1.9] or [12, (7.4)]. □

These are notes from a reading seminar for graduate students that I co-organized with Ben Brubaker at Stanford in 2004.

1.2. **Brauer group.** For a field $K$ to define the Brauer group $\mathrm{Br}(K)$ one first needs to prove the following fact: If $A$ and $B$ are central simple algebras over $K$ then $A \otimes_K B$ is again a central simple algebra over $K$. (For a proof see either [12] or [9].) This suggests that central simple algebras over a field might form a group. To make this idea precise, define an equivalence relation on the set of such algebras by declaring $A \sim B$ if $A \otimes_K M_n(K) \cong B \otimes_K M_m(K)$ for some $n$ and $m$. This is indeed an equivalence relation as $M_{n_1}(K) \otimes M_{n_2}(K) \cong M_{n_1 n_2}(K)$. Define the *Brauer group* $\mathrm{Br}(K)$ to be the set of equivalence classes of central simple algebras under the above equivalence relation. The group operation is defined by $[A]\cdot[B] = [A \otimes B]$. The equivalence class of a matrix algebra is the unit element, and the inverse of $[A]$ is $[A^{\mathrm{opp}}]$, where $A^{\mathrm{opp}}$ is the algebra having the same set of elements as $A$ with the multiplication in $A^{\mathrm{opp}}$ defined via the multiplication in $A$ by $a * b := b \cdot a$. It is another fact, whose proof is given in *loc.cit.*, that $[A][A^{\mathrm{opp}}] = 1$.

Note that every class $[A]$ in $\mathrm{Br}(K)$ contains a unique, up to an isomorphism, central division algebra. Indeed, this follows from Theorem 1.2 and the fact that for two central division algebras $D_1$ and $D_2$ we have $M_n(D_1) \otimes M_m(D_2) \cong M_{nm}(D_1 \otimes D_2)$. The point of introducing the Brauer group, instead of studying the central division algebras directly, is that the very natural operation of forming tensor product of two such algebras takes us out of the category of division algebras but remains in the category of central simple algebras.

**Theorem 1.3.**

(1) *If $K$ is algebraically closed then* $\mathrm{Br}(K) = 0$.
(2) *There is an isomorphism* $\mathrm{inv}_{\mathbb{R}} : \mathrm{Br}(\mathbb{R}) \cong \mathbb{Z}/2$ *with the Hamiltonians giving the nontrivial element.*
(3) *If $K$ is a non-archimedean local field then there is a natural homomorphism* $\mathrm{inv}_K : \mathrm{Br}(K) \to \mathbb{Q}/\mathbb{Z}$ *which turns out to be an isomorphism. Moreover, the image of a central division algebra $D$ of dimension $n^2$ over $K$ has exact order $n$ in $\mathbb{Q}/\mathbb{Z}$.*
(4) *If $K$ is a global field (that is, either a number field or the field of rational functions on a smooth projective geometrically connected curve over a finite field) then there is an exact sequence*

$$0 \to \mathrm{Br}(K) \to \bigoplus_v \mathrm{Br}(K_v) \to \mathbb{Q}/\mathbb{Z} \to 0,$$

*where the sum in the middle is over all (including archimedean) places of $K$, $K_v$ is the completion of $K$ at the place $v$ and the map to $\mathbb{Q}/\mathbb{Z}$ is given by $\oplus_v \mathrm{inv}_{K_v}$.*

*Proof.* First part is clear since there are no division algebras over an algebraically closed field except $K$ itself. Indeed, if $D$ is such an algebra then for $\alpha \in D$, $K(\alpha)$ is a commutative integral domain over $K$ and hence is a finite field extension of $K$. This implied $\alpha \in K$. Second part is classical (known as Frobenius theorem). For the proof see [5, Cor. IX.6.8]. Parts (3) and (4) are among the main statements of local and global class field theory respectively, and are very non-trivial to prove; see [9]. The only remark to make is why the dimension of a central simple algebra is a square. In fact, given a central simple algebra $A$ over a field $K$, $A \otimes_K K^{\mathrm{alg}} \cong M_n(K^{\mathrm{alg}})$. Hence $\dim_K(A) = \dim_{K^{\mathrm{alg}}}(A \otimes K^{\mathrm{alg}}) = n^2$. □

1.3. **Split and ramified central simple algebras.** A central simple algebra $A$ over a field $K$ is said to be *split* if it is isomorphic to $M_n(K)$ for some $n$, otherwise it is called *ramified*. This terminology is used in [4], and although it is not very important we take a moment to explain the terminology.

A field $L$ containing $K$ is said to be a *splitting field* of $A$ if $A \otimes_K L$ is a split algebra over $L$. We have seen that $K^{\mathrm{alg}}$ is such a field. It is not hard to show that some finite extension of $K$ will have the same property. Indeed, using Theorem 1.2 and dimension comparisons, it is enough to show that given a central division algebra $D$ over $K$ there is a finite extension $L$ such that $D \otimes_K L$ is no longer a division algebra. Fix a basis of $D$ over $K$, and consider the multiplication table of $D$ with respect to this basis. The equation $a \cdot b = 0$, $a, b \neq 0$ will have a solution in $D \otimes L$ for some finite extension $L$ of $K$, as this is equivalent to the existence of a non-zero solution of a system of polynomial equations which come from the multiplication table. Hence it is natural to ask what is the minimal degree of a field $L$ that splits $D$. The answer is the following [12, (7.15)]:

**Theorem 1.4.** *Suppose $D$ has dimension $n^2$ over $K$. Every maximal subfield of $D$ contains $K$ and is a splitting field for $D$. Such subfields always exist and are characterized by $[L : K] = n$. Hence $D \otimes_K L \cong M_n(L)$. Moreover, every field which splits $D$ and has degree $n$ over $K$ is isomorphic to a maximal subfield of $D$. There exists a splitting field which is separable over $K$.*

Now assume the base field $K$ is a local non-archimedean field. Let $D$ be a division algebra over $K$. For any $\alpha \in D$, $L = K(\alpha)$ is a finite field extension of $K$ and hence by a well-known property the valuation on $K$ can be uniquely extended to $L$. Since $L$ is a subfield of $D$, by Theorem 1.4 $[K(\alpha) : K] \leq n$, and hence the valuation on $L$ takes values in $\mathbb{Z}[n^{-1}]$. In this manner we can attach a valuation (that is, an element in $\mathbb{Z}[n^{-1}]$) to every element of $D$. A crucial (non-trivial) fact is that this can be done compatibly; see [12, §12]. More precisely, the map $w : D \to \mathbb{Z}[n^{-1}]$ defined above has the following properties:

(1) $w(\alpha) = \infty$ if and only if $\alpha = 0$,
(2) $w(\alpha\beta) = w(\alpha) + w(\beta) = w(\beta\alpha)$.
(3) $w(\alpha + \beta) \geq \min(w(\alpha), w(\beta))$.

Now we can define a natural subring of $D$, the subring of integers $\mathcal{O}_D = \{\alpha \in D \mid w(\alpha) \geq 0\}$, which contains the ring of integers $\mathcal{O}_K$ of $K$. There is a two-sided maximal ideal $\mathfrak{P} = \{\alpha \in D \mid w(\alpha) > 0\}$ of $\mathcal{O}_D$, and every two-sided ideal of $\mathcal{O}_D$ is a power of $\mathfrak{P}$; see [12, (13.2)]. Let $\mathfrak{p}$ be the maximal ideal $\mathcal{O}_K$, then $\mathfrak{p}\mathcal{O}_D$ is a two sided ideal of $\mathcal{O}_D$ and hence for some $e \geq 0$ we have $\mathfrak{P}^e = \mathfrak{p}\mathcal{O}_D$. It is known [12, (13.3)] that $e = n$. In particular, in every central simple algebra of the form $M_n(D)$, with $D$ a non-commutative division algebra, $\mathfrak{p}$ "ramifies".

On the other hand, if the central simple algebra $A \cong M_n(K)$, then a natural subring is $\mathcal{O}_M = M_n(\mathcal{O}_K)$. In this case, the left ideal $\mathcal{O}_M\mathfrak{p}$ is a product of $n$ left ideals $I_i$ of determinant $\mathfrak{p}$, where $I_i$ is generated by $\mathrm{diag}(a_1, a_2, \ldots, a_n)$ with $a_j = 1, j \neq i$, and $a_i$ equal to a uniformizer of $\mathfrak{p}$. In particular, $\mathfrak{p}$ "splits" in $\mathcal{O}_M$.

1.4. **Quaternion algebras.** We are ready to explain the first mathematical sentence in [4]: "Let $p$ be a rational prime, and let $B$ be *the* quaternion algebra over $\mathbb{Q}$ ramified at the two places $p$ and $\infty$."

**Definition 1.5.** Let $K$ be a field. A central simple algebra $A$ over $K$ is a *quaternion algebra* if $\dim_K A = 4$.

From Theorem 1.2 we have that a quaternion algebra is either isomorphic to $M_2(K)$ or is a division algebra. The next theorem is an immediate consequence of the fundamental theorem 1.3:

**Theorem 1.6.** *If $K$ is a local non-archimedean field or the field of real numbers then there is a unique quaternion algebra $A$ over $K$ which is a division algebra. If $K$ is a global field then $A$ is split at all but finitely many places and the number of places where it is ramified is even. Conversely, any even set of places which does not contain $\mathbb{C}$ arises as the set of ramification places of a unique quaternion algebra over $K$. In particular, $A$ is split at all places of $K$ if and only if it is split over $K$.*

To better understand quaternions we will write down explicitly (in terms of generators and relations) the quaternions over $\mathbb{Q}$ ramified at $p$ and $\infty$.

Assume $\mathrm{char}(K) \neq 2$. For $a, b \in K^\times$, let $H(a,b)$ be the $K$-algebra with basis $1, i, j, k$ (as a $K$-vector space) and relations $i^2 = a, j^2 = b, ij = k = -ji$. With these relations the multiplication on $H(a,b)$ is well-defined.

**Lemma 1.7.** *The algebra $H(a,b)$ is a central simple algebra of dimension 4 over $K$ (i.e., is a quaternion algebra). Any quaternion algebra is isomorphic to $H(a,b)$ for some $a, b \in K^\times$.*

*Proof.* Let $\alpha \in H(a,b)$. Write $\alpha = w + xi + yj + zij$ with $w, x, y, z \in K$. First note that there is a natural involution on $H(a,b)$ given by $\alpha \mapsto \bar{\alpha} = w - xi - yj - zij$, and moreover, $\alpha \cdot \bar{\alpha} = w^2 - ax^2 - by^2 + abz^2$. Next, from the multiplication table of $H(a,b)$ it is clear that $\alpha$ will be in the center if and only if $x, y, z = 0$, i.e., $\alpha \in K$. Hence $H(a,b)$ is central. To prove that it is simple assume $I$ is a non-zero two-sided ideal. We will show $I = H(a,b)$. To do this it is enough to show that some basis element $1, i, j, ij$ is in $I$. Let $\alpha \neq 0 \in I$ with $\alpha = w + xi + yj + zij$. Note that $\frac{1}{2}(\alpha i + i\alpha) = ax + wi \in I$, and similarly $by + wj, abz + wij \in I$. If one of $w, x, y, z$ is 0 then we are done. Otherwise,

$$\begin{aligned} &w\alpha - x(ax + wi) - y(by + wj) - z(abz + wij) \\ &= w^2 - ax^2 - by^2 - abz^2 \in I. \end{aligned}$$

If this last element is non-zero we are done. Otherwise, $\alpha \cdot \bar{\alpha} = w^2 - ax^2 - by^2 + abz^2 \in I$ is non-zero.

Now let $A$ be a 4-dimensional central simple algebra over $K$. We want to show that $A \cong H(a,b)$ for some $a, b \in K^\times$. Let $\alpha \in A$, and assume $\alpha \notin K$. Then $K(\alpha)$ is a field extension of $K$ inside $A$ and hence must be a quadratic extension by Theorem 1.4. Without loss of generality we can assume $\alpha^2 = a \in K^\times$. Take $i := \alpha$. Next, $i \mapsto -i$ is an automorphism of $A$. Hence by Skolem–Noether Theorem [12, (7.21)] there is an invertible $j \in A$ such that $jij = -i$. Hence $ij = -ji$. We claim that $\{1, i, j, ij\}$ are linearly independent over $K$. Indeed, assume $w + xi + yj + zij = 0$. If $y = z = 0$ then we must also have $w = x = 0$ as $w + xi \in K(i)$ which is a field. Otherwise, $j = -\frac{w+xi}{y+zi} \in K(i)$ which is a contradiction as $j$ does not commute with $i$. Finally we must show $j^2 \in K^\times$. Observe that $ij = -ji$ implies $j^2 i = ij^2$, hence $j^2$ is in the center of $A$. Since $A$ is central and $j$ invertible, we must have $j^2 = b \in K^\times$. □

*Example* 1.8. When $K = \mathbb{R}$ then $H(-1,-1)$ is nothing else than the Hamiltonians.

*Example* 1.9. The matrix algebra $M_2(K)$ is isomorphic to $H(1,1)$. Indeed, take $i = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $j = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Then $1, i, j, ij = -ji$ form a basis of $M_2(K)$ and $i^2 = j^2 = 1$.

*Remark* 1.10. It is easy to check that for a given quaternion algebra there are infinitely many $H(a,b)$ isomorphic to it.

The proof of Lemma 1.7 gives a method for constructing quaternions over global fields with prescribed ramification locus. Indeed, given a quaternion algebra $A$, in the proof we introduced an involution $\alpha \mapsto \bar{\alpha}$ with $\alpha \cdot \bar{\alpha} = w^2 - ax^2 - by^2 + abz^2$, where $\alpha = w + xi + yj + zij$. If $w^2 - ax^2 - by^2 + abz^2 = 0$ has no non-zero solutions in $K$, with $a$ and $b$ being fixed, then any $\alpha \neq 0 \in A = H(a,b)$ has an inverse given by $\bar{\alpha}/(\alpha \cdot \bar{\alpha})$. On the contrary, if this equation has a non-zero solution then $A$ is clearly not an integral domain, and hence cannot be a division algebra (with the only remaining possibility being the matrix algebra due to dimension considerations in Theorem 1.2). We summarize this into a lemma:

**Lemma 1.11.** *Let $K$ be a field of characteristic not equal to 2. Then $H(a,b) \cong M_2(K)$ if and only if $X^2 - aY^2 - bZ^2 + abT^2 = 0$ has a non-zero solution in $K$.*

**Proposition 1.12.** *Let $p$ be a prime number. The quaternion algebra over $\mathbb{Q}$ ramified exactly at $p$ and $\infty$ is isomorphic to $H(a,b)$ with the following possibilities for $a, b$:*

1. *$p = 2$, $\{a,b\} = \{-1,-1\}$,*
2. *$p \equiv -1 \bmod 4$, $\{a,b\} = \{-1,-p\}$,*
3. *$p \equiv 5 \bmod 8$, $\{a,b\} = \{-2,-p\}$,*
4. *$p \equiv 1 \bmod 8$, $\{a,b\} = \{-q,-2p\}$, where $q$ is a prime with $\left(\frac{q}{p}\right) = -1$ and $q \equiv 5 \bmod 8$.*

*Proof.* Using Lemma 1.11, we need to find two rational numbers $a$ and $b$ such that $X^2 - aY^2 - bZ^2 + abT^2$ has a non-zero solution over $\mathbb{Q}_\ell$ for $\ell \neq p$, but no such solutions over $\mathbb{R}$ and $\mathbb{Q}_p$. To take care of the restriction over $\mathbb{R}$ it is obviously sufficient (and necessary) to choose $a, b$ to be negative. Since we are dealing with a quadratic equation, to determine whether $X^2 - aY^2 - bZ^2 + abT^2 = 0$ has a solution over $\mathbb{Q}_\ell$ for $\ell \neq 2$ it is enough (by Hensel's lemma) to reduce this equation mod $\ell$ and consider the corresponding question over $\mathbb{F}_\ell$. Since according to Theorem 1.3 the number of places where $H(a,b)$ is ramified must be even, if we have complete information about the ramification of $H(a,b)$ at the odd primes and over $\mathbb{R}$, then the ramification over $\mathbb{Q}_2$ is uniquely determined. Hence we will concentrate on the former cases and use Hensel's lemma. The following simple fact will be useful:

**Lemma 1.13.** *For any finite field $\mathbb{F}_q$ and $a, b, c \in \mathbb{F}_q^\times$, the equation $ay^2 + bz^2 = c$ has at least one non-zero solution.*

*Proof.* For $q$ even this is trivial since any element is a square. Otherwise consider $S = \{ac_1^2 \mid c_1 \in \mathbb{F}_q\}$ and $T = \{c - bc_2^2 \mid c_2 \in \mathbb{F}_q\}$. We have $\#S = \#T = \frac{q+1}{2}$. Hence $\#S + \#T > q$, and $S \cap T \neq \emptyset$. So $ay^2 + bz^2 = c$ has a solution, and it must be non-zero as $c \neq 0$. □

The first three cases immediately follow from this lemma, along with the Hensel's lemma and the quadratic reciprocity law. Namely,

$$\left(\frac{-1}{p}\right)=(-1)^{(p-1)/2},\quad \left(\frac{2}{p}\right)=(-1)^{(p^2-1)/8}.$$

When $p\equiv 1 \bmod 8$ take $a=-q, b=-2p$, where $q$ is a prime with $\left(\frac{q}{p}\right)=-1$ and $q\equiv 5 \bmod 8$ (such primes exist by Dirichlet's density theorem). By Lemma 1.13 $H(a,b)$ will split over $\mathbb{Q}_\ell$ for any $\ell\neq p,q,2$. Over $\mathbb{Q}_p$ our quaternion will ramify as $X^2+qY^2+2pZ^2+2pqT^2=0$ has no solutions mod $p$. Indeed, $\left(\frac{-q}{p}\right)=\left(\frac{-1}{p}\right)\left(\frac{q}{p}\right)=1(-1)=-1$. Finally, mod $q$ the equation $X^2+qY^2+2pZ^2+2pqT^2=0$ has a non-zero solution as

$$\left(\frac{-2p}{q}\right)=\left(\frac{-1}{q}\right)\left(\frac{2}{q}\right)\left(\frac{p}{q}\right)=1(-1)(-1)^{(p-1)(q-1)/4}\left(\frac{q}{p}\right)=1.$$

□

## 2. Ideal theory

By its very definition simple algebras have no two-sided ideals. Nevertheless, similar to the theory of fractional ideals of number fields one introduces a certain notion again called "an ideal", and this leads to a rich arithmetic theory of central simple algebras. The theorems and the proofs in this theory are motivated by the corresponding theorems over number fields, the main technical difficulty being that the rings we are dealing with are usually not commutative.

2.1. **Orders.** Let $R$ be a noetherian integral domain with quotient field $K$ and let $A$ be a central simple $K$-algebra. For any finite dimensional $K$-vector space $V$, a *full $R$-lattice* in $V$ is a finitely generated $R$-submodule $M$ in $V$ such that $K\otimes_R M\cong V$. An *$R$-order* in the $K$-algebra $A$ is a subring $\Lambda$ of $A$, having the same unity element as $A$, and such that $\Lambda$ is a full $R$-lattice in $A$. A *maximal $R$-order* in $A$ is an $R$-order which is not contained in any other $R$-order in $A$.

*Example* 2.1. Let $A=M_n(K)$ be the algebra of $n\times n$ matrices over $K$. Then one easily verifies that $\Lambda=M_n(R)$ is an order in $A$.

*Example* 2.2. Let $M$ be a full $R$-lattice in $A$. (For example, if we choose a basis of $A$, $A=\oplus_{i=1}^n Ke_i$, then $M=\oplus_{i=1}^n Re_i$ is such a lattice.) Define *left order of $M$* to be

$$O_l(M)=\{a\in A\mid aM\subseteq M\}.$$

This is clearly a subring of $A$ with the same unity element, and moreover it is an $R$-order. Indeed, since $K\otimes_R M=A$ we have $K\otimes_R O_l(M)=\{a\in A\mid a(K\otimes M)\subseteq (K\otimes M)\}=A$. It remains to show that $O_l(M)$ is a finitely generated $R$-module. Since $K\otimes M=A$, there are two elements $r_1,r_2\in R$ and $m\in M$ such that $\frac{r_1}{r_2}m=1$. Hence $r_2=r_1m\in M$, and $r_2O_l(M)\subseteq M$. As $M$ is finitely generated $R$-module and $R$ is noetherian, every $R$-submodule of $M$ is also finitely generated (i.e., $M$ is noetherian). This implies $O_l(M)$ is finitely generated as $R$-module; see [7, §3]. Similarly to $O_l(M)$ one also defines the right order $O_r(M)$ of $M$.

This last example shows that orders always exist, and if we assume $R$ is integrally closed then a standard argument shows that every order is contained in a maximal order (in particular, maximal orders exist); see [12, (10.4)].

**Definition 2.3.** Let $\Lambda$ be an $R$-order in $A$. A full $R$-lattice $I$ in $A$ is called a *left ideal of* $\Lambda$ (resp. *right ideal*, *two-sided ideal*) if it is stable under the left multiplication by $\Lambda$ (resp. under the right multiplication, under multiplication on the right and on the left).

2.2. **Norms.** Given an arbitrary finite dimensional $K$-algebra $A$, any $a \in A$ gives a $K$-linear transformation $\varphi_a$ of $A$, regarded as a $K$-vector space, by $\varphi_a(b) = ba$ (this is not an algebra homomorphism). In fact, it not hard to see that $a \mapsto \varphi_a$ defines an injection of $A$ into the algebra of $K$-linear transformations $\mathrm{Hom}_K(A, A)$ of $A$. Consider the characteristic polynomial of $\varphi_a = X^m - c_1 X^{m-1} + \cdots + (-1)^m c_m$, $c_i \in K$, after choosing a basis for $A$, where $m = \dim_K A$. Define $\mathcal{N}_{A/K}(a) = c_m$. Since $\mathcal{N}_{A/K}(a)$ is the determinant of $\varphi_a$ as a linear transformation, it is independent of the choice of basis (in fact, the whole characteristic polynomial is independent of such a choice). The well-known multiplicative properties hold, that is, $\mathcal{N}(ab) = \mathcal{N}(a)\mathcal{N}(b)$, $a, b \in A$, and $\mathcal{N}(ka) = k^m \mathcal{N}(a)$, $k \in K$.

If we assume $A$ is central simple then it is possible to introduce a different notion of a norm. Indeed, by Theorem 1.4 there is a finite separable extension $L$ of $K$ with $\tau : A \otimes_K L \cong M_n(L)$, where $\dim_K A = n^2$. For $a \in A$ define $\mathrm{Nr}_{A/K}(a) := \det(\tau(a \otimes 1))$. This definition is independent of the choice of isomorphism $\tau$. Indeed, any other isomorphism differs from $\tau$ by an automorphism of $M_n(L)$, and hence (by Noether-Skolem) is a conjugate of $\tau$ by an element in $\mathrm{GL}_n(L)$. Since for $\sigma \in \mathrm{Gal}(L/K)$, $\tau \circ \sigma$ is another isomorphism $A \otimes_K L \cong M_n(L)$ we also conclude $\mathrm{Nr}_{A/K}$ takes values in $K$. Moreover, one can show [12, (9.3)] that this definition is independent of the choice of the splitting field $L$. We call Nr the *reduced norm* on $A$. This norm is related to $\mathcal{N}$ introduced previously by the formula $(\mathrm{Nr}_{A/K})^n = \mathcal{N}_{A/K}$, which explains the terminology; see [12, (9.5)].

*Example* 2.4. Consider the quaternions $A = H(a, b)$; c.f. Lemma 1.7. If $A$ is split then the reduced norm is nothing else but the determinant. Now suppose $A = K + Ki + Kj + Kij$ is a division algebra. Then $L := K(i)$ is a maximal subfield in $A$, and hence splits $A$. Let $\alpha \in A$ be written as $\alpha = x + yj$, $x, y \in L$. An isomorphism $\tau : L \otimes A \cong M_2(L)$ is given by

$$\alpha \mapsto \begin{pmatrix} x & y \\ b\bar{y} & \bar{x} \end{pmatrix},$$

where for $x = w + zi$ we denote $\bar{x} = w - zi$. Hence, $\mathrm{Nr}(\alpha) = x\bar{x} - by\bar{y}$. If we write $\alpha = w + zi + sj + tij$ then $x = w + zi$ and $y = s + ti$. One easily verifies

$$\mathrm{Nr}(\alpha) = w^2 - a^2 z^2 - bs^2 + t^2 ab = \alpha\bar{\alpha}.$$

We already extensively used this norm in §1.4.

**Definition 2.5.** With notations as in §2.1, suppose we are given a full $R$-lattice $M$ in the central simple $K$-algebra $A$. Define $\mathrm{Nr}_{A/K}(M)$ the *reduced norm of* $M$ to be the $R$-ideal in $K$ generated by $\{\mathrm{Nr}_{A/K}(a) \mid a \in I\}$. (Here $K$ is regarded as a central simple algebra, $R$ as an order in $K$, and ideal as in Definition 2.3.)

This is well-defined. Indeed, we need to show that $\mathrm{Nr}(M)$ is a finitely generated $R$-module in $K$. It is enough to show that for some $r \in R$, $r\mathrm{Nr}(M) \subset R$, and without loss of generality we can assume $A \cong M_n(K)$ is split. Write $M = Ra_1 + Ra_2 + \cdots + Ra_{n^2}$ for some fixed $a_i \in A$. Then one easily checks that treating the coefficients of $a_i$'s as variables taking values in $R$, the determinant of any element

of $M$ can be expresses as an $R$-value of a fixed polynomial in $n^2$ variables with coefficients in $K$.

2.3. **Orders over Dedekind domains.** To start, we assume $R$ is a complete discrete valuation ring, that is, $R$ is a principal ideal domain with a unique maximal ideal $P = \pi R \neq 0$, and $R$ is complete relative to the $P$-adic valuation. In this situation the ideal theory of central simple algebras over the fraction field $K$ of $R$ is quite simple. First consider a division algebra $D$ over $K$.

**Theorem 2.6.** *The integral closure $\Lambda$ of $R$ in $D$ is the unique maximal order of $D$. ($\Lambda$ also can be characterized as the ring $\mathcal{O}_D$ in §1.3.) Let $\mathfrak{P}$ be the prime ideal of $\mathcal{O}_D$ as in §1.3. Then any one-sided ideal of $\mathcal{O}_D$ is of the form $\mathfrak{P}^m$ and these ideals are necessarily two-sided.*

*Proof.* See [12, (12.8), (13.2)]. □

The case of an arbitrary central simple algebra $A$ over $K$ is handled by the following theorem [12, (17.3)].

**Theorem 2.7.** *Write $A = M_n(D)$ for a division algebra $D$ over $K$ (c.f. Theorem 1.2). Let $\Lambda = M_n(\mathcal{O}_D)$. Then $\Lambda$ is a maximal $R$-order in $A$, and has a unique maximal two-sided ideal $\mathfrak{P}_D\Lambda$. The powers*

$$(\mathfrak{P}_D\Lambda)^m = \mathfrak{P}_D^m\Lambda, \quad m = 0, 1, 2, \dots$$

*give all the non-zero two-sided ideals of $\Lambda$. Every maximal $R$-order in $A$ is a conjugate of $\Lambda$ by an invertible element in $A$, and every such conjugate is a maximal order. Finally, every one-sided ideal of a maximal $R$-order in $A$ is principal.*

The same theorems also hold without assuming $R$ is complete; see [12, §18].

With this last theorem at hand we now can take on the study of ideals in central simple algebras over general Dedekind domains. Just like in the study of number fields, a crucial technique is localization. That is, given a multiplicative set $S$ of a Dedekind domain $R$ and an $R$-order $\Lambda$ in $A$, consider the $S^{-1}R$-order $S^{-1}\Lambda$ in $A$. The following theorem [12, (11.2),(11.6)] in many cases allows to reduce questions about maximal orders to questions about orders over local rings.

**Theorem 2.8.** *Let $R$ be a Dedekind domain, and $\Lambda$ an $R$-order in $A$. Then $\Lambda$ is a maximal order if and only if for each prime ideal $P$ of $R$, $\Lambda_P$ (resp. the $P$-adic completion completion $\hat{\Lambda}_P$) is a maximal $R_P$-order in $A$ (resp. is a maximal $\hat{R}_P$-order in $\hat{A}_P$).*

**Proposition 2.9.** *Let $R$ be a Dedekind domain, and $\Lambda$ a maximal $R$-order in a central simple algebra $A$.*

1. *A left $\Lambda$-ideal $I$ is lattice in $A$ such that for any prime ideal $P$ of $R$, $I_P = \Lambda_P a_P$ for some invertible element $a_P \in A$.*
2. $\mathrm{Nr}_{A/K}(I)_P = \mathrm{Nr}_{A_P/K}(I_P)$.
3. *If $I$ is an integral ideal of $\Lambda$, i.e., $I \subseteq \Lambda$, then*

$$\mathrm{Nr}_{A/K}(I)^n = \prod_{P \in \mathrm{Ass}_R(\Lambda/I)} P^{l_P},$$

*where the product is over the associated primes of the torsion $R$-module $\Lambda/I$, $l_P$ is the length of $R_P$-module $(\Lambda/I)_P$, and $n^2 = \dim_K A$.*

(4) *Let $M$ be any full $R$-lattice in a central simple algebra $A$. Then $O_l(M)$ is maximal if and only if $O_r(M)$ is maximal.*
(5) *For any $\alpha \in \Lambda$, $\mathrm{Nr}(\alpha) \in R$.*

*Proof.*

(1) If $I$ is a left $\Lambda$-ideal then $I_P$ is a left $\Lambda_P$-ideal, hence by Theorem 2.8 and Theorem 2.7 it is principal: $I_P = \Lambda_P a_P$, $a_P \in A$. To show that $a_P$ is invertible, note that since both $I_P$ and $\Lambda_P$ are full $R_P$-lattices in $A$ we have $A = K \otimes I_P = (K \otimes \Lambda_P)a_P = Aa_P$. In particular, there is $b \in A$ such that $1 = ba_P$. Conversely, if $I$ is a lattice such that for every prime $P$, $I_P = \Lambda_P a_P$, then $(\Lambda I)_P = \Lambda_P I_P = I_P$ for every prime $P$. This implies $\Lambda I = I$, that is, $I$ is a left $\Lambda$-ideal. Indeed, localization of the noetherian $R$-module $\Lambda I/I$ at every prime is 0, hence the module itself is zero [7, Thm. 4.6].
(2) This is clear; see also [7, Thm. 4.4].
(3) It is enough to show that both sides are equal to each other after localizing at every prime ideal $P$ of $R$. (Consider both sides as $R$-submodules $N$ and $M$ of $K$. If $(N + M/M)_P = 0$ then by [7, Thm.4.6] $N + M = M$, which implies $N \subseteq M$. A similar argument gives the reverse inclusion.) Hence by Theorem 2.7 we simply can assume $I = \Lambda\lambda$ for some $\lambda \in \Lambda$. Choose an $R$-basis of $\Lambda$ and consider the $R$-linear transformation on $\Lambda$ given by left multiplication by $\lambda$. Using the relation between norms and reduced norms, we need to show
$$\det(\lambda) = \prod_{P \in \mathrm{Ass}_R(\Lambda/\lambda)} P^{l_P}.$$
This becomes clear after diagonalizing $\lambda$ by elementary row and column operations.
(4) Using Theorem 2.8, we can assume $R$ is a DVR. Denote $\mathcal{O} = O_l(M)$ and assume it is a maximal $R$-order in $A$. According to Theorem 2.7, since $M$ is a left $\mathcal{O}$ ideal, $M = \mathcal{O}a$ for some $a \in A$. Moreover, we have shown in (1) that $a$ has to be invertible in $A$. We clearly have $\mathcal{O}_r(M) \supseteq a^{-1}\mathcal{O}a$. But again by Theorem 2.7, $a^{-1}\mathcal{O}a$ is a maximal order, hence so is $\mathcal{O}_r(M)$.
(5) This is clear from (3).

□

*Remark* 2.10. Part (3) of the proposition says that the order of $\Lambda/I$ as a finite group is equal to the idelic norm of the ideal $\mathrm{Nr}(I^n)$.

**Lemma 2.11.** *Let $\Lambda$ be a not necessarily maximal order. Then there is a left $\Lambda$-ideal $I$ such that $O_r(I)$ is maximal.*

*Proof.* Let $I_0$ be any left $\Lambda$-ideal (for example, $\Lambda$ itself). Let $\mathcal{O} = O_r(I)$. There is a maximal order $\Lambda_1$ containing $\mathcal{O}$. It is easy to see that $I_1 = I\Lambda_1$ is again a full $R$-lattice in $A$. This lattice $I_1$ is clearly a left $\Lambda$-ideal and $O_r(I_1) = \Lambda_1$ is maximal. □

Given a full $R$-lattice $M$ in a central simple algebra $A$ define
$$M^{-1} = \{a \in A \mid MaM \subseteq M\}.$$

**Proposition 2.12.** *Let $I$ be a left ideal of some maximal order $\Lambda$ in $A$. Then*

(1) $I^{-1}$ *is a right* $\Lambda$*-ideal, and left* $O_r(I)$*-ideal.*
(2) $II^{-1} = \Lambda$, $I^{-1}I = O_r(I)$, $(I^{-1})^{-1} = I$.
(3) *If* $I$ *is a left* $\Lambda$*-ideal and* $J$ *is a left* $O_r(I)$*-ideal then* $\mathrm{Nr}(IJ) = \mathrm{Nr}(I)\mathrm{Nr}(J)$. *In particular,* $\mathrm{Nr}(I^{-1}) = \mathrm{Nr}(I)^{-1}$.
(4) *If* $J \subset I$ *is another left* $\Lambda$*-ideal and* $\mathrm{Nr}(J) = \mathrm{Nr}(I)$ *then* $J = I$. *In particular, if there is an element* $\alpha \in I$ *such that* $\mathrm{Nr}(\alpha) = \mathrm{Nr}(I)$ *then* $I = \Lambda\alpha$.

*Proof.*

(1) We need to show that $I^{-1}$ is a full $R$-lattice. It is clear that there are $r_1, r_2 \in R$ such that $r_1\Lambda \subseteq I \subseteq r_2\Lambda$ (for example, express the generators of one lattice in terms of the generators of the other with $K$ coefficients and clear the denominators; the required $r$'s are the lcm of these denominators). Then $r_1^{-1}\Lambda \supseteq I^{-1} \supseteq r_2^{-1}\Lambda$. Indeed, since $I^{-1} = \{a \in A \mid Ia \subseteq \Lambda\}$, we have
$$r_1^{-1}\Lambda = \{a \in A \mid r_1\Lambda a \subseteq \Lambda\} \supseteq \{a \in A \mid Ia \subseteq \Lambda\} = I^{-1}.$$
Similarly for the other inclusion. Hence $I^{-1}$ is a finitely generated $R$-module with $K \otimes I^{-1} = A$, i.e., it is a full $R$-lattice. That $I^{-1}$ is preserved under the left multiplication by $O_r(I)$ and right multiplication by $\Lambda$ is clear.
(2) It is enough to prove this after localizing at an arbitrary prime ideal of $R$. But over a DVR the statements are obvious as every ideal is principal and
$$(\Lambda\alpha)^{-1} = \{a \in A \mid \Lambda\alpha a \subseteq \Lambda\} = \alpha^{-1}\Lambda.$$
(3) Denote $\Lambda_1 = O_r(I)$. We know from Proposition 2.9 that $\Lambda_1$ is a maximal order, and clearly $IJ$ is a left $\Lambda$-ideal. Consider the exact sequence of left $\Lambda$-modules
$$0 \to I/IJ \to \Lambda/IJ \to \Lambda/I \to 0.$$
Since localization is exact, to prove the claim we may assume $R$ is a local DVR with prime ideal $P$. On the other hand, $l_P(\Lambda/IJ) = l_P(I/IJ) + l_P(\Lambda/I)$. Hence, using Proposition 2.9, it is enough to show that $I/IJ \cong \Lambda_1/J$. Indeed, $I$ is principal as right $\Lambda_1$-ideal, $I = \alpha\Lambda_1$. Thus,
$$\alpha\Lambda_1/\alpha\Lambda_1 J = \alpha\Lambda_1/\alpha J \cong \Lambda_1/J,$$
where the indicated isomorphism is an $R$-isomorphism.
(4) Without loss of generality we can assume $I$ is integral left $\Lambda$-ideal. Proposition 2.9(3) implies $I/J = 0$ since we assume $\Lambda/I = \Lambda/J$ as $R$-modules. For the last sentence of the proposition, note that $\mathrm{Nr}(\Lambda\alpha) = \mathrm{Nr}(\alpha)$.

□

2.4. **Some finiteness properties.** Let $K$ be a global field, and let $D_K$ be a central division algebra over $K$ of dimension $n^2$. We make an extra assumption that at the places where $D_K$ is ramified $D_v$ is a division algebra (this is automatic if $n = 2$). Let $S$ be a non-empty finite set of places which necessarily include the archimedean ones. Let $R$ be the ring of $S$-integers in $K$. This is a Dedekind domain with fraction field $K$. Let $\Lambda$ be a fixed maximal $R$-order in $D_K$. We know from Theorem 2.8 that $\Lambda_v$ is a maximal order for every place $v \notin S$. If $v$ is a ramification prime then $\Lambda_v$ is uniquely determined, whereas for places where $D$ is split $\Lambda_v$ is a conjugate of $M_n(R_v)$; c.f. Theorem 2.7. Choosing the isomorphisms $D_v \cong M_n(K_v)$ appropriately for split places we can assume that $\Lambda_v$ is in fact equal

to $M_n(R_v)$. Now we give a somewhat ad hoc definition of the *idele* group $D_{\mathbb{A}}^{\times}$ of $D$:

$$D_{\mathbb{A}}^{\times} = \{\tilde{a} = (a_v) \in \prod_v D_v^{\times} \mid a_v \in \Lambda_v^{\times} \text{ for almost all } v\}.$$

It is clear that this definition is independent of the choice of the finite set $S$ used to define $R$. It is not hard to check that for any two maximal $R$-orders $\Lambda$ and $\Gamma$ in $D_K$ there is an equality $\Lambda_v = \Gamma_v$ for almost all $v$. Hence this definition is also independent of the choice of $\Lambda$. (There is a more intrinsic way to define $D_{\mathbb{A}}^{\times}$ as the group of units in the ring of adele valued points $\mathrm{Hom}(\mathbb{A}_K, D)$ of $D$, with $D$ treated as an algebraic group.) There is a natural norm defined on $D_{\mathbb{A}}^{\times}$ as

$$\|\tilde{a}\| = \prod_v |\mathcal{N}(a_v)|_v,$$

where $|\cdot|_v$ is the canonical $v$-adic norm, and the product is over all places of $K$. We will denote by $D_{\mathbb{A}}^{(1)}$ the kernel of $\|\cdot\| : D_{\mathbb{A}}^{\times} \to \mathbb{C}^{\times}$. The group of non-zero elements $D_K^{\times}$ embeds diagonally into $D_{\mathbb{A}}^{(1)}$ thanks to the product formula. Finally define

$$\Lambda_{\mathbb{A}}^{\times} := \{\tilde{a} = (a_v) \in D_{\mathbb{A}}^{(1)} \mid a_v \in \Lambda_v^{\times} \text{ for all } v \notin S\}.$$

**Definition 2.13.** Two left $\Lambda$-ideals $I$ and $J$ are said to be *equivalent* if there is $a \in D_K^{\times}$ with $J = Ia$. The number of equivalence classes of left $\Lambda$-ideals is called the *class number of* $D_K$ (with respect to $S$).

**Proposition 2.14.** *With $S$ being fixed, the class number of $D_K$ is finite and independent of the choice of maximal order $\Lambda$.*

*Proof.* Every ideal $I$ is principal at all the places where $D$ is split, i.e., $I_v = \Lambda_v a_v$ with $a_v \in D_v^{\times}$. Moreover, for almost all places $I_v = \Lambda_v$, that is, $a_v \in \Lambda_v^{\times}$, and we conclude that there is an element $\tilde{b} \in D_{\mathbb{A}}^{(1)}$ with $b_v = a_v$ for almost all $v$. Hence $D_{\mathbb{A}}^{(1)}$ acts from the left transitively on the set of equivalence classes of ideals. The action of $\Lambda_{\mathbb{A}}^{\times}$ is clearly trivial, and so is the action of $D_K$ (by definition). Hence the class number is less or equal to the number of double cosets in $\Lambda_{\mathbb{A}}^{\times} \setminus D_{\mathbb{A}}^{(1)} / D_K^{\times}$. By a standard argument $D_{\mathbb{A}}^{(1)}/D_K^{\times}$ is compact and $\Lambda_{\mathbb{A}}^{\times}$ is open in $D_{\mathbb{A}}^{(1)}$. So the class number is finite. In fact, it is not hard to show that the class number is equal to the number of double cosets. As for the independence of class number from a particular choice of $\Lambda$, one observes that any two maximal $R$-orders are conjugate by an element of $D_{\mathbb{A}}^{(1)}$. The independence now is clear. □

**Proposition 2.15.** *Let $\Lambda$ be a fixed maximal $R$-order, and let $\{I_1, I_2, \ldots, I_n\}$ be a set of left $\Lambda$-ideals representing the distinct ideal classes. Then each conjugacy class of maximal $R$-orders in $D_K$ is represented in the set of right orders $\{O_r(I_1), O_r(I_2), \ldots, O_r(I_n)\}$.*

*Proof.* Let $\Gamma$ be a maximal order in $D_K$. Since $\Gamma_v$ and $\Lambda_v$ are maximal $R_v$-orders for all $v \notin S$, there is $\alpha_v \in D_v^{\times}$ for each such $v$ with $\alpha_v^{-1}\Lambda_v\alpha_v = \Gamma_v$. Moreover, $\alpha_v \in \Lambda_v^{\times}$ for almost all $v$. Hence there is $\tilde{\alpha} \in D_{\mathbb{A}}^{(1)}$ such that $\tilde{\alpha}_v = \alpha_v$ for all $v \notin S$. Let $J = \Lambda\tilde{\alpha}$. This is a left $\Lambda$-ideal with $\Gamma$ as its right order. As $\{I_1, I_2, \ldots, I_n\}$ represent all the left $\Lambda$-ideal classes, there is $a \in D_K^{\times}$ and some $I_i$ with $J = I_i a$. We clearly have $O_r(I_i) = a\Gamma a^{-1}$. □

**Lemma 2.16.** *Assume the group of $S$-units is finite. Then for any maximal $R$-order $\Lambda$ in $D_K$ the group $\Lambda^\times$ is finite.*

*Proof.* Let $Z$ be the center of $D$ as an algebraic group. Then $G = D^\times/Z^\times$ as an algebraic variety over $K$ is projective. Let $v$ be a place in $S$. Consider $G(K_v)$. This is compact in $v$-adic topology, and contains $\Lambda^\times/R^\times$ as a discrete subgroup. Hence $\Lambda^\times/R^\times$ is finite. □

## 3. Endomorphism rings of super-singular elliptic curves

In this section we study super-singular (s.s.) elliptic curves over fields of positive characteristic and their endomorphism rings. These latter rings are very closely related to quaternion algebras.

3.1. **The structure of End(E).** Let $p$ be a prime number. We will denote by $\mathbb{F}_p$ the finite field of $p$ elements ($= \mathbb{Z}/p\mathbb{Z}$), and will denote by $\overline{\mathbb{F}}_p$ the algebraic closure of $\mathbb{F}_p$. Let $E$ be s.s. elliptic curve over $\overline{\mathbb{F}}_p$. Recall that such curves are characterized by the property that their $p$ torsion is connected as a group-scheme (or equivalently, the map $[p] : E \to E$ is purely inseparable); see [15, §V.3]. Let $\mathrm{End}(E)$ be the endomorphism ring of $E$. It is known that $\mathrm{End}(E)$ is a free $\mathbb{Z}$-module of rank at most 4 [15, §III.7], and in fact, it is an order in a definite quaternion algebra over $\mathbb{Q}$, so $\mathrm{rank}_{\mathbb{Z}}(\mathrm{End}(E)) = 4$, [15, §V.3]. (This last property also uniquely characterizes s.s. elliptic curves.)

Consider $\mathrm{End}^0(E) := \mathrm{End}(E) \otimes_{\mathbb{Z}} \mathbb{Q}$. From what was said, we know that $\mathrm{End}^0(E)$ is a quaternion algebra over $\mathbb{Q}$, and we intend to describe this algebra. As we have already discussed §1.4, to uniquely characterize a quaternion algebra it is enough to specify the places where it is ramified. A priori, $\mathrm{End}^0(E)$ is ramified over $\mathbb{R}$ (since it is definite). So this algebra must be ramified at least over some finite prime (the number of ramified places, including the archimedean one, must be even). We will show that for any prime $\ell \neq p$, $\mathrm{End}^0(E) \otimes \mathbb{Q}_\ell \cong M_2(\mathbb{Q}_\ell)$. This implies

$$\mathrm{End}^0(E) \cong B_p,$$

where $B_p$ is the unique quaternion algebra over $\mathbb{Q}$ ramified exactly at $p$ and $\infty$, c.f. Proposition 1.12.

**Lemma 3.1.** *With notations as above, for any $\ell \neq p$ we have an isomorphism $\mathrm{End}^0(E) \otimes \mathbb{Q}_\ell \cong M_2(\mathbb{Q}_\ell)$.*

*Proof.* Denote $V_\ell(E) = T_\ell(E) \otimes_{\mathbb{Z}_\ell} \mathbb{Q}_\ell$, where $T_\ell(E) := \varprojlim E[\ell^n] \cong \mathbb{Z}_\ell \times \mathbb{Z}_\ell$ is the Tate module of $E$. Since $\mathbb{Q}_\ell$ is flat over $\mathbb{Z}_\ell$, by [15, Thm.III.7.4] there is a natural injection

$$\mathrm{End}^0(E) \otimes \mathbb{Q}_\ell \hookrightarrow \mathrm{Hom}_{\mathbb{Q}_\ell}(V_\ell(E)) \cong M_2(\mathbb{Q}_\ell).$$

But both sides have dimension 4 as vector spaces over $\mathbb{Q}_\ell$, hence the injection must be an isomorphism. □

**Theorem 3.2.** *The ring $\Lambda = \mathrm{End}(E)$ is a maximal order in $B_p$.*

*Remark* 3.3. There are (at least) two different proofs of this fact. First we give the "modern" proof which uses some results of Tate from 60's. The theorem was originally proven by Deuring in 40's. We will present Deuring's proof in the next subsection; see Theorem 3.7.

*Proof.* Let $\ell \neq p$ be a prime number. Let $k$ be a finite extension of $\mathbb{F}_p$ which is large enough so that both $E$ and all endomorphisms of $E/\overline{\mathbb{F}}_p$ are defined over $k$. Then according to Tate's theorem the natural map

$$\Lambda \otimes \mathbb{Z}_\ell \to \mathrm{End}_{\mathbb{Z}_\ell[\mathrm{Gal}(\bar{k}/k)]}(T_\ell(E))$$

is an *isomorphism*. Since $\mathrm{rank}_{\mathbb{Z}_\ell}(\Lambda \otimes \mathbb{Z}_\ell) = 4$, and $\mathbb{Z}_\ell$-rank of the right-hand side is at most 4, we easily see that the Galois group must acts through scalars on $T_\ell(E)$. Hence

$$\Lambda \otimes \mathbb{Z}_\ell \cong \mathrm{End}_{\mathbb{Z}_\ell}(\mathbb{Z}_\ell \times \mathbb{Z}_\ell) \cong M_2(\mathbb{Z}_\ell).$$

There is a version of Tate's theorem for $p$, from which one deduces

$$\Lambda \otimes \mathbb{Z}_p \xrightarrow{\sim} \mathrm{End}(\widehat{E}_{/\overline{\mathbb{F}}_p}),$$

where $\widehat{E}$ is the height 2 and dimension 1 commutative formal group over $\overline{\mathbb{F}}_p$ associated to $E$ (the height is 2 because $E$ is s.s.). It is known that the endomorphism ring of such a formal group is a maximal order in the unique division algebra over $\mathbb{Q}_p$ of dimension 4. (Brian Conrad tells me that this last fact can be found in Hazewinkel's book.)

We proved that the localization of $\Lambda$ at any prime is a maximal order. Hence by Theorem 2.8, $\Lambda$ is a maximal order in $B_p$. □

3.2. **Ideal theory and arithmetic.** The purpose of this subsection is to relate the ideal theory of the quaternion algebra $B_p$ to the arithmetic of s.s. elliptic curves over $\overline{\mathbb{F}}_p$. After that, one is able to rephrase most of the questions about ideals and maximal orders in $B_p$ as questions about s.s. elliptic curves and attack them using algebra-geometric methods.

Fix a s.s. elliptic curve $E$ over $\overline{\mathbb{F}}_p$, and denote $\Lambda = \mathrm{End}(E)$. Let $I$ be a left (resp. right) integral $\Lambda$-ideal (i.e., $I$ is a full $\mathbb{Z}$-lattice in $B_p$, $\Lambda I \subseteq I$ (resp. $I\Lambda \subseteq I$) and $I \subseteq \Lambda$). To such an ideal one can associate another elliptic curve $E^{(I)}$ and an isogeny $\varphi_I : E \to E^{(I)}$ (because $E$ and $E^{(I)}$ are isogenous, $E^{(I)}$ is necessarily s.s.). Indeed, consider the group-scheme

$$C_I = \bigcap_{i \in I} \ker(i : E \to E),$$

with the scheme-theoretic intersection taken inside of $E$. This is easily seen to be a finite subgroup-scheme of $E$, hence gives an isogeny $\varphi_I : E \to E/C_I$. We take $E^{(I)} = E/C_I$.

For a finite group(-scheme) $M$ we will denote by $\#M$ its order (its order as a group scheme). Given an isogeny $\varphi : A \to B$ of abelian varieties over a field, its *degree* $\deg(\varphi)$ as a finite flat map is equal to the order of the finite flat group-scheme $\ker(\varphi)$. For an integral left $\Lambda$-ideal $I$ we will denote $\ker(I) := C_I$, and $\deg(I) := \deg(\varphi_I) = \#C_I$.

**Lemma 3.4.** *With previous notations, assume $\Lambda$ is a maximal order. Then we have*

$$\deg(I) = \#\mathbb{Z}/\mathrm{Nr}(I).$$

*Proof.* Let $\alpha \in \Lambda$. Recall that in the proof of the fact that $\Lambda$ is an order in a quaternion algebra one shows that $\mathrm{Nr}(\alpha) = \alpha \circ \hat{\alpha}$, where $\hat{\alpha}$ is the dual isogeny of $\alpha$; c.f. [15, §III.9]. On the other hand, $\alpha \circ \hat{\alpha} = \deg(\alpha)$. Since $\mathrm{Nr}(\Lambda\alpha) = \mathbb{Z}\mathrm{Nr}(\alpha)$, the claim follows for principal ideals.

Now suppose $I$ is (not necessarily principal) left integral ideal. Let $J$ be a right integral $\Lambda$-ideal. One easily checks that $\deg(IJ) \leq \deg(I)\deg(J)$. Indeed, note that

$$\ker(IJ) \subseteq \bigcap_{i\in I, j\in J} \ker(ij).$$

Next, for any $I$ we clearly have $\deg(I) \leq \text{g.c.d.}(\#\ker(i))_{i\in I}$. Thus $\deg(I) \leq \#\mathbb{Z}/\mathrm{Nr}(I)$, since $\mathbb{Z}$-ideal $\{\#\ker(i) \mid i \in I\} = (\text{g.c.d.}(\#\ker(i))_{i\in I})\mathbb{Z}$ is equal to $\{\mathrm{Nr}(i) \mid i \in I\} = \mathrm{Nr}(I)$. Now assume $J$ is such that $IJ$ is principal. Then

$$\begin{aligned}\deg(IJ) &\leq \deg(I)\deg(J)\\ &\leq \#\mathbb{Z}/\mathrm{Nr}(I)\mathrm{Nr}(J)\\ &= \#\mathbb{Z}/\mathrm{Nr}(IJ) \quad \text{(by Proposition 2.12).}\end{aligned}$$

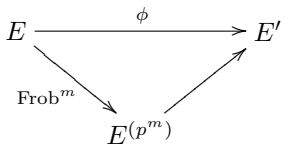
Since $IJ$ is principal, from what we already proved, $\deg(IJ) = \#\mathbb{Z}/\mathrm{Nr}(IJ)$. Hence equality holds through out, and $\deg(I) = \#\mathbb{Z}/\mathrm{Nr}(I)$ as desired.

It remains to show that there exists a right integral $\Lambda$-ideal $J$ with $IJ$ principal. For this we can take an appropriate scalar multiple of $I^{-1}$, so that $nI^{-1} \subseteq \Lambda$. (Geometrically all we are doing is taking the dual isogeny to $\varphi_I$.) □

**Proposition 3.5.** *With previous notations, assume $\Lambda$ is a maximal order. Then every isogeny from $E$ to another elliptic curve has the form $\varphi_I$ for some left integral ideal $I$.*

*Proof.* Suppose $I$ and $J$ are two integral left ideals, and $\ker(I) = \ker(J)$. It is easy to check that $\ker(I+J) = \ker(I)\cap\ker(J) = \ker(J)$. Hence by Lemma 3.4 we must have $\mathrm{Nr}(I+J) = \mathrm{Nr}(J)$. Since $I + J \supseteq J$, we know from Proposition 2.12 that the equality of the norms implies $I + J = J$, i.e., $I \subseteq J$. Similarly one proves the converse inclusion. The conclusion is that distinct ideals define distinct isogenies.

Since every isogeny, up to an automorphism, is uniquely determined by its kernel, it is enough to show that the number of order $n$ subgroup-schemes of $E$ is equal to the number of left integral ideals of reduced norm $n\mathbb{Z}$. It is easy to check that one can assume $n$ is a prime power. We distinguish two cases. First assume $n = p^m$. Since $E$ is s.s. every isogeny $\phi : E \to E'$ of order $p^m$, being purely inseparable, must factor through $E^{(p^m)}$




where $E^{(p^m)}$ is the image of $E$ by the $p^m$-power Frobenius map. Comparing the degrees, we get $E' \cong E^{(p^m)}$. Hence there is a unique subgroup-scheme of order $p^m$ in $E[p^m]$ (the kernel of $\mathrm{Frob}^m$). We know that to count the number of ideals of reduced norm $p^m$ we may localize at $p$. But $\Lambda_p$ is a maximal ideal in a division algebra, and hence has a unique ideal of reduced norm $p^m$; see [12, (24.13)].

Now assume $n = \ell^m$, where $\ell \neq p$. Again localizing at $\ell$, we need to show that modulo $\mathrm{GL}_2(\mathbb{Z}_\ell)$ the number of elements in $M_2(\mathbb{Z}_\ell)$ of determinant $\ell^m$ is equal to the number of subgroups of $\mathbb{Z}/\ell^m\times\mathbb{Z}/\ell^m$ of order $\ell^m$ (we are using here the fact that $B_p$ is split at $\ell$, $\Lambda$ is a conjugate of $M_2(\mathbb{Z}_\ell)$, and left integral $\Lambda$-ideals are principal). There is a one-to-one correspondence between the subgroups of $\mathbb{Z}/\ell^m \times \mathbb{Z}/\ell^m$ of

order $\ell^m$ and $\mathbb{Z}_\ell$-submodules of $\mathbb{Z}_\ell \times \mathbb{Z}_\ell$ of cokernel of order $\ell^m$. Indeed, as $\mathbb{Z}_\ell^2/\Gamma$ is killed by $\ell^m$, $\ell^m\mathbb{Z}_\ell^2 \subset \Gamma$. Next, the image $S$ of $\Gamma$ under the homomorphism

$$0 \to \ell^m\mathbb{Z}_\ell^2 \to \mathbb{Z}_\ell^2 \to (\mathbb{Z}/\ell^m)^2 \to 0$$

is a subgroup of order $\ell^m$. Conversely, for any subgroup $S$ of order $\ell^m$ in $(\mathbb{Z}/\ell^m)^2$, the preimage of $S$ in $\mathbb{Z}_\ell^2$ is a submodule with the desired property. One easily checks that there is a one-to-one correspondence between the submodules $\Gamma$ of $\mathbb{Z}_\ell^2$ of index $\ell^m$ and elements modulo $\mathrm{GL}_2(\mathbb{Z}_\ell)$ in $M_2(\mathbb{Z}_\ell)$ of determinant $\ell^m$. Indeed, take the matrix with $\mathbb{Z}_\ell$ coefficients which transforms the lattice $\mathbb{Z}_\ell^2$ in the $\mathbb{Q}_\ell$-vector space $\mathbb{Q}_\ell \otimes (\mathbb{Z}_\ell^2)$ into the lattice $\Gamma$. □

**Lemma 3.6.** *With previous notations let $I$ be a left integral $\Lambda$-ideal (here we do not assume $\Lambda$ is maximal). Then there is an injective ring homomorphism $O_r(I) \hookrightarrow \mathrm{End}(E^{(I)})$.*

*Proof.* Let $n$ be a large enough natural number so that $nO_r(I) \subseteq \Lambda$, and let $\gamma \in O_r(I)$. Consider the composite $\varphi_I \circ n\gamma$. Since $In\gamma \subseteq nI$, $\ker(\varphi_I \circ n\gamma) \supseteq \ker(nI)$. Hence $\varphi_I \circ n\gamma$ factors through $n\varphi_I$, and $\gamma$ induces an endomorphism of $E^{(I)}$. This clearly defines a homomorphism $O_r(I) \to \mathrm{End}(E^{(I)})$, which is easy to check is injective. □

**Theorem 3.7.** *If $E$ is a s.s. elliptic curve over $\overline{\mathbb{F}}_p$ then $\Lambda = \mathrm{End}(E)$ is a maximal order in $B_p$.*

*Proof.* By Lemma 2.11, there is a left $\Lambda$-ideal (which we can assume is integral after multiplying by an appropriately large natural number) such that $\Lambda_1 = O_r(I)$ is maximal. By Lemma 3.6, $\Lambda_1 \hookrightarrow \mathrm{End}(E^{(I)})$, so $\mathrm{End}(E^{(I)}) = \Lambda_1$ is maximal. We know that $E^{(I)}$ is isogenous to $E$, for example, under the dual isogeny of $\varphi_I$. On the other hand, we know from Proposition 3.5 that such an isogeny must be of the form $\varphi_J$ for some left integral $\Lambda_1$-ideal $J$. Again applying Lemma 3.6, this time to $\varphi_J$, we get $O_r(J) \hookrightarrow \Lambda$. Finally, from Proposition 2.9 we know that $O_r(J)$ is maximal since $O_l(J) = \Lambda_1$ is maximal. Hence $\Lambda$ must be a maximal order. □

**Theorem 3.8.** *There is a one-to-one correspondence between the isomorphism classes of s.s. elliptic curves over $\overline{\mathbb{F}}_p$ and the left ideal classes in $B_p$.*

*Proof.* Recall that the left ideal classes of $B_p$ are defined as the equivalence classes of left ideals of some fixed maximal order $\Lambda$ in $B_p$, where two left $\Lambda$-ideals $I$ and $J$ are equivalent if there is $a \in B_p^\times$ with $J = Ia$. (One shows that the number of equivalence classes is independent of the choice of $\Lambda$.)

Fix a s.s. elliptic curve $E$ over $\overline{\mathbb{F}}_p$ and let $\Lambda = \mathrm{End}(E)$. We know from Theorem 3.7 that $\Lambda$ is a maximal order in $B_p$. Let $I$ and $J$ be two left $\Lambda$-ideals which are equivalent. After scaling $I$ and $J$ by appropriately large natural numbers, we can assume $I$ and $J$ are integral and there is $a \neq 0 \in \Lambda$ with $J = Ia$ (note that $I$ and $nI$ are in the same equivalence class for any $n \neq 0 \in \mathbb{Z}$). We claim that $E^{(I)} \cong E^{(J)}$. Indeed, $\varphi_J = \varphi_I \circ a$. Thus there is an injective map from the set of ideal classes into the set of isomorphism classes of s.s. elliptic curves. On the other hand, any two s.s. elliptic curves over $\overline{\mathbb{F}}_p$ are isogenous, for example, by Tate's theorem [17, Thm.2]. Using this fact and Proposition 3.5, the map we constructed is also surjective. □

**Corollary 3.9.** *Let $\{E_1, E_2, \cdots, E_n\}$ be a set of s.s. elliptic curves over $\overline{\mathbb{F}}_p$ representing the distinct isomorphism classes of such curves. Then every conjugacy class of maximal orders in $B_p$ is represented in the set $\{\mathrm{End}(E_1), \ldots, \mathrm{End}(E_n)\}$.*

*Proof.* Let $\Lambda = \mathrm{End}(E_1)$ and let $\{I_1, I_2, \ldots, I_n\}$ be a set of left integral $\Lambda$-ideals representing all distinct ideal classes with $I_1 = \Lambda$. By Lemma 3.6 and Theorem 3.8 we know that $\{O_r(I_1), \ldots, O_r(I_n)\}$ and $\{\mathrm{End}(E_1), \ldots, \mathrm{End}(E_n)\}$ represent the same set of conjugacy classes of maximal orders in $B_p$. Hence the corollary follows form Proposition 2.15. □

**Proposition 3.10.** *Let $E_1$ and $E_2$ be two s.s. elliptic curves over $\overline{\mathbb{F}}_p$, well-defined up to an isomorphism. Then $\mathrm{End}(E_1)$ and $\mathrm{End}(E_2)$ are conjugate as maximal orders in $B_p$ if and only if $E_1$ and $E_2$ are conjugate by an automorphism of the field $\overline{\mathbb{F}}_p$.*

*Proof.* Denote $\Lambda_i := \mathrm{End}(E_i)$, $i = 1, 2$. Suppose $E_1$ and $E_2$ are conjugate by an element $\sigma$ of $\mathrm{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_p)$. Applying $\sigma$ to the coefficients of $\psi \in \Lambda_1$, we get an isomorphism $\Lambda_1 \cong \Lambda_2$. Tensoring this isomorphism with $\mathbb{Q}$, we get an automorphism of $B_p$. Since all the automorphisms of a central simple algebra are inner (by Skolem-Noether theorem) we get that $\Lambda_1$ and $\Lambda_2$ are conjugate maximal orders in $B_p$.

Conversely, suppose $\alpha\Lambda_1\alpha^{-1} = \Lambda_2$ for some $\alpha \in B_p^\times$. As in the proof of Theorem 3.8, there is a left integral $\Lambda_1$-ideal $I$ such that $\varphi_I(E_1) = E_2$ and $O_r(I) = \Lambda_2$. Since $O_r(I\alpha) = \alpha^{-1}O_r(I)\alpha$, $O_l(I\alpha) = O_l(I)$, and $I\alpha$ is in the same ideal class as $I$, possibly replacing $E_2$ by an isomorphic elliptic curve, we can assume $O_l(I) = O_r(I)$ (since $\varphi_I \cong \varphi_{I\alpha}$), that is, $I$ is a two-sided integral $\Lambda_1$-ideal. Factor $\varphi_I = \varphi_s \circ \varphi_i$ into a composite of a purely inseparable morphism and a separable one; c.f. [15, §II.2]. The morphism $\varphi_i$ is given by a power of the Frobenius map $x \mapsto x^p$, which can be considered as an element of $\mathrm{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_p)$. Hence if we care only about Galois conjugacy classes of elliptic curves, we can assume $\varphi_I$ is separable. Since $E_1$ is s.s. for $\varphi_I$ to be separable, it is necessary and sufficient for $\deg(\varphi_I)$ to be coprime to $p$. On the other hand, by Lemma 3.4 $\deg(\varphi_I) = \#\mathbb{Z}/\mathrm{Nr}(I)$, so the generator $n$ of the ideal $\mathrm{Nr}(I)$ must be coprime to $p$. In view of the fact that $B_p$ is split at every prime $\ell \neq p$, the structure theorems of two-sided ideals of maximal orders [12, (22.10), (22.14)] imply $I$ is principal and is equal to $n\Lambda_1$. Hence $\varphi_I \in \mathrm{End}(E_1)$ and $E_1 = E_2$. This finishes the proof. □

Let $\Lambda$ be a fixed maximal order in $B_p$, $\{I_1, \ldots, I_n\}$ be representatives of distinct left ideal classes of $\Lambda$ with $I_1 = \Lambda$, and $\Lambda_i = O_r(I_i)$. From Corollary 3.9 and its proof we know that there exist s.s. elliptic curves $E_1, \ldots, E_n$ with $E_i \cong E_1^{(I_i)}$ and $\mathrm{End}(E_i) = \Lambda_i$. Denote $M_{ij} = I_j^{-1}I_i$. Since $O_r(I_j^{-1}) = \Lambda$, this is a lattice in $B_p$, and it is naturally a left $\Lambda_j$ and right $\Lambda_i$ module.

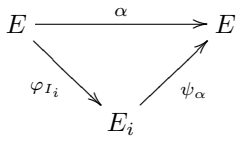
**Proposition 3.11.** *With previous notations, there is an isomorphism*

$$M_{ij} \cong \mathrm{Hom}(E_i, E_j)$$

*as left $\Lambda_j$ and right $\Lambda_i$ modules.*

*Proof.* Without loss of generality we can assume the ideals $I_i$, $i = 1, \ldots, n$, are integral. Let $\alpha \in I_i$. Consider the endomorphism of $E := E_1$ induced by $\alpha$. Since

$\ker(\alpha) \supseteq \ker(I_i)$, $\alpha : E \to E$ must factor through $\varphi_{I_i}$




where $\psi_\alpha \in \mathrm{Hom}(E_i, E)$. One easily checks that $\alpha \mapsto \psi_\alpha$ defines an injection $I_i \hookrightarrow \mathrm{Hom}(E_i, E)$ of left $\Lambda$ and right $\Lambda_i$ modules. Now let $\psi \in \mathrm{Hom}(E_i, E)$. The composite $\psi \circ \varphi_{I_i}$ is an endomorphism $\gamma$ of $E$. We claim that $\gamma \in I_i$. Indeed, consider the left integral $\Lambda$-ideal generated by $\gamma$ and $I_i$, $\Lambda\gamma + I_i$. We have $\ker(\Lambda\gamma + I_i) = \ker(\gamma) \cap \ker(I_i)$. Since by construction $\ker(\gamma)$ contains $\ker(I_i)$, we get $\ker(\Lambda\gamma + I_i) = \ker(I_i)$. Hence by Lemma 3.4 and Proposition 2.12, $\gamma + I_i = I_i$, i.e., $\gamma \in I_i$ as required. We conclude that the map $\alpha \mapsto \psi_\alpha$ constructed above is also surjective, in particular

$$I_i \cong \mathrm{Hom}(E_i, E). \tag{3.1}$$

Let $I$ be a left integral $\Lambda_j$-ideal. The lattice $I_j I$ is a left $\Lambda$-ideal. We claim that

$$E^{(I_j I)} \cong E_j^{(I)}. \tag{3.2}$$

As usual, we can assume that after scaling, $I_j I$ is an integral $\Lambda$-ideal. Since $O_r(I_j) = \Lambda_j$ and $I \subseteq \Lambda_j$, we have $I_j I \subseteq I_j$. Hence $\ker(I_j I) \supseteq \ker(I_j)$, and

$$\ker(I_j I) \bmod \ker(I_j) \cong \ker(I).$$

We conclude $\varphi_I \circ \varphi_{I_j} = \varphi_{I_j I}$, which is equivalent to the claim.

Now apply (3.1) with $E$ replaced by $E_j$ and $E_i$ replaced by $E_j^{(M_{ij})}$. (Here we consider $M_{ij}$ as a left $\Lambda_j$-ideal.) We have

$$\begin{aligned} M_{ij} &\cong \mathrm{Hom}(E_j^{(M_{ij})}, E_j) \quad \text{(by (3.1))} \\ &\cong \mathrm{Hom}(E^{(I_j M_{ij})}, E_j) \quad \text{(by (3.2))} \\ &\cong \mathrm{Hom}(E^{(I_i)}, E_j) \quad \text{(by Proposition 2.12)} \\ &\cong \mathrm{Hom}(E_i, E_j), \end{aligned}$$

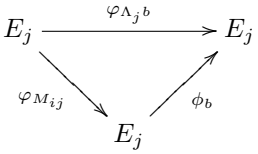
where all the isomorphisms are isomorphisms of left $\Lambda_j$ and right $\Lambda_i$ ideals. □

**Proposition 3.12.** *With notations as in Proposition 3.11, the degree of of an isogeny $\phi_b : E_i \to E_j$ corresponding to a non-zero element $b \in M_{ij}$ is given by the formula*

$$\deg(\phi_b) = \mathrm{Nr}(b)/m_{ij},$$

*where $m_{ij}$ is the generator of the fractional ideal $\mathrm{Nr}(M_{ij})$.*

*Proof.* Consider the left principal ideal $\Lambda_j b$. Then by the construction implicit in the proof of Proposition 3.11, $\phi_b$ is the isogeny for which $\varphi_{\Lambda_j b} = \phi_b \circ \varphi_{M_{ij}}$



From this it is easy to see that $\phi_b$ is invariant under scaling $M_{ij}$ by integers, so we can assume $M_{ij}$ is an integral left $\Lambda_j$-ideal. Moreover,

$$\deg(\phi_b) = \deg(\varphi_{\Lambda_j b})/\deg(\varphi_{M_{ij}}).$$

The claim follows from Lemma 3.4 and Proposition 2.12. Note that the expression $\mathrm{Nr}(b)/m_{ij}$ is invariant under scaling $M_{ij}$ by integers, since $\mathrm{Nr}(nM_{ij}) = \mathrm{Nr}(n)\mathrm{Nr}(M_{ij})$ and $\mathrm{Nr}(nb) = \mathrm{Nr}(n)\mathrm{Nr}(b)$. □

3.3. **Brandt matrices.** Let $\Lambda$ be a fixed maximal order in $B_p$. Let $\{I_1, I_2, \dots, I_n\}$ be a set of left $\Lambda$-ideals representing the distinct ideal classes, with $I_1 = \Lambda$. For $1 \leq i \leq n$ we let $O_r(I_i) = \Lambda_i$, and let $w_i$ be the order of the finite group of invertible elements $\Lambda_i^\times/(\pm 1)$ in $\Lambda_i$ modulo the subgroup generated by $\pm 1$; c.f. Lemma 2.16. Note that geometrically $2w_i = \#\mathrm{Aut}(E_i)$. Let

$$B_{ij}(m) = \frac{1}{2w_j}\#\{b \in M_{ij} \mid \mathrm{Nr}(b)/m_{ij} = m\}.$$

**Definition 3.13.** The *Brandt matrices* $B(m)$, $m = 0, 1, 2, \dots$ are the $n \times n$ matrices

$$B(m) = (B_{ij}(m))_{1 \leq i,j \leq n}.$$

Thanks to Proposition 3.12 we can study Brandt matrices using the geometry of s.s. elliptic curves.

**Proposition 3.14.** *With notations as in Corollary 3.9, the entry $B_{ij}(m)$ is equal to the number of subgroup-schemes $C$ of order $m$ in $E_i$ such that $E_i/C \cong E_j$.*

*Proof.* [4, Prop.2.3]. □

**Proposition 3.15.** *With previous notations, the curves $E_i$ and $E_j$ are conjugate by an automorphism of $\overline{\mathbb{F}}_p$ if and only if $i = j$ or $B_{ij}(p) = 1$. Moreover, there is an equality*

$$\#\{E_i \mid 1 \leq i \leq n, j(E_i) \in \mathbb{F}_p\} = \mathrm{Trace}(B(p)).$$

*Proof.* It is well-known that if $E$ is a s.s. elliptic curve over $\overline{\mathbb{F}}_p$ then, in fact, $E$ can be defined over the quadratic extension of $\mathbb{F}_p$, i.e., the $j$-invariant $j(E)$ of $E$ lies in $\mathbb{F}_{p^2}$. Indeed, we took as a definition of s.s. to be "$E[p^m]$ is connected". This easily implies that the dual of $\mathrm{Frob} : E \to E^{(p)}$ is the Frobenius map on $E^{(p)}$, that is, $E^{(p^2)} \cong E$ (c.f. proof of Proposition 3.5). As $j(E^{(p)}) = j(E)^p$, we have $j(E)^{p^2} = j(E)$. Hence $j(E) \in \mathbb{F}_{p^2}$. The conjugation of s.s. elliptic curve $E$ by the topological generator $x \mapsto x^p$ of $\mathrm{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_p)$ has the same effect as the Frobenius morphism. Hence $E_i$ and $E_j$ are conjugate by an automorphism of $\overline{\mathbb{F}}_p$ if and only if $i = j$ or $E_i^{(p)} \cong E_j$. Now the claim follows from Proposition 3.14, as the kernel of the Frobenius morphism is the unique order $p$ subgroup-scheme of a s.s. elliptic curve over $\overline{\mathbb{F}}_p$. □

**Proposition 3.16.** *The Brandt matrices $B(m)$ for $m \geq 1$ generate a commutative subring $\mathbb{B}$ of $M_n(\mathbb{Z})$. Moreover, the commutative algebra $\mathbb{B} \otimes_{\mathbb{Z}} \mathbb{Q}$ is semi-simple, and is isomorphic to a product of totally real number fields.*

*Proof.* The first part is proven in [4, Prop.2.7]. We would like to clarify the last statement.

First recall the definition of semi-simple rings (algebras). There are several (equivalent) ways to define this notion, c.f. [6, XVII.4]. Let $K$ be a field. We will

say that a finite dimensional $K$-algebra $A$ is *semi-simple* if it is isomorphic to a direct product of simple $K$-algebras, where simple $K$-algebra is as defined in §1.1. When $A$ is commutative this is equivalent to the definition given in [6], thanks to XVII.4.3 in *loc.cit.* We will need the following well-known fact whose proof we recall for the sake of completeness:

**Lemma 3.17** (Spectral Theorem)**.** *Let $V$ be a finite-dimensional vector space over $\mathbb{C}$ with a positive definite Hermitian form $\langle , \rangle$. Let $T_1, T_2, \dots$ be a sequence of commuting Hermitian operators (i.e., $T_i$ are linear transformations of $V$ such that for any $i$ and $j$ we have $T_iT_j = T_jT_i$ and $\langle T_iv, v' \rangle = \langle v, T_iv' \rangle$). Then $V$ has a basis consisting of vectors that are eigenvectors for all $T_i$. In other words, the $T_i$ are simultaneously diagonalizable.*

*Proof.* Since $\mathbb{C}$ is algebraically closed, $T_1$ has an eigenvector $e_1$. Let $V_1 = \mathbb{C}e_1$ be the subspace of $V$ spanned by $e_1$, and let $V_2 = V_1^{\perp}$ be the orthogonal complement of $V_1$ with respect to $\langle , \rangle$. Since the pairing is positive definite $V = V_1 \oplus V_2$, and moreover this decomposition is stable under the action of $T_1$. Indeed, for $v_2 \in V_2$ we have

$$\langle e_1, T_1v_2 \rangle = \langle T_1e_1, v_2 \rangle = \lambda_1 \langle e_1, v_2 \rangle = 0,$$

where $\lambda_1$ is the eigenvalue of $e_1$ for the action of $T_1$. Applying the same argument to $V_2$, and induction on the dimension of $V$ we get that $T_1$ is diagonalizable.

Now write $V = \oplus V(\lambda_i)$, where the $\lambda_i$ are the *distinct* eigenvalues of $T_1$ and $V(\lambda_i)$ is the eigenspace for $\lambda_i$; thus $T_1$ acts as multiplication by $\lambda_i$ on $V(\lambda_i)$. Since $T_2$ commutes with $T_1$, each $V(\lambda_i)$ is stable under the action of $T_2$. So we can decompose each $V(\lambda_i)$ further as a direct sum of eigenspaces for $T_2$. Continuing in this manner we will arrive at a decomposition $V = \oplus W_j$ such that each $T_i$ acts as a scalar on each $W_j$. Now choose a basis for each $W_j$ and take the union. □

Let $V = \oplus \mathbb{Q}e_i$ be the $n$-dimensional $\mathbb{Q}$-vector space, where each $e_i$ corresponds to an isomorphism class $E_i$ of s.s. elliptic curves over $\overline{\mathbb{F}}_p$. The Brandt matrices can be described as linear transformations of $V$, with the action of $B(m)$ given by

$$B(m)e_i = \sum_{C_m} e_i/C_m,$$

where the sum is over all subgroup-schemes of order $m$ of $E_i$ and $e_i/C_m$ is the basis vector of $V$ corresponding to the isomorphism class of $E_i/C_m$. As in [4], we define a pairing on $V$ given by $\langle e_i, e_j \rangle = w_i\delta_{ij}$, where $\delta_{ij}$ is the Kronecker function. This pairing is obviously positive definite, and each $B(m)$ satisfies

$$\langle B(m)e_i, e_j \rangle = \langle e_i, B(m)e_j \rangle.$$

Indeed, the left hand side is equal to $w_j$ multiplied by the number of order $m$ subgroup-schemes $C_m$ of $E_i$ with $E_i/C_m \cong E_j$. As in the proof of [4, Prop.2.3], this is equal to the number of degree $m$ morphisms from $E_i$ to $E_j$, $\#\mathrm{Hom}(E_i, E_j)_m$. Similarly, the left hand-side is $\#\mathrm{Hom}(E_j, E_i)_m$. Since taking the duals of isogenies identifies $\mathrm{Hom}(E_i, E_j)_m$ with $\mathrm{Hom}(E_j, E_i)_m$, both sides are equal.

We can extend $\langle , \rangle$ to a positive definite Hermitian pairing on $V_{\mathbb{C}} := V \otimes \mathbb{C}$. According to Lemma 3.17, $V_{\mathbb{C}}$ has a basis of eigenvectors for all $B(m)$, $m \geq 1$. Hence $V_{\mathbb{C}}$ is obviously a semi-simple module for the algebra $\mathbb{B} \otimes \mathbb{C}$ generated by the $B(m)$ over $\mathbb{C}$. By [6, XVII.4.7], $\mathbb{B}_{\mathbb{C}}$ is a semi-simple $\mathbb{C}$-algebra, and by [6, XVII.6.1] $B_{\mathbb{Q}}$ is a semi-simple $\mathbb{Q}$-algebra. (If the Jacobson radical $N$ of $\mathbb{B}_{\mathbb{Q}}$ is non-zero then the

radical $N \otimes \mathbb{C}$ of $\mathbb{B}_{\mathbb{C}}$ is also non-zero. But $\mathbb{B}_{\mathbb{C}}$ obviously has no nilpotent elements, which leads to contradiction.) It remains to show that $\mathbb{B}_{\mathbb{Q}}$ is a direct product of totally real number fields. Note that a priori we know

$$\mathbb{B}_{\mathbb{Q}} \cong K_1 \times \cdots \times K_r,$$

where $K_i$ are number fields, since the only commutative simple $\mathbb{Q}$-algebras are the finite field extensions of $\mathbb{Q}$ (for example, from Theorem 1.2).

Let $V_{\mathbb{C}} = \oplus \mathbb{C} v_i$ be the decomposition of $V_{\mathbb{C}}$ into simultaneous eigenspaces for all $B(m)$ provided by the spectral theorem. Let

$$K_{v_i} = \mathbb{Q}(\lambda_i(1), \lambda_i(2), \dots),$$

where $B(m)v_i = \lambda_i(m)v_i$. Since $\mathbb{B}_{\mathbb{Q}}$ is finitely generated over $\mathbb{Q}$ and all $\lambda_i(m)$ are algebraic numbers, $K_{v_i}$ is a number field. Moreover, it is totally real since the eigenvalues of Hermitian operators are real. Consider the homomorphism of $\mathbb{Q}$-algebras given by

$$\begin{aligned} \mathbb{B}_{\mathbb{Q}} &\longrightarrow K_{v_1} \times \cdots \times K_{v_n} \\ B(m) &\mapsto (\lambda_1(m), \dots, \lambda_n(m)). \end{aligned}$$

This homomorphism is clearly injective. Hence $\mathbb{B}_{\mathbb{Q}}$ must be a direct product of totally real number fields (note that $K_{v_i} = K_{v_j}$ if and only if $v_i$ and $v_j$ are in the same $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-orbit). □

## 4. The heights of special points

The purpose of this section is to explain the calculations in [4, §§10–11]. We start with a discussion of §3 *loc. cit.* In §3 Gross constructs a certain 1-dimensional variety over $\mathbb{Q}$ whose $K$-rational points, with $K$ being an imaginary quadratic field, correspond to the pairs $(\Lambda, f)$ where $\Lambda$ is maximal order in $B_p$ and $f$ is an optimal embedding of the ring of integers of $K$ into $\Lambda$ modulo the conjugation by $\Lambda^\times$. The existence of this "moduli space" makes the whole approach of the paper very similar to the calculations in Gross-Zagier. Nevertheless, for the actual calculations in §10 one does not need this space; it is possible to simply fix an embedding of $K$ into $B_p$ and then directly to define an action of $\mathrm{Pic}(\mathcal{O})$ on the s.s. elliptic curves over $\overline{\mathbb{F}}_p$ which contain $\mathcal{O}$ in their endomorphism rings. This being done, the rest boils down to computing the number of isogenies of given degree between certain s.s. elliptic curves. We will take this somewhat simpler approach.

4.1. **Field embeddings into a quaternion algebra.** Let $B := B_p$ be the quaternion over $\mathbb{Q}$ ramified at $p$ and $\infty$. Let $K$ be an imaginary quadratic field of discriminant $D = -d$, where $d$ is a prime with $d \equiv 3 \pmod 4$ and $d > 3$.

**Lemma 4.1.** *The field $K$ embeds into $B$ if and only if $\left(\frac{p}{d}\right) \neq 1$.*

*Proof.* By Lemma 1.7 we know that $B \cong H(a, b)$ for some $a, b \in \mathbb{Q}^\times$. If $K \hookrightarrow B$ then from the proof of Lemma 1.7 it is clear that we can take $a = -d$. If $d = p$ then $\left(\frac{p}{d}\right) = 0$ and we are done. Assume $d \neq p$. For $H(-d, b)$ to be ramified at $p$ the quadratic form $X^2 + dY^2 - bZ^2 - dbT^2$ should not have non-zero solutions mod $p$. Hence $b$ must be divisible by $p$, c.f. Lemma 1.13, and $\left(\frac{-d}{p}\right) = -1$. But

$$\left(\frac{-d}{p}\right) = (-1)^{(p-1)/2}\left(\frac{d}{p}\right) = (-1)^{\frac{p-1}{2}}(-1)^{\frac{p-1}{2}\frac{d-1}{2}}\left(\frac{p}{d}\right) = \left(\frac{p}{d}\right).$$

Conversely, if $\left(\frac{p}{d}\right) \neq 1$ then one easily checks that $H(-d,-p)$ (resp. $H(-d,-1)$) is ramified exactly at $p$ and $\infty$ when $d \neq p$ (resp. $d = p$). Since $B = K \oplus Kj$, where $j\alpha = \overline{\alpha}j$ for $\alpha \in K$ and $j^2 = -p$ (resp. $j^2 = -1$), our field $K = \mathbb{Q}(\sqrt{-d})$ obviously embeds into this quaternion algebra. □

If we have two embeddings $f_1$ and $f_2$ of $K$ into $B$ then a $\mathbb{Q}$-isomorphism $f_1(K) \cong f_2(K)$ extends to an automorphism of $B$, hence by Noether-Skolem must be given by conjugation. Fixing the canonical embedding of $K = \mathbb{Q}(i)$, $i^2 = -d$, into $H(-d,-p)$, all the other embeddings are in one-to-one correspondence with the conjugacy classes $hih^{-1}$, $h \in H^\times$, of $i$.

Let $\mathcal{O}$ be the ring of integers of $K$, and let $\Lambda$ be a fixed maximal order in $B$. Consider an embedding $f$ of $K$ into $B$. We say that $f$ gives an *optimal embedding* of $\mathcal{O}$ into $\Lambda$ if $f(K) \cap \Lambda = \mathcal{O}$. It is clear that for any $f$ the ring $\mathcal{O}$ is optimally embedded into some maximal order. As in §3 we denote by $\{I_1, \ldots, I_n\}$ a set of ideals representing the distinct left ideal classes of $\Lambda$, and we denote $\Lambda_i = \mathcal{O}_r(I_i)$ the right order of $I_i$. We have shown that this is a maximal order in $B$. Denote by $h_i(d)$ the number of optimal embeddings of $\mathcal{O}$ into $\Lambda_i$, modulo conjugation by $\Lambda_i^\times$.

Let $\hat{\Lambda} = \prod_v \Lambda_v$ and $\hat{B} = \prod_v' B_v$, where the restricted product is with respect to $\hat{\Lambda}$. Every ideal $I$ of $\Lambda$ is locally principal, $I_v = \Lambda_v g_v$ with $g_v \in \Lambda_v^\times \setminus B_v^\times$. The element $g_I = (\cdots g_v \cdots)$ lies in $\hat{\Lambda}^\times \setminus \hat{B}^\times$, and every such coset determines a left ideal $I = (\hat{\Lambda} g) \cap B$ of $\Lambda$. From this description it is clear that $\Lambda_i = B \cap g_i^{-1} \hat{\Lambda} g_i$. Hence to give an optimal embedding of $\mathcal{O}$ into $\Lambda_i$ is equivalent to giving a field homomorphism $f : K \to B$ such that $f(K) \cap g_i^{-1} \hat{\Lambda} g_i = f(\mathcal{O})$ in $\hat{B}$. Let $f : K \to B$ be a fixed embedding. Since all other embeddings are obtained by conjugating this one, we conclude that all optimal embeddings of $\mathcal{O}$ into *some* $\Lambda_i$, modulo conjugation by $\Lambda_i^\times$, are classified by the elements of the double coset space $\hat{\Lambda}^\times \setminus \hat{N}^\times / f(K^\times)$, where

$$\hat{N}^\times = \{g \in \hat{B}^\times \mid f(K) \cap g^{-1} \hat{\Lambda} g = f(\mathcal{O})\}.$$

**Theorem 4.2** (Eichler)**.** *Let $h(d)$ be the class number of $K$. Then*

$$\sum_{i=1}^{n} h_i(d) = \left(1 - \left(\frac{p}{d}\right)\right) h(d).$$

*Proof.* See [4, pp.132-134]. First of all, if $\left(\frac{p}{d}\right) = 1$ then by Lemma 4.1 there are no embeddings of $K$ into $B$ so both sides of the equality in the theorem are zero. Assume $\left(\frac{p}{d}\right) \neq 1$ from now on.

Let $f : K \to B$ be an embedding, and denote by $\hat{f}$ its extension to an embedding $\hat{f} : \hat{K}^\times \to \hat{B}^\times$. Here $\hat{K}^\times$ denotes the group of finite ideles over $K$. Denote $\hat{\mathcal{O}} = \prod_v \mathcal{O}_v$ with the product over the finite places of $K$, and consider $K^\times$ as being diagonally embedded in $\hat{K}^\times$. The sum $\sum_{i=1}^n h_i(d)$ is the order of the double coset space $\hat{\Lambda}^\times \setminus \hat{N}^\times / f(K^\times)$. There is an exact sequence

$$0 \to \hat{\mathcal{O}} \setminus \hat{K}^\times / K^\times \xrightarrow{f} \hat{\Lambda}^\times \setminus \hat{N}^\times / f(K^\times) \to \hat{\Lambda}^\times \setminus \hat{N}^\times / \hat{f}(\hat{K}^\times) \to 0.$$

The fact that $f(\hat{K}^\times) \subset \hat{N}^\times$ follows from the commutativity of $K$. Indeed, $f(K) \cap g^{-1} \hat{\Lambda} g = f(\mathcal{O})$ if and only if $g f(K) g^{-1} \cap \hat{\Lambda} = g f(\mathcal{O}) g^{-1}$. When $g \in f(\hat{K}^\times)$ the second condition becomes $f(K) \cap \hat{\Lambda} = f(\mathcal{O})$. It is well-known that $\hat{\mathcal{O}} \setminus \hat{K}^\times / K^\times$ is isomorphic to the class group $\mathrm{Pic}(\mathcal{O})$ of $K$. On the other hand, the double coset space on the right is a direct product of the local double coset spaces $\Lambda_v^\times \setminus N_v^\times / f(K_v^\times)$

over all finite places $v$ of $\mathbb{Z}$. These latter ones classify the optimal embeddings of $\mathcal{O}_v$ into the maximal order $\Lambda_v$ modulo conjugation by $\Lambda_v^\times$. It remains to count the cardinalities of these local terms.

There are three cases to consider. First let $v = p$. The local algebra $B_v$ is a division algebra and there is a canonical valuation on $B_v$. The ring $\Lambda_v$ is the unique maximal order in $B_v$ consisting exactly of the elements which have non-negative valuation. Also, since this valuation is the unique valuation which extends the valuation from the quadratic subfields of $B_v$, it is clear that any embedding of $K_v \to B_v$ gives an optimal embedding of $\mathcal{O}_v$ into $\Lambda_v$. Hence $N_v^\times = B_v^\times$. The existence of the valuation also easily implies that $\Lambda_v^\times$ has two left cosets in $B_v^\times$ given by $\Lambda_v^\times$ and $\Lambda_v^\times \alpha$ with $\mathrm{val}(\alpha) = 1/2$. Hence the set $\Lambda_v^\times \setminus N_v^\times / f(K_v^\times)$ has either 1 or 2 elements depending on whether $\alpha$ is in $f(K_v^\times)$ or not. This is equivalent to whether $K_v/\mathbb{Q}_p$ is ramified or not (by assumption it is not split). So the number of elements is $\left(1 - \left(\frac{p}{d}\right)\right)$.

Next, let $v \neq p$ be a prime which stays prime in $K$. We can write $\mathcal{O}_v = \mathbb{Z}_v[u]$, where $u = \frac{1+\sqrt{-d}}{2}$. The local algebra $B_v$ is isomorphic to $M_2(\mathbb{Q}_v)$ with a standard maximal order $M_2(\mathbb{Z}_v)$. We claim that there is a unique way to embed $\mathcal{O}_v$ into $M_2(\mathbb{Z}_v)$, modulo conjugation by $\mathrm{GL}_2(\mathbb{Z}_v)$. Intuitively, the reason for this is that any such embedding is uniquely determined by the image of $u$. On the other hand, since $f_v : K_v \to B_v$ is a field embedding the unique minimal polynomial of $u$ in $\mathbb{Z}_v[X]$ will be the characteristic polynomial of the matrix $f_v(u)$. Hence the conjugacy class of $f_v(u)$ is uniquely determined. Then one tries to show that the matrix by which we are conjugating can be chosen with integral coefficients (this is rather tedious); see [18, II.3]. Another way to prove this statement is to note that we are asking the following question: given a free rank-2 $\mathbb{Z}_v$-module $S$ in how many ways we can extend $\mathbb{Z}_v$ structure to a $\mathcal{O}_v$-algebra structure (assuming there is at least one). $S$ is automatically a free $\mathcal{O}_v$ module of rank one as $S$ is torsion free and is over a local ring; see [7, Thm. 2.5]. Hence, up to a $\mathbb{Z}_v$-module isomorphism, there is a unique $\mathcal{O}_v$-algebra structure on $S$.

Finally, let $v$ be a place which splits in $K$. Then $K_v \cong \mathbb{Q}_v \oplus \mathbb{Q}_v$. Let $e_1$ and $e_2$ be the idempotents corresponding to this splitting. The images of $e_1$ and $e_2$ under any embedding of $K_v$ into $M_2(\mathbb{Q}_v)$ are two matrices $T_1$ and $T_2$ which satisfy the following conditions: $T_1^2 = T_1$, $T_2^2 = T_2$ and $T_1 + T_2 = I$. Then one has to show that $T_1$ and $T_2$ can be simultaneously conjugated by an element of $\mathrm{GL}_2(\mathbb{Z}_v)$ into $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. Alternatively, any free rank-2 $\mathbb{Z}_v$-module $S$ has a unique structure of a faithfull $S$-algebra. Hence any two such algebras are isomorphic over $\mathbb{Z}_v$.

The above analysis shows that $\hat{\Lambda}^\times \setminus \hat{N}^\times / \hat{f}(\hat{K}^\times)$ has $\left(1 - \left(\frac{p}{d}\right)\right)$ elements which finishes the proof. □

4.2. **Heights.** In this subsection we assume $\left(\frac{p}{d}\right) = -1$. By Lemma 4.1 $K$ embeds in $B$ and gives a decomposition $B = K \oplus Kj$, where $j^2 = -p$, $j\alpha = \bar{\alpha} j$ for all $\alpha \in K$. Let $\Lambda$ be a maximal order which contains $\mathcal{O}$, and let $\mathfrak{a}$ be an integral ideal of $\mathcal{O}$. Consider a s.s. elliptic curve $E$ over $\overline{\mathbb{F}}_p$ with endomorphism ring $\Lambda$. Since $I = \Lambda\mathfrak{a}$ is a left $\Lambda$-ideal, we can form $E_\mathfrak{a} := E^{(I)}$. If $\mathfrak{a}$ and $\mathfrak{a}'$ are in the same ideal class $A \in \mathrm{Pic}(\mathcal{O})$ then $\Lambda\mathfrak{a}$ and $\Lambda\mathfrak{a}'$ are also in the same left ideal class. Hence $E_\mathfrak{a}$ is well-defined by $A$, and we can denote it by $E_A$. In this way we get an action of the group $\mathrm{Pic}(\mathcal{O})$ on the set of isomorphism classes of s.s. elliptic curves over $\overline{\mathbb{F}}_p$ whose

endomorphism rings contain $\mathcal{O}$ under the above fixed embedding of $K$ into $B$. To see how $\mathcal{O}$ embeds in $\mathrm{End}(E_A) = \mathcal{O}_r(\Lambda\mathfrak{a})$, c.f. Lemma 3.6, note that $\mathcal{O}$ naturally acts on the right of $\Lambda\mathfrak{a}$.

Let $\Lambda$ be a maximal order which contains $\mathcal{O} \oplus \mathcal{O}j$. It is clear that under the obvious embedding $f$ of $K$ into $B$, given by sending $K$ into the first factor of the decomposition $B = K \oplus Kj$, we have $f(K) \cap \Lambda = f(\mathcal{O})$. Let $\{I_1, I_2, \ldots, I_n\}$ represent the left ideal classes of $\Lambda$ with $I_1 = \Lambda$. Let $M = \oplus_{i=1}^n \mathbb{Z}x_i$ be the free $\mathbb{Z}$-module of rank $n$, where $x_i$ correspond to distinct isomorphism classes of s.s. elliptic curves over $\overline{\mathbb{F}}_p$. We have defined an action of $\mathrm{Pic}(\mathcal{O})$ on $x_1$; the image of $x_1$ under the action of ideals in class $A$ will be denoted by $x_A$. In our later calculations we will need an explicit description of $\Lambda$.

**Lemma 4.3.** *Explicitly, $\Lambda$ can be chosen to be*

$$\{\alpha + \beta j \mid \alpha, \beta \in \mathfrak{d}^{-1},\ \alpha \equiv \varepsilon\beta \pmod{\mathcal{O}_\mathfrak{d}}\},$$

*where $\varepsilon^2 \equiv -p \pmod{d}$, $\mathfrak{d}^{-1}$ is the inverse different of $\mathcal{O}$, and $\mathcal{O}_\mathfrak{d}$ is the localization of $\mathcal{O}$ at the prime ideal $\mathfrak{d}$.*

*Proof.* It is clear that the set in the lemma contains $\mathcal{O}\oplus\mathcal{O}j$. What is not so obvious is why $\Lambda$ is an order and why it is maximal (after all, $\Lambda$ must be a ring but $\mathfrak{d}^{-1}$ is not a ring). Recall that

$$\mathfrak{d}^{-1} = \{x \in K \mid \mathrm{Tr}_{K/\mathbb{Q}}(xy) \in \mathbb{Z} \text{ for all } y \in \mathcal{O}\}.$$

It is clear that $\Lambda$ is a full $\mathbb{Z}$-lattice in $B$. To check that it is a ring let $\alpha_1 + \beta_1 j$ and $\alpha_2 + \beta_2 j$ be two elements of $\Lambda$. Then

$$\begin{aligned}(\alpha_1 + \beta_1 j)(\alpha_2 + \beta_2 j) &= (\alpha_1\alpha_2 - p\beta_1\bar{\beta}_2) + (\alpha_1\beta_2 + \beta_1\bar{\alpha}_2)j \\ &\equiv (\varepsilon^2\beta_1\beta_2 - p\beta_1\bar{\beta}_2) + (\varepsilon\beta_1\beta_2 + \beta_1\bar{\varepsilon}\bar{\beta}_2)j \\ &\equiv \beta_1\varepsilon^2\mathrm{Tr}_{K/\mathbb{Q}}(\beta_2) + \beta_1\varepsilon\mathrm{Tr}_{K/\mathbb{Q}}(\beta_2)j \pmod{\mathcal{O}_\mathfrak{d}}.\end{aligned}$$

In the second congruence we are using the assumption that $-p = \varepsilon^2 + nd$, for some $n \in \mathbb{Z}$, and the fact that $d\beta_1\beta_2$ is in $d\mathfrak{d}^{-2}$, with the latter an integral $\mathcal{O}$-ideal. Since $\mathrm{Tr}_{K/\mathbb{Q}}(\beta_2) \in \mathbb{Z}$, the last expression of the congruences is in $\Lambda$, so $\Lambda$ is an order.

To show that it is a maximal order we will use a somewhat ad hoc argument. Denote $\Delta = \mathcal{O}\oplus\mathcal{O}j$; this is an order in $B$. A $\mathbb{Z}$-basis for $\Delta$ is given by $1, \frac{1+i}{2}, j, \frac{j+ij}{2}$, where $i^2 = -d$ (here we are using $d \equiv 3 \pmod 4$). As one easily checks, the discriminant of $\Delta$ with respect to the reduced trace on $B$ is equal to $p^2d^2$; that is, $\det(\mathrm{Tr}_B(v_i\bar{v}_j))_{i,j} = p^2d^2$, where $v_1, \ldots, v_4$ is some ordering of the above basis and conjugation is the canonical involution of $B$. One then needs the notion of a discriminant of an arbitrary order (which is not necessarily a free $\mathbb{Z}$-module); see [14, §2]. Discriminants localize, and using this fact one easily shows that the discriminants are squares of integers. Moreover, the maximal orders have discriminants equal to $p^2$. (The situation is very similar to number field case where the ramified primes divide the discriminant of the extension.) We have a strict containment $\Delta \subset \Lambda$. Hence $p^2d^2 = [\Lambda : \Delta]^2\mathrm{disc}(\Lambda)$ and $p^2|\mathrm{disc}(\Lambda)$. Since $d$ is a prime, we must have $\mathrm{disc}(\Lambda) = p^2$. This last equality implies $\Lambda$ is maximal. □

We had an action of the Hecke algebra $\mathbb{T}$ on $M$ which was canonically isomorphic to the action of the algebra generated over $\mathbb{Z}$ by the Brandt matrices; see [4, §§4–5]. The central subject of [4, §10] is to compute

$$\langle x_B, T_m x_{AB}\rangle,$$

where $\langle\cdot,\cdot\rangle$ is the pairing in §4 of *loc. cit.*, and $T_m \in \mathbb{T}$. From the definitions

$$\langle x_B, T_m x_{AB}\rangle = \frac{1}{2}\#\{\phi \in \mathrm{Hom}(x_{AB}, x_A) \mid \deg\phi = m\}. \tag{4.1}$$

**Proposition 4.4.** *Let $\mathfrak{a}$ and $\mathfrak{b}$ be ideals in classes $A$ and $B$ which are relatively prime to $\mathfrak{d} = (\sqrt{-d})$. Then we have a bijection*

$$\mathrm{Hom}(x_{AB}, x_B) \cong \{\alpha + \beta j \mid \alpha \in \mathfrak{d}^{-1}\mathfrak{a},\ \beta \in \mathfrak{d}^{-1}\mathfrak{b}^{-1}\bar{\mathfrak{b}}\bar{\mathfrak{a}},\ \alpha \equiv \varepsilon\beta \pmod{\mathcal{O}_\mathfrak{d}}\}.$$

*If $\phi$ corresponds to $\alpha + \beta j$ then* $\deg\phi = (\mathrm{Nr}(\alpha) + p\mathrm{Nr}(\beta))/\mathrm{Nr}(\mathfrak{a})$.

*Proof.* By the definition of the action of $\mathrm{Pic}(\mathcal{O})$ the left ideal class of $x_{AB}$ is $\Lambda\mathfrak{a}\mathfrak{b}$, and similarly the left ideal class of $x_B$ is $\Lambda\mathfrak{b}$. By Proposition 3.11

$$\mathrm{Hom}(x_{AB}, x_B) \cong (\Lambda\mathfrak{b})^{-1}(\Lambda\mathfrak{a}\mathfrak{b}) = \mathfrak{b}^{-1}\Lambda\mathfrak{a}\mathfrak{b}.$$

From Lemma 4.3 we get the desired expression for $\mathrm{Hom}(x_{AB}, x_B)$. Indeed, the factors $\bar{\mathfrak{b}}\bar{\mathfrak{a}}$ are due to the relation $jx = \bar{x}j$ for all $x \in K$, and since we assumed $\mathfrak{a}$ and $\mathfrak{b}$ to be relatively prime to $\mathfrak{d}$, locally at $\mathfrak{d}$ these ideals are the unit ideals so the congruence relation is preserved.

The final statement of the proposition follows from Proposition 3.12. In fact, $\mathrm{Nr}(\alpha + \beta j) = \mathrm{Nr}(\alpha) + p\mathrm{Nr}(\beta)$ and $\mathrm{Nr}(\mathfrak{b}^{-1}\Lambda\mathfrak{a}\mathfrak{b}) = \mathrm{Nr}(\mathfrak{b})^{-1}\mathrm{Nr}(\mathfrak{b})\mathrm{Nr}(\mathfrak{a}) = \mathrm{Nr}(\mathfrak{a})$. (There is a small discrepancy with our earlier notations; here we denote the generator of the ideal $\mathrm{Nr}(I)$ by same symbol.) □

For an ideal class $A$ let $r_A(n)$ be the number of integral ideals $\mathfrak{a}$ in the class of $A$ with norm $n$ when $n \geq 1$, and let $r_A(0) = 1/2$. Let $R(n) = \sum_A r_A(n)$, where the sum is over all ideal classes of $K$, and $n \geq 0$. Finally, let $\sigma(n) = 1$ if $(n, d) = 1$ and $\sigma(n) = 2$ otherwise.

**Proposition 4.5.** *We have the equality*

$$\langle x_B, T_m x_{AB}\rangle = \sum_{n=0}^{[md/p]} \sigma(n)\cdot r_{A^{-1}}(md - np)\cdot r_{AB^2}(n).$$

*Proof.* By (4.1) and Proposition 4.4 we wish to count half the number of solutions to the identity

$$\mathrm{Nr}(\alpha) + p\mathrm{Nr}(\beta) = m\mathrm{Nr}(\mathfrak{a}), \tag{4.2}$$

with $\alpha \in \mathfrak{d}^{-1}\mathfrak{a}$, $\beta \in \mathfrak{d}^{-1}\mathfrak{b}^{-1}\bar{\mathfrak{b}}\bar{\mathfrak{a}}$, $\alpha \equiv \varepsilon\beta \pmod{\mathcal{O}_\mathfrak{d}}$. First we solve a similar question which does not require the last congruence condition to hold. Consider the ideals

$$\mathcal{L} = (\alpha)\mathfrak{d}\mathfrak{a}^{-1} \quad \text{and} \quad \mathcal{L}' = (\beta)\mathfrak{d}\mathfrak{b}\bar{\mathfrak{b}}^{-1}\bar{\mathfrak{a}}^{-1}, \tag{4.3}$$

which are integral since $\alpha \in \mathfrak{d}^{-1}\mathfrak{a}$, and $\beta \in \mathfrak{d}^{-1}\mathfrak{b}^{-1}\bar{\mathfrak{b}}\bar{\mathfrak{a}}$. Moreover, as $\mathrm{Nr}(\bar{\mathfrak{b}}) = \mathrm{Nr}(\mathfrak{b})$ and $\mathrm{Nr}(\mathfrak{d}) = d$, we have

$$\mathrm{Nr}(\mathcal{L}) + p\mathrm{Nr}(\mathcal{L}') = \mathrm{Nr}(\alpha)d\mathrm{Nr}(\mathfrak{a})^{-1} + p\mathrm{Nr}(\beta)d\mathrm{Nr}(\mathfrak{a})^{-1}.$$

Hence (4.2) holds if and only if

$$\mathrm{Nr}(\mathcal{L}) + p\mathrm{Nr}(\mathcal{L}') = md. \tag{4.4}$$

Since the ideal $\bar{\mathfrak{a}}\mathfrak{a}$ is principal, $\bar{\mathfrak{a}}$ is in $A^{-1}$. Similarly $\bar{\mathfrak{b}} \in B^{-1}$. As $\mathfrak{d} = (\sqrt{-d})$ is principal, we conclude that the ideals $\mathcal{L}$ and $\mathcal{L}'$ lie in the classes of $A^{-1}$ and $AB^2$

respectively, and the number of solutions to (4.4) with ideals in these classes is equal to

$$r_{A^{-1}}(md) + \sum_{n>0} r_{A^{-1}}(md - np) \cdot r_{AB^2}(n),$$

where $n = \mathrm{Nr}(\mathcal{L}')$ and $md - np = \mathrm{Nr}(\mathcal{L})$.

A solution to (4.4) in ideals gives a solution to (4.2) in elements by inverting formula (4.3). The ideal $(\alpha)$ determines the element $\alpha$ up to a unit in $K$. Since we are assuming $d > 3$, the only units are $\pm 1$. So each solution to (4.4) in ideals gives 4 solutions $(\alpha, \beta)$ to (4.2), except when $n = 0$ when we get only 2 solutions $(\alpha, 0)$. We still have to determine how many of those solutions satisfy the congruence $\alpha \equiv \varepsilon\beta \pmod{\mathcal{O}_\mathfrak{d}}$. Locally at $\mathfrak{d}$ we have $\mathrm{Nr}(\beta) \equiv n/d$ and $\mathrm{Nr}(\alpha) \equiv -np/d$ modulo $\mathcal{O}_\mathfrak{d}$. If $n$ is divisible by $d$ then $\alpha$ and $\beta$ are integral at $\mathfrak{d}$ hence the congruence relation is satisfied for any choice of the signs. On the other hand, if $(n, d) = 1$ then $\alpha \equiv \pm\varepsilon/\sqrt{-d}$ and $\beta \equiv \pm 1/\sqrt{-d}$, so only two choices of $(\alpha, \beta)$ satisfy the congruence. This finishes the proof. □

**Corollary 4.6.** *For all $m \geq 1$*

$$\sum_B \langle x_B, T_m x_{AB} \rangle = \sum_{n=0}^{[md/p]} \sigma(n) \cdot r_A(md - np) \cdot R(n).$$

*Proof.* An integral ideal $\mathfrak{a}$ is in $A$ if and only if its conjugate $\bar{\mathfrak{a}}$ is in $A^{-1}$. Moreover, $\mathrm{Nr}(\mathfrak{a}) = \mathrm{Nr}(\bar{\mathfrak{a}})$, so we conclude that $r_{A^{-1}}(k) = r_A(k)$. Next, it is well-known that the 2-torsion of class groups of quadratic imaginary fields is generated by the ideal classes above the ramified primes; see [3, p.181]. Since $\mathfrak{d}$ is principal we get $\mathrm{Pic}(\mathcal{O})[2] = 1$. In particular, $\sum_B r_{AB^2}(n) = R(n)$. Now the claim follows from Proposition 4.5. □

4.3. **The main identity.** Let $f(z) = \sum_{n\geq 1} a_n e^{2\pi i n z}$ be a holomorphic cusp form of weight 2 for the group $\Gamma_0(p)$. Define its $L$-function as $L(f, s) = \sum_{n\geq 1} a_n n^{-s}$. From now on we will assume that $f$ is also an eigenform for the Hecke operators and is normalized by the condition $a_1 = 1$. Let $\varepsilon = \left(\frac{\cdot}{d}\right)$ be the Legendre character. Define the twisted form $f \otimes \varepsilon$ by the sum $\sum_{n\geq 1} a_n \varepsilon(n) e^{2\pi i n z}$; this is a cusp form of weight 2 for the group $\Gamma_0(pd^2)$. Denote

$$L_K(f, s) = L(f, s)L(f \otimes \varepsilon, s).$$

This is the finite part of the $L$-function of the lifting of $\pi_f$ to $K$, where $\pi_f$ is the automorphic representation of GL(2) over $\mathbb{Q}$ attached to $f$. When $f$ has rational coefficients, and hence $L(f, s) = L(E, s)$ for a certain elliptic curve over $\mathbb{Q}$ of level $p$, $L_K(f, s)$ is the $L$-function of $E$ over $K$. In this case $L(f \otimes \varepsilon, s) = L(E_D, s)$ is the $L$-function of the quadratic twist $E_D$ of $E$.

**Lemma 4.7.** *With our previous notations there is the identity*

$$L_K(f, s) = \sum_{\substack{n\geq 1 \\ (n,p)=1}} \varepsilon(n) n^{1-2s} \cdot \sum_{n\geq 1} a_n R(n) n^{-s}.$$

*Proof.* We have three Euler product expansions

$$L(f,s)=\prod_{\ell}[(1-\alpha_\ell\ell^{-s})(1-\beta_\ell\ell^{-s})]^{-1},$$

$$L(f\otimes\varepsilon,s)=\prod_{\ell}[(1-\alpha_\ell\varepsilon(\ell)\ell^{-s})(1-\beta_\ell\varepsilon(\ell)\ell^{-s})]^{-1},$$

$$\zeta_K(s)=\sum_{n\geq 1}R(n)n^{-s}=\sum_{n\geq 1}n^{-s}\cdot\sum_{n\geq 1}\varepsilon(n)n^{-s}=\prod_{\ell}[(1-\ell^{-s})(1-\varepsilon(\ell)\ell^{-s})]^{-1},$$

where $\alpha_\ell\beta_\ell=\ell$ for $\ell\neq p$, and $\alpha_p=\pm 1$, $\beta_p=0$. The first two series have Euler product expansion due to the assumption of $f$ being a newform, and the last one is Dirichlet's factorization formula. Hence by [1, Lem. 1.6.1]

$$\sum_{n\geq 1}a_nR(n)n^{-s}=L(f,s)L(f\otimes\varepsilon,s)\prod_{\ell}(1-\alpha_\ell\beta_\ell\varepsilon(\ell)\ell^{-2s}).$$

Since

$$\sum_{\substack{n\geq 1\\(n,p)=1}}\varepsilon(n)n^{1-2s}=\prod_{\ell}(1-\alpha_\ell\beta_\ell\varepsilon(\ell)\ell^{-2s})^{-1}$$

the claim follows. □

For a ring $F$ we denote $M_F:=M\otimes_{\mathbb{Z}}F=\bigoplus_{i=1}^n Fx_i$. Recall that we defined a bilinear $\mathbb{Z}$-valued positive definite pairing on $M$ by $\langle x_i,x_j\rangle=w_i\delta_{ij}$. This pairing gives an isomorphism of $M^\vee=\mathrm{Hom}(M,\mathbb{Z})$ with the subgroup of $M_{\mathbb{Q}}$ with basis $\{x_i^\vee=x_i/w_i\}_{i=1}^n$. Extend $\langle\cdot,\cdot\rangle$ to a positive definite Hermitian pairing on $M_{\mathbb{C}}$.

Denote by $E_p(z)$ the normalized Eisenstein series of weight-2 for $\Gamma_0(p)$

$$E_p(z)=\frac{p-1}{12}+\sum_{n\geq 1}\sigma(n)_pq^n,\quad\text{where}\quad\sigma(n)_p=\sum_{\substack{m|n\\(m,p)=1}}m\,. \tag{4.5}$$

There is a $\mathbb{T}$-module isomorphism $M_{\mathbb{C}}\cong M_2(p)$, where $M_2(p)$ is the $\mathbb{C}$-space of modular forms of weight 2; see [4, §5]. Moreover, this isomorphism identifies the submodule $M_{\mathbb{C}}^0$ of degree 0 elements in $M_{\mathbb{C}}$ with the subspace of cusp forms $S_2(p)$. Let $f_1,\dots,f_{n-1}$ be the newforms in $S_2(p)$. Since $E_p$ is an eigenform, we have a decomposition $M_2(p)=\mathbb{C}E_p\oplus\mathbb{C}f_1\oplus\cdots\oplus\mathbb{C}f_{n-1}$ into $\mathbb{T}$-eigenspaces. This decomposition induces the corresponding decomposition

$$M_{\mathbb{C}}=M^{\mathrm{Eis}}\oplus M^{f_1}\oplus\cdots\oplus M^{f_{n-1}}, \tag{4.6}$$

where $M^{f_i}$ (resp. $M^{\mathrm{Eis}}$) is the $\mathbb{T}$-eigenspace with eigenvalues $a_n(f_i)$ (resp. $\sigma(n)_p$). Recall that $M_{\mathbb{C}}$ is a semi-simple $\mathbb{T}$-module; c.f. Proposition 3.16. Since the action of $\mathbb{T}$ on $M_{\mathbb{C}}$ is self-adjoint with respect to $\langle\cdot,\cdot\rangle$, the decomposition (4.6) is orthogonal for this pairing. Using this decomposition, write $c_K=\sum_A x_A$ as

$$c_K=c_{E,K}+\sum_{i=1}^{n-1}c_{f_i,K}, \tag{4.7}$$

where $c_{f_i,K}$ is the projection of $c_K$ to the $f_i$-isotypical component $M^{f_i}$, and similarly for $c_{E,K}$. Note that $\deg c_{f,K}=0$, as $f$ is a cusp form.

**Theorem 4.8.**

$$L_K(f,1) = \frac{(f,f)}{\sqrt{d}}\langle c_{f,K}, c_{f,K}\rangle.$$

*In particular,* $L_K(f,1) \geq 0$ *with equality if and only if* $c_{f,K} = 0$.

*Proof.* See [4, Prop. 11.2]. Following [4, §5], for any $x \in M$ and $y \in M^\vee$ define the weight-2 modular form

$$\phi(x,y) = \frac{\deg x \cdot \deg y}{2} + \sum_{m\geq 1}\langle T_m x, y\rangle q^m.$$

Lemma 4.7, Corollary 4.6 and the analytic calculations in [4] imply

$$L_K(f,1) = \sum_{A,B}\frac{(f,\phi(x_B, x_{AB}))}{\sqrt{d}} = \frac{(f,\phi(c_K,c_K))}{\sqrt{d}}.$$

We need to show that the $f$-eigencomponent of the modular form $\phi(c_K,c_K)$ is $\langle c_{f,K}, c_{f,K}\rangle f$. The map $\phi : M_{\mathbb{C}} \times M_{\mathbb{C}}^\vee \to M_2(p)$ is $\mathbb{T}$-bilinear, so $\phi(c_{E,K}, c_{f_i,K}) = 0$ and $\phi(c_{f_j,K}, c_{f_i,K}) = 0$ unless $i = j$. Hence, as one easily checks, the $f$-eigencomponent is $\phi(c_{f,K}, c_{f,K})$. On the other hand,

$$\begin{aligned}\phi(c_{f,K}, c_{f,K}) &= \sum_{m\geq 1}\langle T_m c_{f,K}, c_{f,K}\rangle q^m \\ &= \sum_{m\geq 1}\langle c_{f,K}, c_{f,K}\rangle a_m q^m \\ &= \langle c_{f,K}, c_{f,K}\rangle f.\end{aligned}$$

□

*Remark* 4.9. In Theorem 4.8 we implicitly assume $\varepsilon(p) = -1$. When $\varepsilon(p) = 1$ there is a similar result due to Gross and Zagier which relates the special value of the *derivative* of $L_K(f,s)$ to the heights of Heegner points. Note that the functional equation forces $L_K(f,1) = 0$ when $\varepsilon(p) = 1$. In the case of prime level, Gross-Zagier formula can be stated as follows:

$$L'_K(f,1) = \frac{(f,f)}{\sqrt{d}}\langle H_{f,K}, H_{f,K}\rangle,$$

here $\langle\cdot,\cdot\rangle$ denotes the Néron-Tate pairing on $\mathrm{Jac}(X_0(p))(K)$ and $H_{f,K}$ is the $f$-isotypical component of a certain special point related to $K$, called Heegner point.

**Corollary 4.10.** *Assume* $\ell$ *is a prime which divides the numerator of* $\frac{p-1}{12}$ *but does not divide* $h$. *Then there is a cusp form* $f$ *with* $\mathbb{Z}$*-Fourier coefficients which is congruent to the Eisenstein series* $E_p$ *modulo* $\ell$ *and satisfies* $L(f,1)L(f\otimes\varepsilon) \neq 0$.

*Proof.* First we examine more closely $M^{\mathrm{Eis}}$. Consider $x_E = \sum_{i=1}^n \frac{1}{w_i} x_i$. By [4, Props. 2.7 and 4.4]

$$\begin{aligned} T_m x_E &= \sum_{i=1}^n \frac{1}{w_i} T_m x_i = \sum_{i=1}^n \frac{1}{w_i} \sum_{j=1}^n B_{ij}(m) x_j \\ &= \sum_{j=1}^n x_j \sum_{i=1}^n \frac{1}{w_i} B_{ij}(m) = \sum_{j=1}^n x_j \sum_{i=1}^n \frac{1}{w_j} B_{ji}(m) \\ &= \sum_{j=1}^n \frac{1}{w_j} x_j \sum_{i=1}^n B_{ji}(m) = \sigma(m)_p \sum_{j=1}^n \frac{1}{w_j} x_j \\ &= \sigma(m)_p x_E. \end{aligned}$$

Hence $x_E$ spans $M^{\mathrm{Eis}}$. Since $c_K - c_{E,K} = \sum_{i=1}^{n-1} c_{f_i,K} \in M^0$ has degree 0, the $E_p$-isotypical component of $c_K$ is the multiple of $x_E$ for which $\deg(c_K - \kappa x_E) = 0$. But $\deg(c_K) = h$ and by Eichler's mass formula $\deg(x_E) = \frac{p-1}{12}$. We conclude

$$c_{E,K} = \frac{12h}{p-1} x_E \quad \text{and} \quad \langle c_{E,K}, c_{E,K} \rangle = \left( \frac{12h}{p-1} \right)^2 \langle x_E, x_E \rangle = \frac{12}{p-1} h^2.$$

Since $T_m c_K \in M$, $m \geq 1$, we have $\langle T_m c_K, c_K \rangle \in \mathbb{Z}$. On the other hand,

$$\begin{aligned} \langle T_m c_K, c_K \rangle &= \langle T_m c_{E,K}, c_{E,K} \rangle + \sum_{f_i} \langle T_m c_{f_i,K}, c_{f_i,K} \rangle \\ &= \sigma(m)_p \langle c_{E,K}, c_{E,K} \rangle + \sum_{f_i} a_m(f_i) \langle c_{f_i,K}, c_{f_i,K} \rangle. \end{aligned}$$

If $\ell$ is a prime which divides the numerator of $\frac{p-1}{12}$ but does not divide $h$ then from our previous calculations we conclude that $\ell$ must divide the denominator of $\sum_{f_i} a_m(f_i) \langle c_{f_i,K}, c_{f_i,K} \rangle$. We conclude that the cusp form $f = \alpha_1 f_1 + \cdots + \alpha_{n-1} f_{n-1}$ has Fourier coefficients in $\mathbb{Z}_\ell \cap \mathbb{Q}$, where $\alpha_i = -\frac{\langle c_{f_i,K}, c_{f_i,K} \rangle}{\langle c_{E,K}, c_{E,K} \rangle}$. As $\frac{\langle T_m c_K, c_K \rangle}{\langle c_{E,K}, c_{E,K} \rangle} \in \ell \mathbb{Z}_\ell$

$$E_p \equiv f (\mathrm{mod}\ \ell).$$

The pairing $\langle \cdot, \cdot \rangle$ is $\mathbb{Q}$-valued on $M_{\mathbb{Q}}$. Hence for any $\tau \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ and any $f_i$ we have $\langle c_{f_i^\tau,K}, c_{f_i^\tau,K} \rangle = \langle c_{f_i,K}, c_{f_i,K} \rangle^\tau$. Without loss of generality we can assume that all $\alpha_i$ are invertible over $\ell$; the newforms in the sum forming $f$ with $\alpha$-coefficients having poles or zeros over $\ell$ must cancel each other out. Let $f_1, \ldots, f_s$, $s \leq n-1$, form one Galois conjugacy class. By the above argument $\alpha_1, \ldots, \alpha_s$ are also Galois conjugate. Choose $\beta_1$ integral over $\mathbb{Z}$ and having the same reduction as $\alpha_1$ modulo $\ell$. For $\alpha_i = \alpha_1^\tau$ take $\beta_i = \beta_1^\tau$. Then $\beta_1 f_1 + \cdots \beta_s f_s$ has $\mathbb{Z}$-Fourier coefficients and is congruent to $\alpha_1 f_1 + \cdots \alpha_s f_s$ modulo $\ell$. By repeating the same process for other conjugacy classes, we get a cusp form $g = \beta_1 f_1 + \cdots \beta_{n-1} f_{n-1}$ with $\mathbb{Z}$-Fourier coefficients and $E_p \equiv g (\mathrm{mod}\ \ell)$.

So far we have not used Theorem 4.8. The theorem will allow us to choose $g$ with $L_K(g,1) \neq 1$. Indeed,

$$\begin{aligned} L_K(g,1) &= (\sum_{i=1}^{n-1} \beta_i L(f_i,1))(\sum_{i=1}^{n-1} \beta_i L(f_i \otimes \varepsilon, 1)) \\ &= \sum_{i=1}^{n-1} \beta_i^2 L_K(f_i,1) + (\cdots). \end{aligned}$$

Since at least one $\langle c_{f_i,K}, c_{f_i,K}\rangle$ is non-zero, by Theorem 4.8 at least one $\beta_i^2 L_K(f_i,1)$ is positive. We have a certain freedom in choosing $\beta$'s; for example, we can replace every $\beta_i$ by $\beta_i + \ell$. Then it is not hard to check that we can choose $\beta_1, \ldots, \beta_{n-1}$ such that the sum representing $L_K(g,1)$ does not vanish. □

4.4. **$\mathbb{T}$-generators of modular forms.** This subsection is somewhat unrelated with what was discussed earlier in the section. Its goal is to explain the statement after [4, (12.13)].

Consider the $\mathbb{Q}$-vector space $X := M_2(p)_{\mathbb{Q}}$ of modular forms of level $p$ and weight 2. There is a perfect bilinear $\mathbb{T}_{\mathbb{Q}}$-equivariant pairing $\mathbb{T}_{\mathbb{Q}} \times X \to \mathbb{Q}$, which is constructed using the Fourier expansion of modular forms; c.f. [2, Lem. 1.34]. In particular, $\dim_{\mathbb{Q}} X = \dim_{\mathbb{Q}} \mathbb{T}_{\mathbb{Q}}$ and $\mathbb{T}_{\mathbb{Q}}$ acts faithfully on $X$. The algebra $\mathbb{T}_{\mathbb{Q}}$ is semi-simple; c.f. Proposition 3.16. From all of this we conclude that $X$ is a free $\mathbb{T}_{\mathbb{Q}}$-module of rank 1, i.e., $X \cong \mathbb{T}_{\mathbb{Q}}$ as $\mathbb{T}_{\mathbb{Q}}$-modules. Let $X_{\mathbb{Z}} \subset X$ be the lattice of modular forms with all Fourier coefficients in $\mathbb{Z}$ except possibly for $a_0$. There is an element $v$ in $X_{\mathbb{Z}}$ which generates a $\mathbb{Q}$-basis of $X$ under the action of $\mathbb{T}$, that is, $(\mathbb{T}v) \otimes \mathbb{Q} = X$. Indeed, $\mathbb{T}_{\mathbb{Q}}$ is obviously a cyclic $\mathbb{T}_{\mathbb{Q}}$-module. So $X$ is also a cyclic $\mathbb{T}_{\mathbb{Q}}$-module. Let $w$ be a $\mathbb{T}_{\mathbb{Q}}$-generator. Since $X = X_{\mathbb{Z}} \otimes \mathbb{Q}$, c.f. [2, Thm.1.31], we can assume $w \in X_{\mathbb{Z}}$ by scaling it appropriately. As $(\mathbb{T}w) \otimes \mathbb{Q} = \mathbb{T}_{\mathbb{Q}} w = X$ we can take $v = w$. Note that even though $\mathbb{T}$ preserves $X_{\mathbb{Z}}$, we cannot necessarily choose $v$ such that $\mathbb{T}v = X_{\mathbb{Z}}$; we only get a sublattice of finite index.

Let $f_{ij} \in X_{\mathbb{Z}}$, $1 \leq i,j \leq n$, be the modular forms constructed in [4, §1]. From Prop. 5.6 in *loc.cit.* we know that $\oplus_{i,j=1}^n \mathbb{Q} f_{ij} = X$. Since $\dim_{\mathbb{Q}} X = n$, one is tempted to claim that some natural subsets of $n$ elements among $n^2$ theta series $f_{ij}$ form a $\mathbb{Q}$-basis of $X$, for example, $f_{11}, f_{12}, \ldots, f_{1n}$. In general, this is false, and the (still open) problem of singling out the correct $n$ series is known as *Hecke's basis problem*. From the previous paragraph we know that there is $\mathbb{Z}$-linear combination of $f_{ij}$ which generates $X$ under the action of $\mathbb{T}_{\mathbb{Q}}$.

**Lemma 4.11.** *The modular form $F = f_{11} + f_{22} + \cdots + f_{nn}$ generates $X$ over $\mathbb{T}_{\mathbb{Q}}$.*

*Proof.* Let $g_1, g_2, \ldots, g_n$ be the newforms of level $p$. Then $F = \sum_{i=1}^n g_i$. Indeed, by definition [4, (1.4)]

$$f_{ii} = \sum_{m \geq 0} B_{ii}(m) q^m.$$

Hence by [4, (5.4)], $F = \sum_{m \geq 0} \mathrm{Tr}(T_m) q^m$. But the trace of $T_m$ is the sum of its eigenvalues, so it is equal to the coefficient of $q^m$ in $\sum_{i=1}^n g_i$. It is clear that the matrix $(a_m(g_j))_{m,j}$, $0 \leq m, 1 \leq j \leq n$, has rank $n$. The claim follows. □

## 5. Elliptic curves with prime conductor

In this section we compare the formula in Theorem 4.8 with the conjecture of Birch and Swinnerton-Dyer.

5.1. **Optimal quotients.** Let $A$ be an abelian variety over a field $K$. We say that the abelian variety $B$ over $K$ is an *optimal quotient* of $A$ if there is a surjective homomorphism over $K$ of abelian varieties $A \to B$ whose kernel $C$ is connected and smooth. In other words, there is a short exact sequence of abelian varieties over $K$

$$0 \to C \to A \to B \to 0.$$

**Lemma 5.1.** *Suppose $K$ is a subfield of $\mathbb{C}$. The morphism $\varphi : A \to B$ defined over $K$ is an optimal quotient if and only if the functorially induced homomorphism $\varphi_* : \mathrm{H}_1(A(\mathbb{C}), \mathbb{Z}) \to \mathrm{H}_1(B(\mathbb{C}), \mathbb{Z})$ is surjective.*

*Proof.* The kernel of $\varphi$ is defined over $K$ by assumption. Clearly as a group-scheme $\ker(\varphi)$ is an extension of a finite étale group-scheme $G$ by an abelian variety $C$ defined over $K$. Hence $\ker(\varphi)$ is connected if and only if $G = 1$. This last condition hold if and only if $G \times_K \mathrm{Spec}(\mathbb{C}) = 1$. Hence without loss of generality we assume from now on that $K = \mathbb{C}$.

Let $T_A = \mathrm{H}^0(A, \Omega^1)^\vee$ be the tangent space to $A$ at the identity, and let $\Lambda_A = \mathrm{H}_1(A, \mathbb{Z})$. There is a natural injective homomorphism

$$\begin{aligned} \Lambda_A &\to T_A \\ \gamma &\mapsto (\omega \mapsto \int_\gamma \omega), \end{aligned}$$

where $\omega \in \mathrm{H}^0(A, \Omega^1)$. The quotient $T_A/\Lambda_A$ is naturally identified with $A(\mathbb{C})$; see [11, §1]. Since $\varphi$ is a surjective homomorphism, it lifts to a surjective homomorphism of analytic groups $\varphi_* : T_A \to T_B$; note that $T_A$ and $T_B$ are the universal covering spaces of $A(\mathbb{C})$ and $B(\mathbb{C})$ respectively. Moreover, $\varphi_*$ takes $\Lambda_A$ to $\Lambda_B$ as these are the abelinizations of the fundamental groups of $A(\mathbb{C})$ and $B(\mathbb{C})$. Collecting these functorial maps, we get a commutative diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \Lambda_A & \longrightarrow & T_A & \longrightarrow & A(\mathbb{C}) & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \varphi_*} & & \downarrow{\scriptstyle \varphi_*} & & \downarrow{\scriptstyle \varphi} & & \\ 0 & \longrightarrow & \Lambda_B & \longrightarrow & T_B & \longrightarrow & B(\mathbb{C}) & \longrightarrow & 0. \end{array}$$

The snake lemma gives an exact sequence of analytic groups

$$0 \to \ker(\Lambda_A \to \Lambda_B) \to \ker(T_A \to T_B) \to \ker(A \to B) \to \mathrm{coker}(\Lambda_A \to \Lambda_B) \to 0.$$

Since $\ker(T_A \to T_B)$ is obviously connected, $\ker(A \to B)$ is connected if and only if the finite group $\mathrm{coker}(\Lambda_A \to \Lambda_B)$ is trivial. □

If $\varphi : A \to B$ is an optimal quotient and $\psi : B \to B'$ is an isogeny then the composite $\psi \circ \varphi : A \to B'$ is not an optimal quotient, since its kernel is not connected, unless $\psi$ is an isomorphism.

Now let $X$ be a smooth projective geometrically connected curve over $K$ which has a $K$-rational point. Fix some point in $X(K)$ and denote it by $\infty$. Let $E$ be an elliptic curve and let $\pi : X \to E$ be a covering over $K$ which sends $\infty$ to the origin. By Albanese functorially $\pi$ induces a homomorphism $J \to E$ over $K$, where $J$ is

the jacobian variety of $X$. There is a canonical (Abel-Jacobi) embedding $X \hookrightarrow J$ sending $\infty$ to the origin, and the original map $\pi$ can be recovered as the composite $X \hookrightarrow J \to E$; see [10, §6]. There are canonical isomorphisms $\mathrm{H}^0(X, \Omega^1) \cong \mathrm{H}^0(J, \Omega^1)$ and $\mathrm{H}_1(X, \mathbb{Z}) \cong \mathrm{H}_1(J, \mathbb{Z})$ which identify the functorial homomorphisms $\mathrm{H}_1(X, \mathbb{Z}) \to \mathrm{H}_1(E, \mathbb{Z})$ and $\mathrm{H}_1(J, \mathbb{Z}) \to \mathrm{H}_1(E, \mathbb{Z})$; see [10, §2]. Hence by Lemma 5.1, $J \to E$ is optimal if and only if $\pi_* : \mathrm{H}_1(X, \mathbb{Z}) \to \mathrm{H}_1(E, \mathbb{Z})$ is surjective.

Conversely, let $\varphi : J \to E$ be a quotient map. This implies that $\varphi_* : T_J \to T_E$ is surjective. Consider the composite of the Abel-Jacobi map $\psi : X \hookrightarrow J$ with the quotient $J \to E$. This gives a morphism $\pi : X \to E$. Since $\psi_*$ sends the tangent space around $\infty$ isomorphically to the tangent space of $J$ around the origin, $\pi_* = \psi_* \circ \varphi_*$ is surjective. Hence $\pi$ is non-trivial.

5.2. **Birch and Swinnerton-Dyer conjecture.** Let $K$ be a number field. Denote by $\mathcal{O}$ the ring of integers of $K$. Given a place $v$ of $K$, denote by $K_v$ the completion of $K$ at $v$. If $v$ is non-archimedean then denote by $\mathcal{O}_v$ the ring of integers of $K_v$, and denote the residue field by $k_v$.

Let $E$ be an elliptic curve over $K$. Denote by $\mathcal{E}$ the Néron model of $E$ over $\mathcal{O}$. Let $\mathcal{E}^0$ be the relative connected component of the identity of $\mathcal{E}$, i.e., the largest open subgroup-scheme of $\mathcal{E}$ in which all fibres are connected. Let $\mathcal{E}_v := \mathcal{E} \times_{\mathcal{O}} k_v$ and $\mathcal{E}_v^0 := \mathcal{E}^0 \times_{\mathcal{O}} k_v$. The group of connected components $\Phi_{E,v} = \mathcal{E}_v / \mathcal{E}_v^0$ of $E$ at $v$ is a finite étale group-scheme over $k_v$. The *Tamagawa number* $c_v(E)$ of $E$ at $v$ is the order of the subgroup $\Phi_{E,v}(k_v)$ of $k_v$-rational points in $\Phi_{E,v}(\bar{k}_v)$. The group $\Phi_{E,v}$ is trivial at almost all places (for example, it is trivial at the places where $E$ has good reduction), hence almost all $c_v = 1$, and we can define the product

$$\Omega_{\mathrm{f}}(E) = \prod_{v \text{ finite}} c_v(E).$$

Let $\Omega^1_{\mathcal{E}}$ be the relative canonical sheaf on $\mathcal{E}$. Then $\mathrm{H}^0(\mathcal{E}, \Omega^1_{\mathcal{E}})$ is rank-1 projective $\mathcal{O}$-submodule of $\mathrm{H}^0(E, \Omega^1_E) \cong K$. If we fix a basis $\omega$ of the 1-dimensional $K$-vector space $\mathrm{H}^0(E, \Omega^1_E)$, then

$$\mathrm{H}^0(\mathcal{E}, \Omega^1_{\mathcal{E}}) = \omega \mathcal{I}_E \subset \mathrm{H}^0(E, \Omega^1_E)$$

for some fractional ideal $\mathcal{I}_E$ of $K$. The Tamagawa number $c_v(E)$ of $E$ at an archimedean place $v$ is the real number $\int_{E(K_v)} |\omega|_v$. Denote

$$\Omega_\infty(E) = \mathrm{Nr}_{K/\mathbb{Q}}(\mathcal{I}_E) \cdot \prod_{v|\infty} c_v(E).$$

One can check that $\Omega_\infty$ is independent of the choice of $\omega$. Let $d_K$ be the absolute discriminant of $K$. Finally define

$$\Omega(E) = \Omega_\infty(E) \cdot \Omega_{\mathrm{f}}(E) / |d_K|^{1/2}.$$

Now assume $E(K)$ is finite, and denote its order by $t$. In this situation the conjecture of Birch and Swinnerton-Dyer predicts that

$$L(E, 1) = \Omega(E) \cdot \#Ш(E)/t^2, \tag{5.1}$$

where Ш is the Tate-Shafarevich group of $E$ over $K$.

5.3. **Optimal quotients of modular Jacobians.** Let $X := X_0(p)$ be the modular curve of prime level $p$. Let $J := J_0(p)$ be its Jacobian variety. Since the cusp $\infty$ of $X$ is $\mathbb{Q}$-rational, we can apply the theory in §5.1 to this situation.

Let $f$ be a newform of level $p$ which is also a cuspform. As explained in [2, §1.7] one can associate to $f$ an optimal quotient $A_f$ of $J := J_0(p)$. The abelian variety $A_f$ depends on the Galois conjugacy class of $f$, and its dimension is equal to the absolute degree of the number field formed by the Fourier coefficients of $f$ over $\mathbb{Q}$. In particular, if $f$ has integer coefficients then $A_f$ is an elliptic curve.

**Lemma 5.2.** *With previous notations, let $f$ have integer coefficients. Denote $E_0 := A_f$. Then in its $\mathbb{Q}$-isogeny class $E_0$ is the unique, up to an isomorphism, optimal quotient of $J$. If $E$ is an elliptic curve $\mathbb{Q}$-isogenous to $E_0$, and $\omega_E$ is a $\mathbb{Q}$-generator of $\mathrm{H}^0(E,\Omega^1_E)$, then $\pi^*(\omega_E) = c\cdot f(q)\frac{dq}{q}$.*

*Proof.* This follows from some well-known facts which are given, for example, in [2, Ch.1]. The facts we need are the following: (1) The isomorphisms

$$\begin{aligned} S_2(p)_{\mathbb{Q}} \xrightarrow{\sim} \mathrm{H}^0(X,\Omega^1_X) &\xrightarrow{\sim} \mathrm{H}^0(J,\Omega^1_J)\\ f \mapsto f(q)\frac{dq}{q} & \end{aligned}$$

are $\mathbb{T}$-equivariant. (2) The splitting $J\otimes\mathbb{Q} = \prod_{[f]} A_f$ corresponds to the splitting of $S_2(p)_{\mathbb{Q}}$ into a direct sum of $\mathbb{T}_{\mathbb{Q}}$-invariant vector spaces.

Let $E$ be a quotient of $J$. Then up to an isogeny $E$ is one of $A_f$ and $\pi^*(\omega_E) = c\cdot f(q)\frac{dq}{q}$. Since the Frobenius eigenvalues of $E$ determine $f$, all we need to prove is that the same newform with rational coefficients does not appear twice in $S_2(p)_{\mathbb{Q}}$. This is a consequence of *multiplicity one* theorem. □

For two newforms $f$ and $g$ define the *Petersson inner product*

$$\begin{aligned}(f,g) = \int\limits_{X(\mathbb{C})} f(q)\frac{dq}{q}\wedge\overline{g(q)\frac{dq}{q}} &= \int\limits_{\Gamma_0(p)\backslash\mathcal{H}} 2\pi i f(z)dz\wedge\overline{2\pi i g(z)dz}\\ &= 4\pi^2\int\limits_{\Gamma_0(p)\backslash\mathcal{H}} f(z)\bar{g}(z)dz\wedge\overline{dz}.\end{aligned}$$

Assume $E$ is an elliptic curve over $\mathbb{Q}$, which is a quotient of $J$. Let $f$ be the newform corresponding to $E$. Let $\omega$ be a non-zero differential of $E$.

**Lemma 5.3.** *We have*

$$\deg\pi\int\limits_{E(\mathbb{C})}|\omega| = c^2(f,f),$$

*where $\pi^*(\omega) = c\cdot f(q)\frac{dq}{q}$.*

*Proof.* Indeed, $\deg\pi\int_{E(\mathbb{C})}\omega\wedge\overline{\omega} = \int_{X(\mathbb{C})}\pi^*(\omega)\wedge\overline{\pi^*(\omega)} = c^2(f,f)$. □

Let $K$ be a quadratic imaginary field of discriminant $-d$, where $d$ is a prime with $d\equiv 3\ (\mathrm{mod}4)$ and $\left(\frac{p}{d}\right) = -1$. The elliptic curve $E$ has conductor $p$, hence it has only one place of bad reduction, namely $p$, and the reduction is multiplicative. The prime $p$ remains inert in $K$. From the theory of Néron models, $E_K$ has bad reduction only at $p$ and the reduction is necessarily split multiplicative. In this

situation it is known that $c_p(E_K) = \mathrm{ord}_p \Delta_E$, where $\Delta_E$ is the minimal discriminant of $E$. For all these facts we refer to [16, Ch.IV]. In particular, in the notations of §5.2, $\Omega_{\mathrm{f}}(E_K) = \mathrm{ord}_p(\Delta_E)$. Choose the holomorphic differential $\omega$ to be a global Néron differential – this is possible since $\mathbb{Z}$ is a principal ideal domain; c.f. [15, Prop. VIII.8.2]. Then $\Omega_\infty(E_K) = \int_{E(\mathbb{C})} |\omega|$. If we further assume $E$ is an optimal quotient, then *Manin's conjecture* says that $c = 1$ in Lemma 5.3. With these assumptions

$$\Omega(E_K) = \frac{(f,f)}{\sqrt{d}} \frac{\mathrm{ord}_p(\Delta_E)}{\deg \pi}.$$

Let in the notations of §4.3, $e_f$ be a primitive $f$-isotypical component of $M$ (note that $e_f$ is indeed in $M$ as $f$ has rational coefficients). Then $c_{f,K} = \frac{\langle e_f, c_d\rangle}{\langle e_f, e_f\rangle} \cdot e_f$. One can show [8] that $\deg \pi \cdot \mathrm{ord}_p(\Delta_E) = \langle e_f, e_f\rangle$ (this relies on some deep results of Grothendieck, Mazur and Ribet). We can rewrite

$$\Omega(E_K) = \frac{(f,f)}{\sqrt{d}} \frac{\mathrm{ord}_p(\Delta_E)^2}{\langle e_f, e_f\rangle}. \tag{5.2}$$

Denote $m_d = \langle e_f, c_d\rangle$. Since $\langle e_{f,K}, e_{f,K}\rangle = m_d^2/\langle e_f, e_f\rangle$, we can rewrite the identity in Theorem 4.8 as

$$L(E_K, 1) = \frac{(f,f)}{\sqrt{d}} \frac{m_d^2}{\langle e_f, e_f\rangle}.$$

If $L_K(f,1) \neq 0$ then $m_d \neq 0$. Substituting (5.2) into (5.1), we also have the conjectural equality

$$L(E_K, 1) = \frac{(f,f)}{\sqrt{d}} \frac{\mathrm{ord}_p(\Delta_E)^2}{\langle e_f, e_f\rangle} \frac{\#Ш(E_K)}{\#E(K)^2}$$

Finally, it is known [8] that $\mathrm{ord}_p(\Delta_E) = \#E(\mathbb{Q})$. Hence, for the optimal curve $E$, we have

$$\#Ш(E_K) = \#(E(K) : E(\mathbb{Q}))^2 \cdot \langle e_f, c_d\rangle^2. \tag{5.3}$$

**Lemma 5.4.** *Let $E$ be an elliptic curve over $\mathbb{Q}$ with prime conductor $p$ (not necessarily optimal). Assume $E(K)$ has rank 0, and assume the parity conjecture. Then $E(K) = E(\mathbb{Q})$.*

*Proof.* Let $D = -d$ be the discriminant of $K$. Let $E$ be given in terms of a minimal Weierstrass equation $y^2 = x^3 + Ax + B$ for some $A, B \in \mathbb{Q}$. Let $E_D$ be the quadratic twist $Dy^2 = x^3 + Ax + B$; in terms of the standard form of the Weierstrass equation, $E_D$ is also given by $y^2 = x^3 + AD^2x + BD^3$. Over $K$ there is an isomorphism $E \cong E_D$. We assumed that $E(K)$ is finite. Obviously this implies that $E(\mathbb{Q})$ and $E_D(\mathbb{Q})$ are also finite. First, for any integer $n \neq 2$, we claim

$$E(K)[n] \subseteq E(\mathbb{Q})[n] \oplus E_D(\mathbb{Q})[n].$$

Indeed, let $\sigma$ be the generator of $\mathrm{Gal}(K/\mathbb{Q})$, and consider the homomorphism $E(K) \to E(\mathbb{Q})$ given by $P \mapsto P + P^\sigma$. The point $P$ is in the kernel of this homomorphism if and only if $P^\sigma = -P$. In terms of coordinates if we put $P = (x, y)$ this last condition means $(\sigma x, \sigma y) = (x, -y)$. Hence $x \in \mathbb{Q}$ and $y = \sqrt{D}y'$ with $y' \in \mathbb{Q}$ uniquely determined. If $y' = 0$ then $(x, y) \in E(\mathbb{Q})[2]$, otherwise, as is easy to see, this point gives the point $(x, y')$ in $E_D(\mathbb{Q})$. So it is enough to show that under the assumptions of the lemma we have $E(\mathbb{Q})[2] = E(K)[2]$ and $E_D(\mathbb{Q})[n] = 1$ for $n \neq 2$.

The discriminant of $E_D$ is $\Delta(E_D) = d^6 \cdot \Delta(E)$. The reduction of $E_D$ at $d$ is additive potentially good, and at $p$ the reduction is multiplicative. The conductor of $E_D$ is $pd^2$. The sign of the functional equation of $L(E_D, s)$ is $-w_p \cdot w_d$. Here $w_d = \left(\frac{-1}{d}\right)$, and $w_p = 1$ or $-1$ depending on whether $E_D$ has non-split or split reduction at $p$ respectively; see [13, §19]. Since we have assumed $d \equiv 3 \pmod 4$, and $E_D(\mathbb{Q})$ is finite, under the parity conjecture, $E_D$ must have a multiplicative non-split reduction at $p$. Let $\mathcal{E}$ be the Néron model of $E_D$ over $\mathbb{Z}$. $\mathrm{Frob}_p$ induces a non-trivial automorphism of the finite cyclic group $\Phi_p(\overline{\mathbb{F}}_p) = \mathcal{E}_{\mathbb{F}_p}/\mathcal{E}^0_{\mathbb{F}_p}(\overline{\mathbb{F}}_p)$, such that $\mathrm{Frob}_p^2 = 1$. Hence $\Phi_p(\mathbb{F}_p) = 1$ or $\mathbb{Z}/2$. We conclude that $\mathcal{E}(\mathbb{F}_p)$ is a subgroup of $\mathbb{Z}/2 \times \mathbb{Z}/(p-1)$. From Néron's table [16, p.365] we also have $\mathcal{E}(\mathbb{F}_d) \subseteq \mathbb{Z}/2 \times \mathbb{Z}/2 \times \mathbb{Z}/d$. From Oort-Tate classification of group-schemes of prime order, one can conclude that $E_D(\mathbb{Q})$ injects into $\mathcal{E}(\mathbb{F}_p)$ and $\mathcal{E}(\mathbb{F}_d)$. If there is a rational point which is not a 2-torsion point then it must be a point of order $d$. But then we must have $p \equiv 1 \pmod d$, which contradicts our assumption $\left(\frac{p}{d}\right) = -1$. Hence $E_D(\mathbb{Q}) = E_D(\mathbb{Q})[2]$ as was required.

Next, recall that the non-trivial 2-torsion elements of $E(\overline{\mathbb{Q}})$ are the points $(x, 0)$, where $x$ is a root of $x^3 + Ax + B$. If there is a 2-torsion point in $E(K)$ which is not in $E(\mathbb{Q})$ then it means that $x^3 + Ax + B$ splits over $K$ without being split over $\mathbb{Q}$. But then the discriminant of $K$ must divide the discriminant of this cubic polynomial, which is also the discriminant of $E$. Since $E$ has prime conductor $p$, $\Delta(E)$ is a power of $p$. Hence $d = p$, a contradiction. □

Using this lemma, we can modify (5.3) into

**Conjecture 5.5.** *For the optimal curve $E$ we must have*

$$\#Ш(E_K) = \langle e_f, c_d \rangle^2.$$

## References


1. D. Bump, *Automorphic forms and representations*, Cambridge Univ. Press, 1998.
2. H. Darmon, F. Diamond, and R. Taylor, *Fermat's last theorem*.
3. A. Fröhlich and M. Taylor, *Algebraic number theory*, Cambridge, 1991.
4. B. Gross, *Heights and the special values of $L-$series*, Canad. Math. Soc. Conf. Proc. **7** (1987), 115–187.
5. T. Hungeford, *Algebra*, GTM 73, Springer, 1974.
6. S. Lang, *Algebra*, Addison-Wesley, 1993.
7. H. Matsumura, *Commuative ring theory*, Cambridge University Press, 1986.
8. J.-F. Mestre and J. Oesterlé, *Courbes de weil semi-stables de discriminant une puissance m-ième*, J. reigne und angew. Math. **400** (1989), 173–184.
9. J. Milne, *Class field theory*, available at www.jmilne.org.
10. J. S. Milne, *Jacobian varieties*, in Arithmetic Geometry (1986), 167–212.
11. D. Mumford, *Abelian varieties*, Oxford University Press, 1970.
12. I. Reiner, *Maximal orders*, Academic Press, 1975.
13. D. Rohrlich, *Elliptic curves and the Weil-Deligne group*, CRM Proceedings **4** (1994), 125–157.
14. J-P. Serre, *Local fields*, GTM 67, Springer, 1979.
15. J. Silverman, *The arithmetic of elliptic curves*, GTM 106, Springer, 1986.
16. ———, *Advanced topics in the arithmetic of elliptic curves*, GTM 151, Springer, 1994.
17. J. Tate, *Endomorphisms of abelian varieties over finite fields*, Invent. Math. **2** (1966), 134–144.
18. M.-F. Vignéras, *Arithmétique des algèbres de quaternions*, LNM 800, Springer, 1980.