13 Endomorphism algebras

The key to improving the efficiency of elliptic curve primality proving (and many other algorithms) is the ability to directly construct an elliptic curve $E/\mathbb{F}_q$ with a specified number of rational points, rather than generating curves at random until a suitable curve is found. To do this we need to develop the theory of *complex multiplication*. As a first step in this direction we introduce the notion of the *endomorphism algebra* of an elliptic curve, which is derived from its endomorphism ring, and classify the possible endomorphism algebras.

Recall from Lecture 7 that the endomorphism ring $\mathrm{End}(E)$ of an elliptic curve E/k consists of the isogenies from E to itself, together with the zero morphism; addition is defined point-wise and multiplication is composition. The ring $\mathrm{End}(E)$ is not necessarily commutative, as we saw in Problem set 4, but the multiplication-by-n maps $[n]$ form a subring of $\mathrm{End}(E)$ isomorphic to $\mathbb{Z}$ that lies in the center of $\mathrm{End}(E)$. We may identify this subring with $\mathbb{Z}$, writing n rather than $[n]$ without risk of confusion: note that $n\phi = \phi+\cdots+\phi$ is the same as $[n]\circ\phi$. Thus $\mathbb{Z} \subseteq \mathrm{End}(E)$, but this inclusion is not necessarily an equality.

Definition 13.1. An elliptic curve E/k has *complex multiplication* (CM) if $\mathrm{End}(E) \neq \mathbb{Z}$.

As we shall see in later lectures, the term arises from the fact that endomorphisms of elliptic curves over $\mathbb{C}$ can be viewed as "multiplication-by-α" maps, for some complex number α. If $\mathrm{End}(E) = \mathbb{Z}$ then α is an integer, but in general α is an algebraic integer in a quadratic field, namely, the splitting field of its characteristic polynomial $x^2 - (\mathrm{tr}\,\alpha)x + \deg\alpha \in \mathbb{Z}[x]$.

Our first objective is to classify the different endomorphism rings that are possible. We begin by recalling some basic facts about $\mathrm{End}(E)$ that we proved in Lecture 7:

- $\mathrm{End}(E)$ has no zero divisors (in particular, it is torsion free);
- the subring $\mathbb{Z} \subseteq \mathrm{End}(E)$ lies in the center of $\mathrm{End}(E)$;
- $\deg\alpha \in \mathbb{Z}_{\geq 0}$ is the degree of $\alpha \neq 0$ as a rational map ($\deg 0 = 0$);
- $\deg(\alpha\beta) = \deg(\alpha)\deg(\beta)$ for all $\alpha, \beta \in \mathrm{End}(E)$;
- $\deg n = n^2$ for all $n \in \mathbb{Z} \subseteq \mathrm{End}(E)$;
- every $\alpha \in \mathrm{End}(E)$ has a unique *dual* $\hat{\alpha} \in \mathrm{End}(E)$ for which $\alpha\hat{\alpha} = \hat{\alpha}\alpha = \deg\alpha = \deg\hat{\alpha}$;
- $\hat{n} = n$ for all $n \in \mathbb{Z} \subseteq \mathrm{End}(E)$;
- $\widehat{\alpha+\beta} = \hat{\alpha} + \hat{\beta}$ for all $\alpha, \beta \in \mathrm{End}(E)$;
- the *trace* of α is defined by $\mathrm{tr}\,\alpha = \alpha + \hat{\alpha} = \mathrm{tr}\,\hat{\alpha}$;
- $\mathrm{tr}\,\alpha = \deg\alpha + 1 - \deg(\alpha - 1) \in \mathbb{Z}$ for all $\alpha \in \mathrm{End}(E)$;
- α and $\hat{\alpha}$ are roots of the characteristic equation $x^2 - (\mathrm{tr}\,\alpha)x + \deg\alpha \in \mathbb{Z}[x]$.

Lemma 13.2. *For all $\alpha, \beta \in \mathrm{End}(E)$ we have $\widehat{\alpha\beta} = \hat{\beta}\hat{\alpha}$.*

Proof. We have

$$(\alpha\beta)(\hat{\beta}\hat{\alpha}) = \alpha(\deg\beta)\hat{\alpha} = \alpha\hat{\alpha}(\deg\beta) = \deg\alpha\deg\beta = \deg(\alpha\beta),$$

thus $\widehat{\alpha\beta} = \hat{\beta}\hat{\alpha}$. □

Together with the facts $\widehat{\alpha+\beta} = \hat{\alpha}+\hat{\beta}$, $\hat{\hat{\alpha}} = \alpha$ and $\hat{1} = 1$, the lemma implies that the map $\varphi \mapsto \hat{\varphi}$ is an *involution* of $\mathrm{End}(E)$.

Definition 13.3. An *anti-homomorphism* $\varphi\colon R \to S$ of rings is a homomorphism of their additive groups for which $\varphi(1_R) = 1_S$ and $\varphi(\alpha\beta) = \varphi(\beta)\varphi(\alpha)$ for all $\alpha, \beta \in R$. An *involution* (also called an *anti-involution*) is an anti-homomorphism $\varphi\colon R \to R$ that is its own inverse, meaning that $\varphi \circ \varphi$ is the identity map.

Note that an involution is an isomorphism when the ring is commutative, but not in general (the restriction to the center is always an isomorphism).

13.1 The endomorphism algebra of an elliptic curve

The additive group of $\mathrm{End}(E)$, like all abelian groups, is a $\mathbb{Z}$-module. Recall that if R is a commutative ring, an *R-module* M is an (additively written) abelian group that admits a scalar multiplication by R compatible with its structure as an abelian group. This means that for all $\alpha, \beta \in M$ and $r, s \in R$:

$$(r+s)\alpha = r\alpha + s\alpha, \qquad r\alpha + r\beta = r(\alpha+\beta), \qquad r(s\alpha) = (rs)\alpha, \qquad 1\alpha = \alpha.$$

(one can check these imply $0\alpha = 0$ and $(-1)\alpha = -\alpha$).

The ring $\mathrm{End}(E)$ is not only a $\mathbb{Z}$-module. Like all rings, it has a multiplication that is compatible with its structure as a $\mathbb{Z}$-module, and this makes it a $\mathbb{Z}$-algebra. For any commutative ring R, an (associative unital) *R-algebra* A is a (not necessarily commutative) ring equipped with a ring homomorphism $R \to A$ that maps R into the center of A.[1] In our situation the map $\mathbb{Z} \to \mathrm{End}(E)$ sending n to $[n]$ is injective and we simply view $\mathbb{Z}$ as a subring of $\mathrm{End}(E)$. When we have a ring A with an involution that is also an R-algebra, we typically require the involution to fix R in order to view it as an R-algebra involution; for the involution $\alpha \mapsto \hat{\alpha}$ on our $\mathbb{Z}$-algebra $\mathrm{End}(E)$, this is true.

We now want to "upgrade" our $\mathbb{Z}$-algebra $\mathrm{End}(E)$ to a $\mathbb{Q}$-algebra (in other words, a $\mathbb{Q}$-vector space with a multiplication that is compatible with its structure as a vector space), To do this we take the tensor product of $\mathrm{End}(E)$ with $\mathbb{Q}$.

Definition 13.4. The *endomorphism algebra* of E is $\mathrm{End}^0(E) := \mathrm{End}(E) \otimes_{\mathbb{Z}} \mathbb{Q}$.

For a commutative ring R, the *tensor product* $A \otimes_R B$ of two R-modules A and B can be defined as the R-module generated by the formal symbols $\alpha \otimes \beta$ with $\alpha \in A$ and $\beta \in B$, subject to the relations

$$(\alpha_1+\alpha_2)\otimes\beta = \alpha_1\otimes\beta + \alpha_2\otimes\beta, \qquad \alpha\otimes(\beta_1+\beta_2) = \alpha\otimes\beta_1 + \alpha\otimes\beta_2, \qquad r\alpha\otimes\beta = \alpha\otimes r\beta,$$

for $\alpha_1, \alpha_2 \in A$, $\beta_1, \beta_2 \in B$ and $r \in R$. The third relation lets us pull out scalars from A or B and treat them as scalars of $A \otimes_R B$. Notice that the tensor product is *not* a direct product (or sum); for example, we always have $0 \otimes \beta = \alpha \otimes 0 = 0$, but this is not true in any non-trivial direct product. Another difference is that the general element of $A \otimes_R B$ is a finite sum of monomial tensors $\alpha \otimes \beta$ and cannot necessarily be written as a single $\alpha \otimes \beta$ (but in the special case we are interested in, this is actually true — see Lemma 13.5 below).

[1] Here we consider only associative unital algebras; one can define a more general notion of an R-algebra that is not necessarily a ring (Lie algebras, for example).

When A and B are R-algebras, not just R-modules, we give the tensor product $A \otimes_R B$ the structure of an R-algebra by defining

$$(\alpha_1 \otimes \beta_1)(\alpha_2 \otimes \beta_2) = \alpha_1\alpha_2 \otimes \beta_1\beta_2.$$

This allows us to compute $(\sum_i \alpha_i \otimes \beta_i)(\sum_j \alpha_j \otimes \beta_j)$ via the distributive law (which holds by decree); the multiplicative identity is $1_A \otimes 1_B$. The R-algebras A and B can be canonically mapped into $A \otimes_R B$ via $\alpha \mapsto \alpha \otimes 1_B$ and $\beta \mapsto 1_A \otimes \beta$. These maps need not be injective (indeed, $A \otimes_R B$ may be the zero ring even when A and B are not, consider $\mathbb{Z}/2\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/3\mathbb{Z}$, for example), but in the case we are interested in they are.

If you have never seen tensor products before this might seem like a lot of abstraction to swallow, but fear not! We are actually only interested in what is essentially the simplest non-trivial example of a tensor product of algebras: we are tensoring two $\mathbb{Z}$-algebras, one of which is the fraction field of $\mathbb{Z}$. The net effect is that we are simply extending our ring of scalars $\mathbb{Z}$ to its fraction field $\mathbb{Q}$. This actually simplifies makes things simpler.

With this in mind, we will simply write elements $\alpha \otimes r$ of $\mathrm{End}(E) \otimes_{\mathbb{Z}} \mathbb{Q}$ as $r\alpha$. The only difference between $r\alpha$ with $r \in \mathbb{Q}$ and $n\alpha$ with $n \in \mathbb{Z}$ is that the former is not necessarily an endomorphism, it is a formal element of the endomorphism algebra $\mathrm{End}^0(E)$. But if we multiply $r\alpha$ by the denominator of r we will get an element of $\mathrm{End}(E)$ (viewed as a subring of $\mathrm{End}^0(E)$ via the natural embedding). To justify this notation, we should verify that every element of $\mathrm{End}^0(E) = \mathrm{End}(E) \otimes_{\mathbb{Z}} \mathbb{Q}$ can actually be written in the form $\alpha \otimes r$.

Lemma 13.5. *Let R be an integral domain with fraction field B, and let A be an R-algebra. Every element of $A \otimes_R B$ can be written in the form $a \otimes b$ with $a \in A$ and $b \in B$.*

Proof. It suffices to show that $\alpha_1 \otimes \beta_1 + \alpha_2 \otimes \beta_2$ can be written as $\alpha_3 \otimes \beta_3$. Let $\beta_1 = r_1/s_2$ and $\beta_2 = r_2/s_2$ with $r_1, r_2, s_1, s_2 \in R$. Then

$$\begin{aligned}
\alpha_1 \otimes \beta_1 + \alpha_2 \otimes \beta_2 &= \alpha_1 \otimes \frac{r_1}{s_1} + \alpha_2 \otimes \frac{r_2}{s_2} \\
&= \alpha_1 \otimes \frac{r_1 s_2}{s_1 s_2} + \alpha_2 \otimes \frac{r_2 s_1}{s_1 s_2} \\
&= r_1 s_2 \alpha_1 \otimes \frac{1}{s_1 s_2} + r_2 s_1 \alpha_2 \otimes \frac{1}{s_1 s_2} \\
&= (r_1 s_2 \alpha_1 + r_2 s_1 \alpha_2) \otimes \frac{1}{s_1 s_2},
\end{aligned}$$

so we may take $\alpha_3 = r_1 s_2 \alpha_1 + r_2 s_1 \alpha_2$ and $\beta_3 = 1/(s_1 s_2)$. □

We also note that $\mathbb{Q}$ lies in the center of $\mathrm{End}^0(E)$ (since $\mathbb{Q}$ is commutative and $\mathbb{Z}$ lies in the center of $\mathrm{End}(E)$), and we note that the canonical homomorphisms $\mathrm{End}(E) \to \mathrm{End}^0(E)$ and $\mathbb{Q} \to \mathrm{End}^0(E)$ are injective because $\mathrm{End}(E)$ and $\mathbb{Q}$ are torsion free $\mathbb{Z}$-algebras (but note that the images of these homomorphisms intersect non-trivially), thus we may regard both $\mathbb{Q}$ and $\mathrm{End}(E)$ as subrings of $\mathrm{End}^0(E)$ that intersect in $\mathbb{Z}$.

13.2 The Rosati involution and the reduced norm and trace

We now extend the involution $\alpha \mapsto \hat{\alpha}$ on $\mathrm{End}(E)$ to $\mathrm{End}^0(E)$ by defining $\widehat{r\alpha} = r\hat{\alpha}$ for all $r \in \mathbb{Q}$. This implies that $\hat{r} = r$ for all $r \in \mathbb{Q}$ (take $\alpha = 1$), and therefore $\hat{\hat{\alpha}} = \alpha$ holds for all $\alpha \in \mathrm{End}^0(E)$. We also have $\widehat{\alpha\beta} = \hat{\beta}\hat{\alpha}$ and $\widehat{\alpha + \beta} = \hat{\alpha} + \hat{\beta}$ for all $\alpha, \beta \in \mathrm{End}^0(E)$, since these

hold for elements of $\mathrm{End}(E)$ and scalars are fixed by $\alpha \mapsto \hat{\alpha}$ and commute. Thus the map $\alpha \mapsto \hat{\alpha}$ is an involution of the $\mathbb{Q}$-algebra $\mathrm{End}^0(E)$, and it is known as the *Rosati involution*.

The Rosati involution allows us to extend the notions of degree and trace on $\mathrm{End}(E)$ to a norm and a trace defined on all of $\mathrm{End}^0(E)$.

Definition 13.6. Let $\alpha \in \mathrm{End}^0(E)$. The (reduced) *norm* of α is $\mathrm{N}\alpha = \alpha\hat{\alpha}$ and the (reduced) *trace* of α is $\mathrm{T}\alpha = \alpha + \hat{\alpha}$.

We now show that $\mathrm{N}\alpha$ and $\mathrm{T}\alpha$ lie in $\mathbb{Q}$, and prove some other facts we will need.

Lemma 13.7. *For all $\alpha \in \mathrm{End}^0(E)$ we have $\mathrm{N}\alpha \in \mathbb{Q}_{\geq 0}$, with $\mathrm{N}\alpha = 0$ if and only if $\alpha = 0$. We also have $\mathrm{N}\hat{\alpha} = \mathrm{N}\alpha$ and $\mathrm{N}(\alpha\beta) = \mathrm{N}\alpha\mathrm{N}\beta$ for all $\alpha, \beta \in \mathrm{End}^0(E)$.*

Proof. Write $\alpha = r\phi$, with $r \in \mathbb{Q}$ and $\phi \in \mathrm{End}(E)$. Then $\mathrm{N}\alpha = \alpha\hat{\alpha} = r^2 \deg\phi \geq 0$. If either r or ϕ is zero then $\alpha = 0$ and $\mathrm{N}\alpha = 0$, and otherwise $\mathrm{N}\alpha > 0$. For $\alpha \neq 0$ we have $\alpha\mathrm{N}\hat{\alpha} = \alpha\hat{\alpha}\alpha = (\mathrm{N}\alpha)\alpha = \alpha\mathrm{N}\alpha$, so $\alpha(\mathrm{N}\hat{\alpha} - \mathrm{N}\alpha) = 0$, and multiplying both sides on the left by $\hat{\alpha}$ yields $\mathrm{N}\hat{\alpha}(\mathrm{N}\hat{\alpha} - \mathrm{N}\alpha) = 0$, so $\mathrm{N}\hat{\alpha} = \mathrm{N}\alpha$ (since $\mathrm{N}\alpha, \mathrm{N}\hat{\alpha} \in \mathbb{Q}_{>0}$). For any $\alpha, \beta \in \mathrm{End}^0(E)$,

$$\mathrm{N}(\alpha\beta) = \alpha\beta\widehat{\alpha\beta} = \alpha\beta\hat{\beta}\hat{\alpha} = \alpha(\mathrm{N}\beta)\hat{\alpha} = \alpha\hat{\alpha}\mathrm{N}\beta = \mathrm{N}\alpha\mathrm{N}\beta. \qquad \square$$

Corollary 13.8. *Every nonzero $\alpha \in \mathrm{End}^0(E)$ has a multiplicative inverse α^{-1}.*

Proof. Let $\alpha \in \mathrm{End}^0(E)$ be nonzero. Then $\mathrm{N}\alpha \neq 0$ and $\beta = \hat{\alpha}/\mathrm{N}\alpha \neq 0$, and we have $\alpha\beta = \alpha\hat{\alpha}/\mathrm{N}\alpha = 1$. Thus every nonzero element of $\mathrm{End}^0(E)$ has a right inverse, including β. So let γ be a right inverse of β. Then $\beta\gamma = 1$ and $\gamma = \alpha\beta\gamma = \alpha$, so $\beta\alpha = 1$. Thus β is also a left inverse of α and therefore $\alpha^{-1} = \beta$. $\square$

The corollary implies that $\mathrm{End}^0(E)$ is a *division ring*, meaning that it satisfies all the field axioms except possibly commutativity of multiplication. The division ring $\mathrm{End}^0(E)$ is a field if and only if it is commutative.

Lemma 13.9. *For all $\alpha \in \mathrm{End}^0(E)$ we have $\mathrm{T}\alpha \in \mathbb{Q}$ and $\mathrm{T}\hat{\alpha} = \mathrm{T}\alpha$. For any $r \in \mathbb{Q}$, $\alpha, \beta \in \mathrm{End}^0(E)$ we have $\mathrm{T}(\alpha + \beta) = \mathrm{T}\alpha + \mathrm{T}\beta$, and $\mathrm{T}(r\alpha) = r\mathrm{T}\alpha$.*

Proof. We first note that

$$\mathrm{T}\alpha = \alpha + \hat{\alpha} = 1 + \alpha\hat{\alpha} - (1 - \alpha)(1 - \hat{\alpha}) = 1 + \mathrm{N}\alpha - \mathrm{N}(1 - \alpha) \in \mathbb{Q},$$

and $\mathrm{T}\hat{\alpha} = \hat{\alpha} + \hat{\hat{\alpha}} = \hat{\alpha} + \alpha = \alpha + \hat{\alpha} = \mathrm{T}\alpha$. We also have

$$\mathrm{T}(\alpha + \beta) = \alpha + \beta + \widehat{(\alpha + \beta)} = \alpha + \beta + \hat{\alpha} + \hat{\beta} = \alpha + \hat{\alpha} + \beta + \hat{\beta} = \mathrm{T}\alpha + \mathrm{T}\beta.$$

and

$$\mathrm{T}(r\alpha) = r\alpha + \widehat{r\alpha} = r\alpha + \hat{\alpha}\hat{r} = r\alpha + \hat{\alpha}r = r\alpha + r\hat{\alpha} = r(\alpha + \hat{\alpha}) = r\mathrm{T}\alpha. \qquad \square$$

Lemma 13.10. *For all $\alpha \in \mathrm{End}^0(E)$, α and $\hat{\alpha}$ are roots of the characteristic polynomial*

$$x^2 - (\mathrm{T}\alpha)x + \mathrm{N}\alpha \in \mathbb{Q}[x].$$

Proof. We have

$$0 = (\alpha - \alpha)(\alpha - \hat{\alpha}) = \alpha^2 - \alpha(\alpha + \hat{\alpha}) + \alpha\hat{\alpha} = \alpha^2 - (\mathrm{T}\alpha)\alpha + \mathrm{N}\alpha$$

and similarly for $\hat{\alpha}$, since $\mathrm{T}\hat{\alpha} = \mathrm{T}\alpha$ and $\mathrm{N}\hat{\alpha} = \mathrm{N}\alpha$. $\square$

Corollary 13.11. *For any nonzero $\alpha \in \mathrm{End}^0(E)$, if $\mathrm{T}\alpha = 0$ then $\alpha^2 = -\mathrm{N}\alpha < 0$. An element $\alpha \in \mathrm{End}^0(E)$ is fixed by the Rosati involution if and only if $\alpha \in \mathbb{Q}$.*

Proof. The first statement follows immediately from $\alpha^2 - (\mathrm{T}\alpha)\alpha + \mathrm{N}\alpha = 0$. For the second, we have already noted that $\hat{r} = r$ for $r \in \mathbb{Q}$, and if $\hat{\alpha} = \alpha$ then the discriminant $(\mathrm{T}\alpha)^2 - 4\mathrm{N}\alpha$ of $x^2 - (\mathrm{T}\alpha)x + \mathrm{N}\alpha$ must be zero, in which case $\alpha = (\mathrm{T}\alpha)/2 \in \mathbb{Q}$. □

13.3 Quaternion algebras

In order to give a complete classification of the possible endomorphism algebras $\mathrm{End}^0(E)$ that can arise, we need to introduce quaternion algebras.

Definition 13.12. A *quaternion algebra* over a field k is a k-algebra that has a basis of the form $\{1, \alpha, \beta, \alpha\beta\}$, with $\alpha^2, \beta^2 \in k^\times$ and $\alpha\beta = -\beta\alpha$.

Let H be a quaternion algebra over a field k. Then H is a 4-dimensional k-vector space with basis $\{1, \alpha, \beta, \alpha\beta\}$, and we may distinguish the subspace $k \subseteq H$ spanned by 1, which does not depend on the choice of α and β. The complementary subspace H_0 is the space of *pure quaternions*. Every $x \in H$ has a unique decomposition of the form $a + x_0$ with $a \in k$ and $x_0 \in H_0$. The unique $\hat{x} = a - x_0$ for which $x + \hat{x} = 0$ is the *conjugate* of x.

The map $x \mapsto \hat{x}$ is an involution of the k-algebra H, and we have a reduced trace $\mathrm{T}x := x + \hat{x}$ and reduced norm $\mathrm{N}x := x\hat{x}$, both of which lie in k. It is easy to check that $\mathrm{T}x = \mathrm{T}\hat{x}$, $\mathrm{N}x = \mathrm{N}\hat{x}$ and that the trace is additive and the norm is multiplicative.

Lemma 13.13. *A quaternion algebra is a division ring if and only if $\mathrm{N}x = 0$ implies $x = 0$.*

Proof. Let x be a nonzero element of a quaternion algebra H. Then $\hat{x} \neq 0$ (since $\hat{0} = 0 \neq x$) If H is a division ring, then x has an inverse x^{-1} and $x^{-1}\mathrm{N}x = x^{-1}x\hat{x} = \hat{x} \neq 0$, so $\mathrm{N}x \neq 0$. Conversely, if $\mathrm{N}x \neq 0$ then $x(\hat{x}/\mathrm{N}x) = 1$ and $(\hat{x}/\mathrm{N}x)x = 1$, so x has an inverse $\hat{x}/\mathrm{N}x$, and this implies that H is a division ring. □

Example 13.14. An example of a quaternion algebra that is a division ring is the ring of *Hamiltonians*, which is the $\mathbb{R}$-algebra with basis $\{1, i, j, ij\}$, where $i^2 = j^2 = -1$ and $ij = -ji$ (the product ij is often denoted k).

13.4 Classification theorem for endomorphism algebras

We are now ready to prove our main result, which classifies the possible endomorphism algebras of an elliptic curve.

Theorem 13.15. *Let E/k be an elliptic curve. Then $\mathrm{End}^0(E)$ is isomorphic to one of:*

(i) *the field of rational numbers $\mathbb{Q}$;*

(ii) *an imaginary quadratic field $\mathbb{Q}(\alpha)$ with $\alpha^2 < 0$;*

(iii) *a quaternion algebra $\mathbb{Q}(\alpha, \beta)$ with $\alpha^2, \beta^2 < 0$.*

Quaternion algebras over $\mathbb{Q}$ (or $\mathbb{R}$) with $\alpha^2, \beta^2 < 0$ are said to be *definite*.

Proof. We have $\mathbb{Q} \subseteq \mathrm{End}^0(E)$, and if equality holds we are in case (i). Otherwise, let α be an element of $\mathrm{End}^0(E)$ not in $\mathbb{Q}$. By replacing α with $\alpha - \frac{1}{2}\mathrm{T}\alpha$, we may assume without loss of generality that $\mathrm{T}\alpha = 0$, since

$$\mathrm{T}\left(\alpha - \frac{1}{2}\mathrm{T}\alpha\right) = \mathrm{T}\alpha - \frac{1}{2}\mathrm{T}(\mathrm{T}\alpha) = \mathrm{T}\alpha - \frac{1}{2}(\mathrm{T}\alpha + \widehat{\mathrm{T}\alpha}) = \mathrm{T}\alpha - \frac{1}{2}(\mathrm{T}\alpha + \mathrm{T}\alpha) = 0,$$

where we have $\widehat{\mathrm{T}\alpha} = \mathrm{T}\alpha$ because $\mathrm{T}\alpha \in \mathbb{Q}$. By Corollary 13.11, we have $\alpha^2 < 0$. Thus $\mathbb{Q}(\alpha) \subseteq \mathrm{End}^0(E)$ is an imaginary quadratic field with $\alpha^2 < 0$, and if $\mathbb{Q}(\alpha) = \mathrm{End}^0(E)$ then we are in case (ii).

Otherwise, let β be an element of $\mathrm{End}^0(E)$ not in $\mathbb{Q}(\alpha)$. As with α, we may assume without loss of generality that $\mathrm{T}\beta = 0$ so that $\beta^2 < 0$. Furthermore, by replacing β with

$$\beta - \frac{\mathrm{T}(\alpha\beta)}{2\alpha^2}\alpha \tag{1}$$

we can assume that $\mathrm{T}(\alpha\beta) = 0$ as well (one can check this by multiplying (1) by α and taking the trace, and note that $\mathrm{T}\beta = 0$ still holds). Thus $\mathrm{T}\alpha = \mathrm{T}\beta = \mathrm{T}(\alpha\beta) = 0$. This implies $\alpha = -\hat{\alpha}$, $\beta = -\hat{\beta}$, and $\alpha\beta = -\widehat{\alpha\beta} = -\hat{\beta}\hat{\alpha}$. Substituting the first two equalities into the third, $\alpha\beta = -\beta\alpha$. Applying this together with the fact that $\alpha^2 < 0$ and $\beta^2 < 0$ lie in $\mathbb{Q}$, it is clear that $\{1, \alpha, \beta, \alpha\beta\}$ spans $\mathbb{Q}(\alpha, \beta)$ as a $\mathbb{Q}$-vector space.

To prove that $\mathbb{Q}(\alpha, \beta)$ is a quaternion algebra, it remains only to show that 1, α, β, and $\alpha\beta$ are linearly independent over $\mathbb{Q}$. By construction, 1, α, and β are linearly independent. Suppose for the sake of contradiction that

$$\alpha\beta = a + b\alpha + c\beta,$$

for some $a, b, c \in \mathbb{Q}$. We must have $c \neq 0$, since β and therefore $\alpha\beta$ do not lie in $\mathbb{Q}(\alpha)$ (note that $\alpha^{-1} \in \mathbb{Q}(\alpha)$). Squaring both sides yields

$$(\alpha\beta)^2 = (a^2 + b^2\alpha^2 + c^2\beta^2) + 2a(b\alpha + c\beta) + bc(\alpha\beta + \beta\alpha).$$

The LHS lies in $\mathbb{Q}$, since $\mathrm{T}(\alpha\beta) = 0$, as does the first term on the RHS, since $\mathrm{T}\alpha = \mathrm{T}\beta = 0$. The last term on the RHS is zero, since $\alpha\beta = -\beta\alpha$. Thus $d = b\alpha + c\beta$ lies in $\mathbb{Q}$, but then $\beta = (d - b\alpha)/c$ lies $\in \mathbb{Q}(\alpha)$, which is our desired contradiction. Thus $\mathbb{Q}(\alpha, \beta) \subseteq \mathrm{End}^0(E)$ is a quaternion algebra with $\alpha^2, \beta^2 < 0$, and if $\mathbb{Q}(\alpha, \beta) = \mathrm{End}^0(E)$ then we are in case (iii).

Otherwise, let γ be an element of $\mathrm{End}^0(E)$ that does not lie in $\mathbb{Q}(\alpha, \beta)$. As with β, we may assume without loss of generality that $\mathrm{T}\gamma = 0$ and $\mathrm{T}(\alpha\gamma) = 0$, which implies $\alpha\gamma = -\gamma\alpha$, as above. Then $\alpha\beta\gamma = -\beta\alpha\gamma = \beta\gamma\alpha$, so α commutes with $\beta\gamma$. By Lemma 13.16 below, $\beta\gamma \in \mathbb{Q}(\alpha)$, but then $\gamma \in \mathbb{Q}(\alpha, \beta)$, which is a contradiction. Thus we have addressed every possible case. □

Lemma 13.16. *Suppose that $\alpha, \beta \in \mathrm{End}^0(E)$ commute and that $\alpha \notin \mathbb{Q}$. Then $\beta \in \mathbb{Q}(\alpha)$.*

Proof. As in the proof of the Theorem 13.15, we can linearly transform α and β to some $\alpha' = \alpha + a$ and $\beta' = \beta + b\alpha + c$, where $a, b, c \in \mathbb{Q}$, so that $\mathrm{T}\alpha' = \mathrm{T}\beta' = \mathrm{T}(\alpha'\beta') = 0$, and therefore $\alpha'\beta' = -\beta'\alpha'$ (set $a = \frac{1}{2}\mathrm{T}\alpha$ and use (1) to determine b and c). We also have $\alpha'\beta' = \beta'\alpha'$, since if α and β commute then so do α' and β', since they are polynomials in α and β. But then $2\alpha'\beta' = 0$, which means $\alpha' = 0$ or $\beta' = 0$, since $\mathrm{End}^0(E)$ has no zero divisors. We cannot have $\alpha' = 0$, since $\alpha \notin \mathbb{Q}$, so $\beta' = 0$, which implies $\beta \in \mathbb{Q}(\alpha)$. □

Remark 13.17. In the proofs of Theorem 13.15 and Lemma 13.16 we do not rely on any properties of $\mathrm{End}^0(E)$ that depend on the fact that it is the endomorphism algebra of an elliptic curve. Indeed, one can replace $\mathrm{End}^0(E)$ with any $\mathbb{Q}$-algebra A possessing an involution $\alpha \mapsto \hat{\alpha}$ that fixes every element of $\mathbb{Q}$ such that the associated norm $\mathrm{N}\alpha = \alpha\hat{\alpha}$ maps nonzero elements of A to positive elements of $\mathbb{Q}$; all the other properties of $\mathrm{End}^0(E)$ that we used can be derived from these.

Theorem 13.18. *If E and E' are isogenous elliptic curves then $\mathrm{End}^0(E) \simeq \mathrm{End}^0(E')$.*

Proof. Let $\phi\colon E \to E'$ be an isogeny. For any $\alpha' \in \mathrm{End}(E')$ we may have an endomorphism $\alpha := \hat{\phi}\alpha'\phi \in \mathrm{End}(E)$. The degree of α is

$$\deg \alpha = \deg(\hat{\phi}\alpha'\phi) = (\deg \phi)^2(\deg \alpha')$$

and the trace of α is

$$\mathrm{tr}\,\alpha = \alpha + \hat{\alpha} = \hat{\phi}\alpha'\phi + \widehat{\hat{\phi}\alpha'\phi} = \hat{\phi}\alpha'\phi + \hat{\phi}\hat{\alpha}'\phi = \hat{\phi}(\alpha' + \hat{\alpha}')\phi = \hat{\phi}(\mathrm{tr}\,\alpha')\phi = (\deg\phi)\,\mathrm{tr}\,\alpha'.$$

It follows that the minimal polynomial of α is the same as the minimal polynomial of $(\deg\phi)\alpha'$, thus

$$\mathbb{Q}(\alpha) \simeq \mathbb{Q}((\deg\phi)\alpha') = \mathbb{Q}(\alpha').$$

This holds for every $\alpha' \in \mathrm{End}(E')$, so $\mathrm{End}^0(E')$ is isomorphic to a $\mathbb{Q}$-subalgebra of $\mathrm{End}^0(E)$. Applying the same argument to $\hat{\phi}\colon E' \to E$ shows that $\mathrm{End}^0(E)$ is isomorphic to a $\mathbb{Q}$-subalgebra of $\mathrm{End}(E)$ and therefore $\mathrm{End}^0(E) \simeq \mathrm{End}^0(E)$. □

Warning 13.19. The endomorphism ring $\mathrm{End}(E)$ is not an isogeny invariant.

Having classified the possible endomorphism algebras $\mathrm{End}^0(E)$, our next task is to classify the possible endomorphism rings $\mathrm{End}(E)$. We begin with the following corollary to Theorem 13.15.

Corollary 13.20. *Let E/k be an elliptic curve. The endomorphism ring $\mathrm{End}(E)$ is a free $\mathbb{Z}$-module of rank r, where $r = 1, 2, 4$ is the dimension of $\mathrm{End}^0(E)$ as a $\mathbb{Q}$-vector space.*

Recall that a free $\mathbb{Z}$-module of rank r is an abelian group that is isomorphic to $\mathbb{Z}^r$.

Proof. Let us pick a basis $\{e_1, \dots, e_r\}$ for $\mathrm{End}^0(E)$ as a $\mathbb{Q}$-basis with the property that $\mathrm{T}(e_ie_j) = 0$ unless $i = j$ (so use the basis $\{1, \alpha\}$ when $\mathrm{End}^0(E) = \mathbb{Q}(\alpha)$ and $\{1, \alpha, \beta, \alpha\beta\}$ when $\mathrm{End}^0 = \mathbb{Q}(\alpha, \beta)$, where α and β are constructed as in the proof of Theorem 13.15). After multiplying by suitable integers if necessary, we can assume without loss of generality that $e_1, \dots e_r \in \mathrm{End}(E)$ (this doesn't change $\mathrm{T}(e_ie_j) = 0$ for $i \neq j$).

Let $A \subseteq \mathrm{End}(E)$ be the $\mathbb{Z}$-module spanned by $\{e_1, \dots, e_r\}$, and define the *dual* $\mathbb{Z}$-module

$$A^* := \{\alpha \in \mathrm{End}^0(E) : \mathrm{T}(\alpha\phi) \in \mathbb{Z}\ \forall \phi \in A\}.$$

This definition makes sense for any $\mathbb{Z}$-module $B \subseteq \mathrm{End}^0(E)$, including $B = \mathrm{End}(E)$, and we note that $A \subseteq B$ implies $B^* \subseteq A^*$ (making A bigger makes A^* smaller). We also note that $\mathrm{End}(E) \subseteq \mathrm{End}(E)^*$, since $T(\alpha\phi) \in \mathbb{Z}$ for all $\alpha, \phi \in \mathrm{End}(E)$. Thus

$$A \subseteq \mathrm{End}(E) \subseteq \mathrm{End}(E)^* \subseteq A^*.$$

If we write $\alpha \in A^* \subseteq \mathrm{End}^0(E)$ as $a_1e_1 + \cdots + a_re_r$ then $\mathrm{T}(\alpha e_i) = a_i\mathrm{T}(e_i^2) \in \mathbb{Z}$. If follows that a_i is an integer multiple of $1/\mathrm{T}(e_i^2)$, so $\{e_1/\mathrm{T}(e_1^2), \dots, e_r/\mathrm{T}(e_r^2)\}$ is a basis for A^* as a $\mathbb{Z}$-module. Thus A^* is a free $\mathbb{Z}$-module of rank r, as is A (both are torsion free since $\mathrm{End}^0(E)$ is). It follows that $\mathrm{End}(E)$ (and $\mathrm{End}(E)^*$) are also be free $\mathbb{Z}$-modules of rank r. □

13.5 Orders in $\mathbb{Q}$-algebras

Definition 13.21. Let K be a $\mathbb{Q}$-algebra of finite dimension r as a $\mathbb{Q}$-vector space. An *order* $\mathcal{O}$ in K is a subring of K that is a free $\mathbb{Z}$-module of rank r. Equivalently, $\mathcal{O}$ is a subring of K that is finitely generated as a $\mathbb{Z}$-module and for which $K = \mathcal{O} \otimes_{\mathbb{Z}} \mathbb{Q}$.

Note that an order is required to be both a full lattice (free $\mathbb{Z}$-module of maximal rank) *and* a ring; in particular it must contain 1.

Example 13.22. The integers $\mathbb{Z}$ are the unique example of an order in $\mathbb{Q}$. Non-examples include the even integers, which is a lattice but not a ring, and the set $\{a/2^n : a, n \in \mathbb{Z}\}$, which is a ring but not a lattice (because it is not finitely generated as a $\mathbb{Z}$-module).

It follows from Corollary 13.20 that the endomorphism ring $\mathrm{End}(E)$ is an order in the $\mathbb{Q}$-algebra $\mathrm{End}^0(E)$. Note that if $\mathrm{End}^0(E) = \mathbb{Q}$, then we must have $\mathrm{End}(E) = \mathbb{Z}$, but in general there are many infinitely many possible orders that $\mathrm{End}(E)$ might be.

Corollary 13.23. *If E is an elliptic curve with complex multiplication then $\mathrm{End}(E)$ is either an order in an imaginary quadratic field, or an order in a quaternion algebra.*

Every order lies in some *maximal order* (an order that is not contained in any other); this follows from an application of Zorn's lemma, using the fact that elements of an order necessarily have monic minimal polynomials. In general, maximal orders need not be unique, but when the $\mathbb{Q}$-algebra K is a number field (a finite extension of $\mathbb{Q}$), this is the case. In view of Corollary 13.23, we are primarily interested in the case where K is an imaginary quadratic field, but it is just as easy to prove this for all number fields. We first need to recall a few standard results from algebraic number theory.[2]

Definition 13.24. An *algebraic number* α is a complex number that is the root of a polynomial with coefficients in $\mathbb{Q}$. An *algebraic integer* is a complex number that is the root of a monic polynomial with coefficients in $\mathbb{Z}$.

Two fundamental results of algebraic number theory are (1) the set of algebraic integers in a number field form a ring, and (2) every number field has an *integral basis* (a basis whose elements are algebraic integers). The following theorem gives a more precise statement.

Theorem 13.25. *The set of algebraic integers $\mathcal{O}_K$ in a number field K form a ring that is a free $\mathbb{Z}$-module of rank r, where $r = [K : \mathbb{Q}]$ is the dimension of K as a $\mathbb{Q}$-vector space.*

Proof. See Theorem 2.1 and Corollary 2.30 in [1] (or Theorems 2.9 and 2.16 in [3]).[3] □

Theorem 13.26. *The ring of integers $\mathcal{O}_K$ of a number field K is its unique maximal order.*

Proof. The previous theorem implies that $\mathcal{O}_K$ is an order. To show that it is the unique maximal order, we need to show that every order $\mathcal{O}$ in K is contained in $\mathcal{O}_K$. It suffices to show that every $\alpha \in \mathcal{O}$ is an algebraic integer. Viewing $\mathcal{O}$ as a $\mathbb{Z}$-lattice of rank $r = [K : \mathbb{Q}]$, consider the sublattice generated by all powers of α. Let $[\beta_1, \ldots, \beta_r]$ be a basis for this sublattice, where each β_i is a $\mathbb{Z}$-linear combination of powers of α. Let n be an integer larger than any of the exponents in any of the powers of α that appear in any β_i. Then $\alpha^n = c_1\beta_1 + \cdots + c_r\beta_r$, for some $c_1, \ldots, c_n \in \mathbb{Z}$, and this determines a monic polynomial of degree n with α as a root. Therefore α is an algebraic integer. □

[2] Algebraic number theory is not a prerequisite for this course. We do presume some familiarity with imaginary quadratic fields, which are covered in most undergraduate algebra courses.

[3] The proof of the second part of this theorem is essentially the same as the proof of Corollary 13.20; instead of the reduced trace in $\mathrm{End}^0(E)$, one uses the trace map from K to $\mathbb{Q}$, which has similar properties.

Finally, we characterize the orders in imaginary quadratic fields, which are the number fields we are most interested in.

Theorem 13.27. *Let K be an imaginary quadratic field with ring of integers $\mathcal{O}_K$. The orders $\mathcal{O}$ in K are precisely the subrings $\mathbb{Z} + f\mathcal{O}_K$, where f is any positive integer.*

Proof. The maximal order $\mathcal{O}_K$ is a free $\mathbb{Z}$-module (a lattice) of rank 2 that contains 1, so it has a $\mathbb{Z}$-basis of the form $[1, \tau]$ for some $\tau \notin \mathbb{Z}$. Let $\mathcal{O} = \mathbb{Z} + f\mathcal{O}_K$. It is clear that $\mathcal{O}$ is a sub-lattice of $\mathcal{O}_K$ that properly contains $\mathbb{Z}$, hence it is of rank 2. The $\mathbb{Z}$-module $\mathcal{O}$ is a subset of the ring $\mathcal{O}_K$ and contains 1, so to show that $\mathcal{O}$ is a ring it suffices to show that it is closed under multiplication. So let $a + f\alpha$ and $b + f\beta$ be arbitrary elements of $\mathcal{O}$, with $a, b \in \mathbb{Z}$ and $\alpha, \beta \in \mathcal{O}_K$. Then

$$(a + f\alpha)(b + f\beta) = ab + af\beta + bf\alpha + f^2\alpha\beta = ab + f(a\beta + b\alpha + f\alpha\beta) \in \mathcal{O},$$

since $ab \in \mathbb{Z}$ and $(a\beta + b\alpha + f\alpha\beta) \in \mathcal{O}_K$. So $\mathcal{O}$ is a subring of K. To see that $\mathcal{O}$ is an order, note that $\mathcal{O} \otimes_{\mathbb{Z}} \mathbb{Q} = \mathcal{O}_K \otimes_{\mathbb{Z}} \mathbb{Q} = K$.

Now let $\mathcal{O}$ be any order in K. Then $\mathcal{O}$ is a rank-2 sub-lattice of $\mathcal{O}_K = [1, \tau]$ that contains 1, so $\mathcal{O}$ must contain an integer multiple of τ. Let f be the least positive integer for which $f\tau \in \mathcal{O}$. The lattice $[1, f\tau]$ lies in $\mathcal{O}$, and we claim that in fact $\mathcal{O} = [1, f\tau]$. Any element α of $\mathcal{O}$ must lie in $\mathcal{O}_K$ and is therefore of the form $\alpha = a + b\tau$ for some $a, b \in \mathbb{Z}$. The element $b\tau = \alpha - a$ then lies in $\mathcal{O}$, and the minimality of f implies that f divides b. Thus $\mathcal{O} = [1, f\tau] = \mathbb{Z} + f\mathcal{O}_K$. □

Remark 13.28. In the theorem above we never actually used the fact that the quadratic field K is imaginary; in fact, the theorem holds for real quadratic fields as well.

The integer f in Theorem 13.27 is called the *conductor* of the order $\mathcal{O} = \mathbb{Z} + f\mathcal{O}_K$. It is equal to the index $(\mathcal{O}_K : \mathcal{O})$, which is necessarily finite.

References

[1] J.S. Milne, *Algebraic number theory*, course notes, version 3.06, 2014.

[2] J.H. Silverman, *The arithmetic of elliptic curves*, second edition, Springer 2009.

[3] I. Stewart and D. Tall, *Algebraic number theory and Fermat's last theorem*, third edition, A.K. Peters, 2002.