THE TRACE FORMULA FOR COMPACT QUOTIENTS

JEREMY BOOHER WITH AN APPENDIX BY BRIAN CONRAD

These notes are for the 2013-2014 learning seminar on the Jacquet–Langlands correspondence. We discuss the trace formula in its simplest form: when the space $\underline{G}(F)\backslash\underline{G}(\mathbf{A}_F)$ is compact (for a connected reductive group $\underline{G}$ over a global field F). This allows a simple proof to be given, as the compactness removes the hard analysis and implies that there is no continuous spectrum.

Following the article by Gelbart and Jacquet [GJ], we prove the trace formula for the algebraic unit group of a central division algebra over a global field (see Theorem 2.3.1). It expresses the trace of convolution with a suitably nice function as a sum of orbital integrals. The key idea is to compute the trace using a kernel for the convolution function. This is easily carried out in §2 via the compactness, following a review of trace class operators and measures on quotients in §1.

This version of the formula is useful for proving the Jacquet–Langlands correspondence by matching orbital integrals, but it is not well-suited for explicit computations. In §3 we specialize to the case that the division algebra is a quaternion division algebra over $\mathbf{Q}$ ramified precisely at a prime p and at ∞. We define Hecke operators in this setting, reinterpret the automorphic forms under consideration as quaternionic modular forms, and relate them to supersingular elliptic curves over $\overline{\mathbf{F}}_p$. Finally in §4 we lift the supersingular elliptic curves to characteristic 0 and derive the Eichler Trace Formula (see §4.2); this expresses traces of Hecke operators on a space of modular forms in terms of class numbers of imaginary quadratic orders.

An appendix by Brian Conrad records modern proofs of several classical results for which we do not know a suitable modern reference.

I thank Iurie Boreico, Brian Conrad, Dan Litt, Niccolò Ronchetti, and Akshay Venkatesh for useful conversations and advice.

1. Analytic Preliminaries

In this section we review some facts about trace class operators and measures on quotients.

1.1. **Trace Class Operators.** A general reference for the material in this section is [RS, VI.6]. Let V be a Hilbert space with orthonormal basis $\{v_i\}$. A bounded linear operator $A : V \to V$ is *Hilbert-Schmidt* if $\sum_i \|Av_i\|^2 < \infty$; finiteness of this number is independent of the choice of $\{v_i\}$. We say that A is *trace class* if $\sum_{i=1}^{\infty}((A^*A)^{\frac{1}{2}}v_i, v_i) < \infty$; once again, such finiteness is independent of the choice of $\{v_i\}$. In the latter case, the sum

$$\sum_{i=1}^{\infty}(Av_i, v_i).$$

turns out to be absolutely convergent and independent of the choice of $\{v_i\}$; this sum is then called the *trace* of A, denote as $\mathrm{tr}(A)$. A basic fact is:

Proposition 1.1.1. *The composition of two Hilbert-Schmidt operators on V is trace class.*

Let (X, μ) be a σ-finite measure space with complete measure μ (so there is an unambiguous complete measure $\mu \times \mu$ on $X \times X$, and Fubini's theorem holds relative to this). On $V = L^2(X) :=$

Date: February 6 and 13, 2014.

$L^2(X,\mu)$ Hilbert-Schmidt and trace class operators can be described as operators associated to kernel functions. To be precise, for any $K \in L^2(X \times X)$ and $f \in L^2(X)$, Fubini's theorem ensures that

$$(T_K f)(x) = \int_{X \times X} K(x,y) f(y) dy$$

converges absolutely for all $x \in X$ away from a set of measure 0 and defines a measureable function on X, with $T_K f \in L^2(X)$. The linear operator T_K on $L^2(X)$ is bounded, with operator norm at most $\|K\|_2$.

Proposition 1.1.2. *A bounded linear T on $L^2(X)$ is Hilbert-Schmidt if and only if $T = T_K$ for some $K \in L^2(X \times X)$.*

One can compute the trace of such an integral operator that is trace class:

Lemma 1.1.3. *Let T be a trace class operator on $L^2(X)$ arising from a continuous kernel $K(x,y)$. Then*

$$\mathrm{tr}(T) = \int_X K(x,x) dx.$$

Remark 1.1.4. This is a standard result and is easy to prove, and will suffice for our purposes. However, from a functional analysis viewpoint it is very restrictive to require that the kernel is continuous since a typical Hilbert-Schmidt operator on $L^2(X)$ has a kernel that is merely L^2 on $X \times X$ and hence is only defined almost everywhere on $X \times X$ yet the diagonal is typically of measure 0.

If the σ-finite measure space (X,μ) is countably generated (meaning that the σ-algebra generated by a countable collection of sets, such as a Borel measure on a separable topological space) then one can average in an appropriate way to associate to any $K \in L^2(X \times X)$ a certain pointwise function $\widetilde{K}(x,x)$ defined almost everywhere on the diagonal such that it is measureable and integrable over the diagonal if and only if T_K is of trace class, in which case $\int_X \widetilde{K}(x,x)dx = \mathrm{tr}(T_K)$; see [Br].

1.2. **Invariant Measures on Homogeneous Spaces.** A reference for the material in this section is [Na, III.4]. Let G be a Hausdorff locally compact topological group, and H a closed subgroup. Pick left Haar measures μ_G and μ_H on G and H respectively. Up to positive scaling, these are the unique nonzero left-invariant σ-regular positive Borel measures on G and H. Let the continuous modulus functions for G and H be denoted by Δ_G and Δ_H respectively. (By definition, $\mu_G(Ag) = \Delta_G(g)\mu_G(A)$ for all Borel sets A in G, and similarly for H.) We call the group *unimodular* when the modulus function is identically 1 (equivalently, when left Haar measures are also right-invariant.)

Elementary examples of unimodular groups are compact groups and discrete groups. To give a wider (and very useful) class of examples, let $\underline{G}$ be a smooth group scheme over a field k that is either local (possibly $\mathbf{R}$ or $\mathbf{C}$) or global, with $G = \underline{G}(k)$ in the local case and $G = \underline{G}(\mathbf{A}_k)$ in the global case. Let $|\cdot|$ denote the normalized absolute value on k in the local case (the one which arises in the Changes of Variable formula for a Haar measure on any Euclidean space over k) and the adelic norm in the global case.

The vector space of top-degree left-invariant differential forms on $\underline{G}$ is 1-dimensional over k, and is an algebraic representation space for $\underline{G}$ via the right-translation action of $\underline{G}$ on itself. For any nonzero top-degree left-invariant differential ω on $\underline{G}$ we obtain a left-invariant density $|\omega|$ on G, so we may use this to define a left Haar measure $\mu_{|\omega|}$ on G via the Change of Variables Formula on Euclidean spaces over local fields. Thus, the modulus function on G has the form $|\chi|$ for a canonical algebraic character $\chi : \underline{G} \to \mathrm{GL}_1$ that is trivial on the center of $\underline{G}$. In particular, unimodularity holds for G whenever $\underline{G}^0$ is reductive or unipotent.

Here is an extremely useful Fubini-like theorem for coset spaces.

Theorem 1.2.1. *For all $f \in C_c(G)$, the function $f^H : \overline{x} \mapsto \int_H f(xh)d\mu_H(h)$ on G/H is continuous and compactly supported, and all elements of $C_c(G/H)$ arise in this way.*

Assume $\Delta_G|_H = \Delta_H$. There is a unique σ-regular left G-invariant measure $\mu_{G/H}$ on G/H such that for all $f \in C_c(G)$,

$$\int_G f(x)d\mu_G(x) = \int_{G/H}\left(\int_H f(xh)d\mu_H(h)\right)d\mu_{G/H}(xH).$$

Moreover, if f is measureable function on G then the integral $f^H(\overline{x})$ converges absolutely for almost all $\overline{x} \in G/H$ and f^H is measurable on G/H, with $f^H \in L^1(G/H, \mu_{G/H})$ if and only if $f \in L^1(G, \mu_G)$, in which case the above integration formula holds.

A useful situation is when H is discrete (with counting measure as its Haar measure) and G is unimodular. For example, $H = \underline{G}(F)$ and $G = \underline{G}(\mathbf{A}_F)$ for a global field F. We also record a useful corollary regarding integrating invariant functions.

Corollary 1.2.2. *Suppose H is compact and $f \in C_c(G)$ is right-invariant by H. Then*

$$\int_G f(x)d\mu_G(x) = \mathrm{vol}(H)\int_{G/H} f(xH)d\mu_{G/H}(xH)$$

where $\mathrm{vol}(H)$ is the volume of H with respect to μ_H.

2. A Trace Formula for Division Algebras

In this section we will prove the simplest case of the trace formula, following [GJ].

2.1. **Definitions.** Let F be a global field and D be a central division algebra over F with $[D : F] = n^2$. Let $\underline{G}$ be the connected reductive F-group representing the functor on F-algebras R given by

$$\underline{G}(R) = (D \otimes_F R)^\times.$$

Informally, one often sees it said that $\underline{G}$ is "$D^\times$ as an F-group" or is "the F-group associated to $D^\times$". The center $Z \subset \underline{G}$ is the evident F-subgroup $\mathrm{GL}_1 \subset \underline{G}$ obtained as the Zariski closure of $F^\times \subset D^\times = \underline{G}(F)$ in $\underline{G}$. The adjoint semisimple quotient $\underline{G}/Z$ is denoted $\underline{G}^{\mathrm{ad}}$.

Since Z is a *split* F-torus, it is an instructive exercise with Hilbert 90 over fields and "spreading out" over S-integer rings to show that the natural continuous map $Z(\mathbf{A}_F)\backslash\underline{G}(\mathbf{A}_F) \to \underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$ is a homeomorphism (in particular, surjective); this avoids confusion that may otherwise occur when working with adelic points of coset schemes for algebraic groups.

For calculations it is convenient to use a Haar measure on $\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$ that is a "restricted direct product" of Haar measures on the local factors $\underline{G}^{\mathrm{ad}}(F_v)$ for places v of F. However, we will use an arbitrary choice and later shall indicate how to chose other measures for compatibility.

Observe that $\underline{G}^{\mathrm{ad}}$ is F-anisotropic since D is a central division algebra over F. Indeed, maximal k-tori in the algebraic unit group $\underline{A}^\times$ of a central simple algebra A over a field k are in bijective correspondence with maximal étale commutative k-subalgebras C of A, with the associated k-torus isomorphic to $\mathrm{R}_{C/k}(\mathrm{GL}_1)$. Hence, the maximal k-tori in $\underline{A}^\times/\mathrm{GL}_1$ have the form $\mathrm{R}_{C/k}(\mathrm{GL}_1)/\mathrm{GL}_1$. If A is a central *division* algebra then $C = k'$ is a separable field extension of k, so the maximal k-tori of $\underline{A}^\times/\mathrm{GL}_1$ have the form $\mathrm{R}_{k'/k}(\mathrm{GL}_1)/\mathrm{GL}_1$. By inspection, the corresponding Galois lattice has no nonzero Galois-invariant elements, so such quotient tori are k-anisotropic. The established global anisotropicity of $\underline{G}^{\mathrm{ad}}$ implies that $\underline{G}^{\mathrm{ad}}(F)\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$ is compact, as a special case of:

Proposition 2.1.1 (Mostow)**.** *Let $\underline{H}$ be a connected reductive over a global field k. If $\underline{H}$ is k-anisotropic then $\underline{H}(k)\backslash\underline{H}(\mathbf{A}_k)$ is compact.*

A proof is given in the appendix.

The compactness of $\underline{G}^{\mathrm{ad}}(F)\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$ ensures that it is easy to work with functions on this quotient analytically. We will often view such functions on $\underline{G}(\mathbf{A}_F)$ subject to translation-invariance

under $\underline{G}(F)Z(\mathbf{A}_F)$. Actually, it will be convenient to consider a slightly weaker invariance property relative to the center Z of $\underline{G}$, as follows. For a continuous character

$$\omega : Z(F)\backslash Z(\mathbf{A}_F) = F^\times\backslash\mathbf{A}_F^\times \to S^1,$$

define $L^2(\omega, \underline{G})$ to be the Hilbert space of measurable functions on $\underline{G}(\mathbf{A}_F)$ (modulo those vanishing almost everywhere) such that for $z \in Z(\mathbf{A}_F)$ and $\gamma \in \underline{G}(F)$ we have

$$f(zg) = \omega(z)f(g) \quad \text{and} \quad f(\gamma g) = f(g)$$

(in the sense of functions defined almost everywhere on $\underline{G}(\mathbf{A}_F)$) and

$$\int_{\underline{G}^{\mathrm{ad}}(F)\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)} |f(g)|^2 d\overline{g} < \infty.$$

(Note that $|f|$ is $\underline{G}(F)Z(\mathbf{A}_F)$-invariant and $\underline{G}(F)Z(\mathbf{A}_F)\backslash\underline{G}(\mathbf{A}_F) = \underline{G}^{\mathrm{ad}}(F)\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$.)

Such f can be interpretted as L^2 sections of a line bundle. The map

$$\underline{G}(F)\backslash\underline{G}(\mathbf{A}_F) \to \underline{G}^{\mathrm{ad}}(F)\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$$

is a topological $Z(F)\backslash Z(\mathbf{A}_F)$-torsor, so its pushout along $\omega : Z(F)\backslash Z(\mathbf{A}_F) \to \mathbf{C}^\times$ is a line bundle $\mathcal{L}_\omega$ over $\underline{G}^{\mathrm{ad}}(F)\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$. The space of measureable global sections of $\mathcal{L}_\omega$ taken modulo those vanishing almost everywhere is exactly the space of f as introduced above but without the L^2-condition. This line bundle has a natural metric, and $L^2(\omega, \underline{G})$ is exactly the space of L^2 global sections of $\mathcal{L}_\omega$ (modulo those vanishing almost everywhere).

Remark 2.1.2. When doing analysis with this Hilbert space, it is cleanest to view elements as sections over $\underline{G}^{\mathrm{ad}}(F)\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$ since this is the space that they "naturally" live on and it is compact. However, they are usually treated as L^2 functions on $\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$ which transform in a constrained way. From the first viewpoint, one cannot directly apply a result like Lemma 1.1.3. We would need a generalization to sections of line bundles. By picking a fundamental domain for the actions of $Z(\mathbf{A}_F)$ and $\underline{G}(F)$ on $\underline{G}(\mathbf{A}_F)$, L^2 functions on the fundamental domain can be identified with $L^2(\omega, \underline{G})$ by extending via the transformation laws. Elements of $L^2(\omega, \underline{G})$ become functions on the fundamental domain via restriction. Then identify the fundamental domain with the quotient $\underline{G}^{\mathrm{ad}}(F)\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$. This allows Lemma 1.1.3 to be directly applied to functions on $\underline{G}^{\mathrm{ad}}(F)\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$ (modulo those vanishing almost everywhere). One can use this technique to extend analytic results to include sections of line bundles. We will do this implicitly from now on.

Let ρ_ω be the right regular representation of $\underline{G}(\mathbf{A}_F)$ on $L^2(\omega, \underline{G})$. By unimodularity of $\underline{G}^{\mathrm{ad}}$, it follows that $\rho_\omega(g)$ is unitary for all g. Compactness of $\underline{G}^{\mathrm{ad}}(F)\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$ makes it easy to check that ρ_ω is a *continuous* (unitary) representation of $\underline{G}(\mathbf{A}_F)$ on $L^2(\omega, \underline{G})$ with central character ω.

2.2. **Some Trace Class Operators.** Next we will define a class of operators on $L^2(\omega, \underline{G})$ for which we can define the trace. We would like to convolve with nice functions on $\underline{G}(\mathbf{A}_F)$ that are compactly supported modulo the center $Z(\mathbf{A}_F)$.

To be precise, recall that for each place v of F we have a central simple algebra $D_v = D \otimes_F F_v$ over F_v that is a matrix algebra for all but finitely many v. Let R_v be a maximal order of D_v (unique at the finitely many places where D_v is a division algebra, and unique up to $D_v^\times$-conjugacy elsewhere), so $K_v := R_v^\times$ is a maximal compact subgroup of $D_v^\times = \underline{G}(F_v)$. We impose a "coherence" condition on the choice of R_v's to ensure that $\prod_{v\nmid\infty} K_v$ is a compact *open* subgroup of the group of finite-adelic points $\underline{G}(\mathbf{A}_F^\infty)$, namely we choose a maximal order R in D (or really any $\mathbf{Z}$-lattice Λ in D whatsoever) and demand that the local choice R_v coincides with the completion localization of R at v for all but finitely many v (or equivalently, coincides with Λ_v for all but finitely many v). The center $F_v^\times = Z(F_v)$ of $D_v^\times$ shall be denoted as Z_v.

We consider the set $\mathcal{H}$ of functions that are finite linear combinations of functions ϕ on $\underline{G}(\mathbf{A}_F)$ of the form $\prod_v \phi_v$, where $\phi_v : \underline{G}(F_v) = D_v^\times \to \mathbf{C}$ satisfies the following properties:

- For all v, ϕ_v is smooth (if v is archimedean this means C^∞; if non-archimedean this means locally constant) with the image of its support in $D_v^\times/Z_v$ compact.
- For all v, $\phi_v(z_v g_v) = \omega_v(z_v)^{-1}\phi_v(g_v)$ for $z_v \in Z_v$ and $g_v \in D_v^\times$.
- For almost all v,
$$\phi_v(g_v) = \begin{cases} \omega_v^{-1}(z_v), & \text{if } g_v = k_v z_v \\ 0 & \text{otherwise} \end{cases}$$
where $g_v \in \underline{G}(F_v)$, $k_v \in K_v$, and $z_v \in Z_v$.

Note that $\phi(zg) = \omega(z)^{-1}\phi(g)$ for all $z \in Z(\mathbf{A}_F)$ and $g \in \underline{G}(\mathbf{A}_F)$, so ϕ is $Z(F)$-invariant (as ω is trivial on $Z(F)$).

For $\phi \in \mathcal{H}$, define an operator $\rho_\omega(\phi)$ on $L^2(\omega, \underline{G})$ by
$$\rho_\omega(\phi) = \int_{\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)} \phi(g)\rho_\omega(g)dg$$
as a continuous compactly-supported operator–valued integral. (Note that the integrand is $Z(\mathbf{A}_F)$-invariant, so it descends to a continuous operator-valued function on $Z(\mathbf{A}_F)\backslash\underline{G}(\mathbf{A}_F) = \underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$ that is compactly supported due to the hypotheses on ϕ.) In more concrete terms, for $f \in L^2(\omega, \underline{G})$ the evident formula $\phi(g)f(xg) = \phi(zg)f(x(zg))$ for $z \in Z(\mathbf{A}_F)$ implies that
$$\left(\rho_\omega(\phi)(f)\right)(x) = \int_{\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)} \phi(g)f(xg)dg$$
for almost all x. (Since the integrand descends to $\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$ as a measurable function with compact support, it is integrable on there due to the L^2-hypothesis on $|f|$.)

Let's show that $\rho_\omega(\phi)$ is an integral operator arising from a nice kernel. Define $K(x,y) = \sum_{\gamma \in \underline{G}^{\mathrm{ad}}(F)} \phi(x^{-1}\gamma y)$ for $x, y \in \underline{G}(\mathbf{A}_F)$; the sum "makes sense" since ϕ is $Z(F)$-invariant, and it only depends on the cosets of x and y in $X := \underline{G}(F)\backslash\underline{G}(\mathbf{A}_F) = \underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$. Since $|\phi|$ descends to a continuous compactly supported function on $\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$, and the discrete subset $\underline{G}^{\mathrm{ad}}(F)$ has finite intersection with any compact set, as we vary x and y inside compact subsets of $\underline{G}(\mathbf{A}_F)$ those $\gamma \in \underline{G}^{\mathrm{ad}}(F)$ for which $\phi(x^{-1}\gamma y) \neq 0$ consitute a *finite* set of possibilities independent of such x, y (kept within compact sets). Hence, $K(x,y)$ is a *locally finite* sum in (x,y), so it is continuous in (x,y). As such, we may and do view K as a continuous function on $X \times X$. The absolute value $|K|$ is $Z(\mathbf{A}_F)$-invariant in both variables, so it descends to a continuous function on $\overline{X} \times \overline{X}$ where $\overline{X} := \underline{G}^{\mathrm{ad}}(F)\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$ is compact.

By the Fubini theorem for coset spaces we have
$$\begin{aligned} \left(\rho_\omega(\phi)(f)\right)(x) &= \int_{\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)} \phi(x^{-1}g)f(g)dg \\ &= \int_{\underline{G}^{\mathrm{ad}}(F)\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)} \sum_{\gamma \in \underline{G}^{\mathrm{ad}}(F)} \phi(x^{-1}\gamma y)f(y)dy \\ &= \int_{\overline{X}} K(x,y)f(y)dy, \end{aligned}$$
where the integrand lies in $L^1(\overline{X})$. This shows that $\rho_\omega(\phi)$ an integral operator, modulo the technicalities mentioned in Remark 2.1.2.

Proposition 2.2.1. *The operator $\rho_\omega(\phi)$ is compact and of trace class.*

Proof. We first claim that $\rho_\omega(\phi)$ is Hilbert-Schmidt (and hence compact). It suffices to show that $|K|$ is L^2 on $\overline{X} \times \overline{X}$, which is obvious since it is continuous and $\overline{X}$ is compact. To show that $\rho_\omega(\phi)$ is of trace class, we use the fact that ϕ may be expressed as a finite sum of convolutions of the form $\phi_i * f_i$ where ϕ_i and f_i have the same properties to those of ϕ and f except that they

may only be m times differentiable at the archimedean places as opposed to smooth for some large m. This is explained in [DL, p. 199]. The idea is to use a characteristic function of a compact open subgroup of the finite adelic places of $\underline{G}^{\mathrm{ad}}$ and to produce the archimedean places via an argument using an elliptic operator coming from the group structure. The relationship between composition of integral operators and convolution of the corresponding kernel functions (when working on appropriate coset spaces for groups) implies that integrating against ϕ is a finite sum of compositions of Hilbert-Schmidt operators, so it is of trace class. □

Corollary 2.2.2. *The representation ρ_ω decomposes as a Hilbert direct sum of irreducibles, each occurring with finite multiplicity.*

Proof. This is an application of the spectral theorem for compact operators on Hilbert spaces: the existence of a non-trivial subrepresentation follows from the existence of a generalized eigenspace. The finite-dimensionality of the eigenspaces implies any subrepresentation occurs at most finitely many times. □

Let $m(\pi) \geq 0$ denote the multiplicity of an irreducible unitary representation π of $\underline{G}(\mathbf{A}_F)$ in ρ_ω (with central character ω) and $\pi(\phi)$ denote the composition of $\rho_\omega(\phi)$ with the inclusion of π into ρ_ω and with the projection to π. We can express the trace of the trace-class operator $\rho_\omega(\phi)$ on $L^2(\omega, \underline{G})$ as the absolutely convergent sum

$$\mathrm{tr}\rho_\omega(\phi) = \sum_\pi m(\pi)\mathrm{tr}(\pi(\phi)).$$

2.3. **The Trace Formula.** Using the kernel K, we shall compute the trace via Lemma 1.1.3 when $[D : F] = p^2$ for a prime $p \neq \mathrm{char}(F)$. The significance of the primality condition is that if $\alpha \in D^\times - F^\times$ (i.e., a representative for a nontrivial element in $D^\times/F^\times = \underline{G}^{\mathrm{ad}}(F)$) then the subfield $F(\alpha) \subset D$ has degree p and so is a *maximal* commutative subfield that is *separable* over F.

In the definition $K(x, y) = \sum_{\gamma \in \underline{G}^{\mathrm{ad}}(F)} \phi(x^{-1}\gamma y)$, isolate the term for $\gamma = e$: it is $\phi(x^{-1}y) = \phi(e)$ when $x = y$. This contributes the term $\mathrm{vol}(\underline{G}^{\mathrm{ad}}(F)\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A}_F))\phi(e)$ to the trace. (Here we have fixed a Haar measure on $\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$ and use counting measure on $\underline{G}^{\mathrm{ad}}(F)$, so the coset space $\underline{G}^{\mathrm{ad}}(F)\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$ is equipped with the associated quotient measure.) Therefore

(2.3.1)
$$\int_{\underline{G}^{\mathrm{ad}}(F)\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)} K(x, x)dx = \mathrm{vol}(\underline{G}^{\mathrm{ad}}(F)\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A}_F))\phi(e) + \int_{\underline{G}^{\mathrm{ad}}(F)\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)} \sum_{\gamma \neq e} \phi(x^{-1}\gamma x)dx.$$

Choose $\alpha \in D^\times - F^\times = \underline{G}(F) - Z(F)$ and let $L = F(\alpha) \subset D$. The centralizer of α in $\underline{G}$ is the maximal F-torus $\underline{L}^\times$ of $\underline{G}$. If α is conjugate to $\alpha' \in \underline{G}(F) - Z(F)$ then conjugation gives an F-isomorphism of fields $F(\alpha) \simeq F(\alpha')$ that sends α to α'. Conversely, if $F(\alpha)$ and $F(\alpha')$ are isomorphic with α being sent to α' then the Skolem–Noether theorem implies that α and α' are conjugate in $\underline{G}(F)$. Denote the quotient $\underline{L}^\times/Z$ by $T_L^\times$, a maximal F-torus of $\underline{G}^{\mathrm{ad}}$. We will use this interpretation to understand the remaining terms of the sum.

Let X be a set of representatives for degree-p field extensions of F that embed into D, up to F-isomorphism. Regard the elements $L \in X$ as equipped with a choice of F-embedding in D. Any nontrivial element $\beta \in \underline{G}^{\mathrm{ad}}(F)$ is conjugate to an element $\alpha \in L^\times/F^\times$ for some $L \in X$, and hence can be written in the form

$$\beta = \eta^{-1}\alpha\eta$$

for $\eta \in \underline{G}^{\mathrm{ad}}(F)$ that is uniquely determined through its coset in $L^\times\backslash D^\times = (T_L^\times\backslash\underline{G}^{\mathrm{ad}})(F)$. This description depends on the choice of α (which generates the degree-p extension L of F), and the

number of choices is the number n_L of F-automorphisms of L. Thus, we may rewrite $K(x,x)$ as

$$K(x,x) = \phi(e) + \sum_{L\in X} \frac{1}{n_L} \sum_{\alpha\in L^\times/F^\times - [1]} \sum_{\eta\in L^\times\backslash \underline{G}(F)} \phi(x^{-1}\eta^{-1}\alpha\eta x).$$

Substituting into (2.3.1) we see that

(2.3.2)

$$\int_{\underline{G}^{\mathrm{ad}}(F)\backslash \underline{G}^{\mathrm{ad}}(\mathbf{A}_F)} K(x,x)dx = \mathrm{vol}(\underline{G}^{\mathrm{ad}}(F)\backslash \underline{G}^{\mathrm{ad}}(\mathbf{A}_F))\phi(e) + \sum_{L\in X} n_L^{-1} \sum_{\alpha} \int_{T_L^\times(F)\backslash \underline{G}^{\mathrm{ad}}(\mathbf{A}_F)} \phi(x^{-1}\alpha x)dx.$$

The last step groups the sum over $L^\times\backslash \underline{G}(F)$ to convert the integral over $\underline{G}^{\mathrm{ad}}(F)\backslash \underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$ into one over $T_L^\times(F)\backslash \underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$. There are no problems rearranging the sum because of the Fubini theorem on coset spaces. It is essential that we use the quotient measure on $T_L^\times(F)\backslash \underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$.

Now the integrand depends only on the class of x modulo $T_L^\times(\mathbf{A}_F)$ because any $y \in T_L^\times(\mathbf{A}_F)$ lies in the same commutative subgroup of $\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$ as α does. Therefore we can write

$$\int_{T_L^\times(F)\backslash \underline{G}^{\mathrm{ad}}(\mathbf{A}_F)} \phi(x^{-1}\alpha x)dx = \mathrm{vol}(\mathbf{A}_F^\times L^\times(F)\backslash \mathbf{A}_L^\times) \int_{T_L^\times(\mathbf{A}_F)\backslash \underline{G}^{\mathrm{ad}}(\mathbf{A}_F)} \phi(x^{-1}\alpha x)dx,$$

using the quotient measure. Substituting into (2.3.2), we obtain the following theorem.

Theorem 2.3.1. *With the previous notation and hypotheses,*

$$\sum_{\pi} m(\pi)\mathrm{tr}(\pi(\phi)) = \mathrm{tr}(\rho_\omega(\phi)) = \mathrm{vol}(\underline{G}^{\mathrm{ad}}(F)\backslash \underline{G}^{\mathrm{ad}}(\mathbf{A}_F))\phi(e) + \sum_{L\in X} n_L^{-1}\,\mathrm{vol}(\mathbf{A}_F^\times L^\times\backslash \mathbf{A}_L^\times) \sum_{\alpha\in L^\times/F^\times - [1]} \int_{T_L^\times(\mathbf{A}_F)\backslash \underline{G}^{\mathrm{ad}}(\mathbf{A}_F)} \phi(x^{-1}\alpha x)dx$$

Remark 2.3.2. The volumes and integrals depend on the measures. The measures on $\underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$ and $F^\times(\mathbf{A}_F)\backslash L^\times(\mathbf{A}_F) = T_L^\times(\mathbf{A}_F)$ are chosen arbitrarily and the measure on $T_L^\times(\mathbf{A}_F)\backslash \underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$ is chosen to be the associated quotient measure.

If the measure on $T_L^\times(\mathbf{A}_F)\backslash \underline{G}^{\mathrm{ad}}(\mathbf{A}_F)$ is a "restricted direct product" of local quotient measures then the integral can be expressed as the product

$$\int_{T_L^\times(\mathbf{A}_F)\backslash \underline{G}^{\mathrm{ad}}(\mathbf{A}_F)} \phi(x^{-1}\alpha x)dx = \prod_v \int_{L_v^\times\backslash D_v^\times} \phi_v(x_v^{-1}\alpha x_v)dx_v.$$

3. Quaternionic Modular Forms and Supersingular Elliptic Curves

Let D be the quaternion division algebra over $\mathbf{Q}$ ramified at a prime p and ∞. In this section we will define Hecke operators using a compact open subgroup $K \subset \underline{G}^{\mathrm{ad}}(\mathbf{A})$ and interpret quaternionic modular forms of level K as functions on the finite set $\underline{G}^{\mathrm{ad}}(\mathbf{Q})\backslash \underline{G}^{\mathrm{ad}}(\mathbf{A})/K$. This set will be expressed in terms of supersingular elliptic curves, giving a connection with the classical subject of Brandt matrices (whose traces give traces of Hecke operators in the classical theory).

3.1. **Quaternionic Modular Forms.** An informal reference for this material is [Bu]. Let R be a maximal order in D. For unramified places v of D, $D_v := D\otimes_{\mathbf{Q}} \mathbf{Q}_v$ is $\mathbf{Q}_v$-isomorphic to the matrix algebra $\mathrm{Mat}_2(\mathbf{Q}_v)$. Since maximal orders in D_v are unique up to conjugacy, this isomorphism can be chosen so that the maximal order $R_v := R\otimes_{\mathbf{Z}} \mathbf{Z}_v$ is identified with the maximal order $\mathrm{Mat}_2(\mathbf{Z}_v)$ for all but finitely many finite places v. This ensures that $\prod_{v\neq p,\infty} R_v^\times$ is identified with a compact open subgroup of $\mathrm{GL}_2(\mathbf{A}^{p,\infty})$, so it yields an open subgroup $K(N)$ (analogous to using $\Gamma_0(N)$ for classical modular forms) as follows for fixed $N \geq 1$ relatively prime to p.

Definition 3.1.1. For $\ell \neq p$, define $K_{\ell,0} = \{x \in \mathrm{GL}_2(\mathbf{Z}_\ell) : x \equiv \left(\begin{smallmatrix} * & * \\ 0 & * \end{smallmatrix}\right) \bmod N\}$. Define $K_{p,0} = R_p^\times$, $K_{\infty,0} = D_\infty^\times$, and the open subgroup $K_0(N) = \prod_v K_{v,0} \subset \underline{G}(\mathbf{A})$. The compact open image of $K_0(N)$ in $\underline{G}^{\mathrm{ad}}(\mathbf{A})$ is denoted $\overline{K}_0(N)$.

Remark 3.1.2. Note that the definition of $K_0(N)$ depends on the choice of $R \subset D$ (and there is generally more than one $D^\times$-conjugacy class of such choices). It is also possible to define many variants of $K_0(N)$ associated to level structures analogous to $\Gamma_1(N)$ and other subgroups of $\mathrm{SL}_2(\mathbf{Z})$.

Now we define Hecke operators analogous to T_m for any m prime to p. Let $\eta_m \in (D \otimes \mathbf{A})^\times$ be the element that is the identity at all places v except those dividing m, where it is $\left(\begin{smallmatrix} m & 0 \\ 0 & 1 \end{smallmatrix}\right) \in \mathrm{GL}_2(\mathbf{Q}_v) = D_v^\times$. The $K_0(N)$-double coset defined by η_m decomposes as a disjoint union

$$K_0(N)\eta_m K_0(N) = \coprod_i \alpha_i K_0(N).$$

Let ϕ_m denote the characteristic function of $K_0(N)\eta_m K_0(N)$ (and its image in $\underline{G}^{\mathrm{ad}}(\mathbf{Q})\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A})$). The function ϕ_m is compactly supported of the type considered at the start of §2.2 with trivial central character ω.

Remark 3.1.3. We can apply the trace formula to compute $\operatorname{tr} \rho_\omega(\phi_m)$. What is the general form of the expression? The volume term $\mathrm{vol}(\underline{G}^{\mathrm{ad}}(\mathbf{Q})\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A}))$ is a Tamagawa number, which for a quaternion division algebras is 2 when using the Tamagawa measure on $\underline{G}^{\mathrm{ad}}(\mathbf{A})$ (arising from any choice of nonzero top-degree invariant differential on $\underline{G}^{\mathrm{ad}}$ over $\mathbf{Q}$) [V, Theorem 2.3].

The remaining terms are $\mathrm{vol}(\mathbf{A}^\times L^\times \backslash \mathbf{A}_L^\times)$ times orbital integrals. These latter volumes look like class numbers (the regulator is trivial because L is imaginary quadratic). The orbital integrals can be evaluated locally. This would take us far afield. Details on how to evaluate local orbital integrals for GL_2 using Bruhat-Tits trees (which applies to all places except p and ∞) are given in [K, §5]. The general shape of trace formula is therefore an expression involving class numbers. We will use a different approach to derive Eichler's trace formula in (4.2.3) below.

Suppose $f \in L^2(1, G)$ is an eigenfunction for $\rho_1(\phi_m)$ with eigenvalue λ, so

$$\begin{aligned}\lambda f = \int_{\underline{G}^{\mathrm{ad}}(\mathbf{Q})\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A})} \phi_m(h) f(gh)dh &= \int_{K_0(N)\eta_m K_0(N)} f(gh)dh \\ &= \sum \int_{K_0(N)} f(g\alpha_i h)dh = \mathrm{vol}(K_0(N)) \sum f(g\alpha_i).\end{aligned}$$

This makes $\rho_1(\phi_m)$ look like the classical Hecke operator. This calculation also shows that such functions descend to the quotient $\underline{G}^{\mathrm{ad}}(\mathbf{Q})\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A})/\overline{K}_0(N)$.

Definition 3.1.4. Let $\Omega(N) = \underline{G}^{\mathrm{ad}}(\mathbf{Q})\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A})/\overline{K}_0(N)$ and define $S_2^D(K_0(N), \mathbf{C})$ to be the quotient of the function space $\mathrm{Func}(\Omega(N), \mathbf{C})$ modulo the constant functions.

The set $\Omega(N)$ is obviously finite since the coset space $\underline{G}^{\mathrm{ad}}(\mathbf{Q})\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A})$ is compact and the group $\overline{K}_0(N)$ is open in $\underline{G}^{\mathrm{ad}}(\mathbf{A})$. Elements of $S_2^D(K_0(N), \mathbf{C})$ are called *quaternionic modular forms of level $K_0(N)$ and weight* 2 for D. Up to suitably normalizing the Hecke operators, the above calculation shows $T_m := \rho_1(\phi_m)$ acts on such a form f by

$$(T_m f)(g) = \sum_i f(\alpha_i g). \tag{3.1.1}$$

Remark 3.1.5. There are identifications of $S_2^D(K_0(N), \mathbf{C})$ with subspaces of various spaces of classical modular forms. These are special instances of the Jacquet–Langlands correspondence for GL_2. More precisely, there is a non-canonical isomorphism of Hecke modules

$$S_2^D(K_0(N), \mathbf{C}) \simeq S_0 \subset S_2(\Gamma_0(Np))$$

where S_0 is the subspace of forms new at p. In this situation, the constant functions killed off in the quotient S_2^D correspond to Eisenstein series under the Jacquet–Langlands correspondence.

Example 3.1.6. Let $N = 1$ and $p = 11$ and $m = 5$. What will the trace be in this case? We can compute the trace of T_5 on $S_2^D(K_0(1), \mathbf{C})$ by computing its trace on $S_2(\Gamma_0(11))$. This is a one dimensional space containing

$$\eta(z)^2\eta(11z)^2 = q - 2q^2 - q^3 + 2q^4 + q^5 + \dots$$

Thus the trace is 1. There is an Eisenstein series in the two dimensional space $M_2(\Gamma_0(11))$, with q-expansion

$$\frac{5}{12} + q + 3q^2 + 4q^3 + 7q^4 + 6q^5 + \dots$$

The trace on this space is 7. This is the number we will compute in the next section by working with all functions from $\Omega(K_0(1))$ into $\mathbf{C}$, not throwing out the (nonzero) constant functions.

3.2. **Brandt Matrices and Supersingular Elliptic Curves.** Let E be a supersingular elliptic curve over $\overline{\mathbf{F}}_p$. Since E is finite type over $\overline{\mathbf{F}}_p$ and the endomorphism ring $\mathrm{End}(E)$ is finitely generated over $\mathbf{Z}$, by writing $\overline{\mathbf{F}}_p$ as a direct limit of its finite subfields we may find a finite subfield $k \subset \overline{\mathbf{F}}_p$ and an elliptic curve E_0 over k such that $(E_0)_{\overline{\mathbf{F}}_p} \simeq E$ and $\mathrm{End}_k(E_0) = \mathrm{End}(E)$. In this way, we can access Tate's work on abelian varieties over finite fields to prove:

Proposition 3.2.1 (Deuring)**.** *All such E consistute a single isogeny class over $\overline{\mathbf{F}}_p$, and the common isomorphism class of their endomorphism algebras is that of D. Moreover, the order* $\mathrm{End}(E)$ *in* $\mathrm{End}^0(E)$ *is maximal for all such E.*

Beware that central simple algebras over global fields generally have many conjugacy classes of maximal orders (in contrast with the case over non-archimedean local fields), so the $D^\times$-conjugacy class of the order $\mathrm{End}(E)$ inside D is sensitive to the isomorphism class of E. A modern proof of Proposition 3.2.1 is given in the appendix.

Fix a supersingular elliptic curve E over $\overline{\mathbf{F}}_p$ with $R = \mathrm{End}(E)$ a maximal order in D. Consider the compact open subgroup $K = K_0(1)$ depending on the choice of R (i.e., $K_v = R_v^\times$ in $D_v^\times$ for all places $v \nmid \infty$ of $\mathbf{Q}$). We are going to relate $\underline{G}^{\mathrm{ad}}(\mathbf{Q})\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A})/\overline{K} = \Omega(1)$ to the set Σ of isomorphism classes of supersingular elliptic curves over $\overline{\mathbf{F}}_p$. The following result is mentioned in a letter of Serre [Se], and was probably known to Deuring in some form.

Theorem 3.2.2. *There is a bijection $\Sigma \to \underline{G}^{\mathrm{ad}}(\mathbf{Q})\backslash\underline{G}^{\mathrm{ad}}(\mathbf{A})/K$.*

Given a supersingular elliptic curve E', the idea is to associate a projective rank-1 right R-module $M := \mathrm{Hom}(E, E')$. One checks that this gives a bijection between the set of supersingular elliptic curves over $\overline{\mathbf{F}}_p$ up to isomorphism and the set of isomorphism classes of projective rank-1 right R-modules. Then one uses algebra to relate these modules to the desired double coset, mimicking the description of ideals in a number field in terms of ideles. The proof is given in the appendix.

Using this identification (relative to a fixed E), there is a concrete description of the Hecke operator T_m as follows. The description of T_m in (3.1.1) shows that $T_m f$ at $g \in \Omega(1)$ is the sum of f evaluated on the "neighboring" elements $g\alpha_i \in \Omega(1)$. If g corresponds to a supersingular curve E', then unwinding the identifications in the theorem shows that the $g\alpha_i$ correspond to degree m subgroup schemes of $E'[m]$. These are equivalent to degree-m isogenies out of E' taken up to an *automorphism* of the target. Thus to compute the trace of T_m on the finite dimensional space of complex-valued functions on $\Omega(1)$, we will count how many endomorphisms of order m there are (up to automorphism) for supersingular curves over $\overline{\mathbf{F}}_p$.

Classically, this data was encoded in terms of Brandt matrices, which were first studied purely in terms of quaternion algebras. That perspective is summarized in [Gr, §1-§2], which also reinterprets it in terms of supersingular elliptic curves. We use the latter viewpoint as our foundation:

Definition 3.2.3. Enumerate the set of isomorphism classes of supersingular elliptic curves over $\overline{\mathbf{F}}_p$ as $\{E_1, \dots, E_n\}$, and let $m > 0$ be an integer. The *Brandt matrix* $B(m)$ is the $n \times n$ matrix whose entry $B_{ij}(m)$ is the number of degree-m isogenies in $\mathrm{Hom}(E_i, E_j)$ taken up to the action of automorphisms of E_i and E_j.

Taking $i = j$, $B_{ii}(m)$ can be reinterpreted as the number of endomorphisms of E_i of degree m taken up to composition on either side with an automorphism of E_i, or equivalently as the number of elements of reduced-norm m in $R_i := \mathrm{End}(E_i)$ modulo left and right multiplication against $R_i^\times$.

Example 3.2.4. Let $p = 11$ and $m = 5$ as before. There are only two supersingular elliptic curves over $\overline{\mathbf{F}}_p$, with j-invariants 0 and 1728 and explicit Weierstrass equations, so theoretically it is possible to find the subgroup schemes of order 5 in $E_i[5]$ by hand. It would be easier to do this using the equivalent definition of the Brandt matrix in terms of the arithmetic of quaternion algebras. SAGE does this, showing

$$B(5) = \begin{pmatrix} 4 & 2 \\ 3 & 3 \end{pmatrix}.$$

This matrix gives the action of T_5 on the 2-dimensional space of functions on $\Omega(1)$ relative to the basis of point-mass functions. Note that the constant function is an eigenfunction with eigenvalue 6, so (as we saw before) the trace on the cusp forms is 1 and the total trace is 7; this includes a contribution from the Eisenstein series in $M_2(\Gamma_0(11))$.

4. The Eichler Trace Formula

To count supersingular elliptic curves over $\overline{\mathbf{F}}_p$ with an endomorphism of order m, we will try to lift the curve and endomorphism to characteristic zero. Such a lift will necessarily be CM when m is not a square, so this suggests a connection with class numbers. A careful analysis will give a proof of the Eichler trace formula in (4.2.3).

4.1. **Lifting supersingular elliptic curves.** Deuring's study of elliptic curves in positive characteristic was carried out at a time long before the advent of the methods of Serre and Tate involving deformation theory, p-divisible groups, and the isogeny theorem. His basic method was to use well-chosen liftings to CM elliptic curves in characteristic 0.

Let (E, f) be a supersingular elliptic curve over $\overline{\mathbf{F}}_p$ equipped with an endomorphism $f \notin \mathbf{Z}$, so $\mathscr{O} := \mathbf{Z}[f]$ is an order in an imaginary quadratic subfield $F \subset \mathrm{End}^0(E)$. Every such F has a unique p-adic place since $E[p^\infty]$ is connected of dimension 1 and

$$F_p := \mathbf{Q}_p \otimes_{\mathbf{Q}} F \hookrightarrow \mathbf{Q}_p \otimes_{\mathbf{Q}} \mathrm{End}^0(E) \to \mathrm{End}^0(E[p^\infty])$$

is injective.

We seek to lift (E, f) to a pair $(\widetilde{E}, \widetilde{f})$ over the valuation ring of a finite extension L of $W[1/p]$ where $W = W(\overline{\mathbf{F}}_p)$. By the faithfulness of the reduction map

$$\mathrm{End}_L(\widetilde{E}_L) = \mathrm{End}(\widetilde{E}) \to \mathrm{End}(E),$$

such a lift distinguishes an *injection*

$$F = \mathbf{Q} \otimes_{\mathbf{Z}} \mathscr{O} \hookrightarrow \mathrm{End}_L(\mathrm{Lie}(\widetilde{E}_L)) = L$$

that induces an injection $F_p \to L$. Thus, a necessary condition on L is that it contains the compositum $W[1/p] \cdot F_p$ over $\mathbf{Q}_p$.

This minimal field of definition for the generic fiber of such a lifting can always be achieved. To prove this, first note that it is harmless to enlarge the order $\mathbf{Z}[f]$ to its saturation in $\mathrm{End}^0(E)$, as every order in a quadratic number field is monogenic, so we fix an quadratic order $\mathscr{O} \subset \mathrm{End}(E)$ that is saturated[1] in $\mathrm{End}^0(E)$ and seek to lift the pair $(E, \mathscr{O})$ to an elliptic curve $\widetilde{E}$ over the valuation

[1] In the terminology of quaternion algebras, such an embedding is said to be *optimal*.

ring of $L := W[1/p] \cdot F_p$ such that the $\mathscr{O}$-action lifts to one on $\widetilde{E}$ respecting the canonical inclusion $\mathscr{O} \subset F \subset F_p \subset L$ via the action on generic fiber of the Lie algebra. Note that for any such lift, $\mathscr{O}$ is necessarily saturated in $\mathrm{End}^0_L(\widetilde{E}_L) = \mathrm{End}^0(\widetilde{E})$ since if $h \in \mathrm{End}(\widetilde{E}) = \mathrm{End}_L(\widetilde{E}_L)$ satisfies $nh \in \mathscr{O}$ for some $n > 0$ then the same holds in $\mathrm{End}(E)$, forcing $h \in \mathscr{O}$ due to the saturatendess of $\mathscr{O}$ in $\mathrm{End}(E)$. It follows that $\mathscr{O} = \mathrm{End}(\widetilde{E})$ (as $\mathrm{End}^0_L(\widetilde{E}_L)$ is imaginary quadratic when larger than $\mathbf{Q}$, since $\mathrm{char}(L) = 0$).

Theorem 4.1.1 (Deuring Lifting Theorem)**.** *Let E be a supersingular elliptic curve over $\overline{\mathbf{F}}_p$ and $\mathscr{O} \subset \mathrm{End}(E)$ an order in the imaginary quadratic field F such that $\mathscr{O}$ is saturated in $\mathrm{End}^0(E)$ relative to $\mathrm{End}(E)$.*

For any finite extension L of $W[1/p] \cdot F_p$, there is a lift of $(E, \mathscr{O})$ over $\mathscr{O}_L$ compatible with the inclusion of $\mathscr{O}$ into $F \subset F_p \subset L$, and this lift is unique up to unique isomorphism.

Note that $W[1/p] \cdot F_p$ is the completion of the maximal unramified extension of F_p, equipped with an identification of its residue field with $\overline{\mathbf{F}}_p$. We restrict attention to the supersingular case because (i) it suffices for our needs, and (ii) in the ordinary case one has to introduce some more notation to do bookkeeping with p-adic places on F. The appendix contains a proof of Theorem 4.1.1 using the Serre–Tate deformation theorem to convert the lifting problem into a problem of lifting an associated formal group.

Now we refine the Deuring Lifting Theorem as a dictionary between certain isomorphism classes of CM elliptic curves in characteristic 0 and supersingular elliptic curves in characteristic p. This requires paying careful attention to ground fields and residue fields.

Let F be an imaginary quadratic field with a unique p-adic place v, and let $\mathscr{O}$ be an order in F that is *maximal* at p. Let L be a finite extension of the completion of F_p^{un} (equipped with an identification j of its residue field with $\overline{\mathbf{F}}_p$). Consider elliptic curves E over L with good reduction equipped with an embedding $\iota : \mathscr{O} \hookrightarrow \mathrm{End}_L(E)$ that is "normalized" in the sense that the induced action on the 1-dimensional Lie algebra over L is the canonical embedding $\mathscr{O} \hookrightarrow F \hookrightarrow L$.

For any such pair $(E, \mathscr{O})$ over L, the reduction E_0 over $\overline{\mathbf{F}}_p$ is supersingular since the p-divisible group of an ordinary elliptic curve over $\overline{\mathbf{F}}_p$ is $\mu_{p^\infty} \times (\mathbf{Q}_p/\mathbf{Z}_p)$, which does not support an action (in the isogeny category) by a quadratic p-adic field such as F_p. The induced embedding $\iota_0 : \mathscr{O} \to \mathrm{End}(E_0)$ is "normalized" in the sense that its action on the $\overline{\mathbf{F}}_p$-line $\mathrm{Lie}(E_0)$ is given by the natural map $\mathscr{O} \to \kappa(v) \hookrightarrow \overline{\mathbf{F}}_p$ (using j), since $\mathrm{Lie}(E_0)$ is an $\mathscr{O}$-linear quotient of the Lie algebra $\mathrm{Lie}(\mathcal{E})$ of the Néron model $\mathcal{E}$ over $\mathscr{O}_L$ (whose generic fiber $\mathrm{Lie}(E)$ has normalized $\mathscr{O}$-action by hypothesis).

Corollary 4.1.2. *Let $\mathscr{O}$ be an order in an imaginary quadratic field such that $\mathscr{O}$ is maximal at p. The construction $(E, \iota) \rightsquigarrow (E_0, \iota_0)$ is a bijection from the set of isomorphism classes of CM elliptic curves over L with good reduction and normalized $\iota : \mathscr{O} \hookrightarrow \mathrm{End}_L(E)$ to the set of isomorphism classes of supersingular elliptic curves over $\overline{\mathbf{F}}_p$ equipped with normalized $\iota_0 : \mathscr{O} \hookrightarrow \mathrm{End}(E_0)$.*

Proof. Given any (E_0, ι_0), to build a lift with normalized action it is harmless to increase $\mathscr{O}$ to its saturation in $\mathrm{End}(E_0)$, as this only makes the task harder. The saturated case is the content of the Deuring Lifting Theorem as formulated above.

It remains to show that if (E, ι) and (E, ι') have $\mathscr{O}$-linearly isomorphic reductions then such an isomorphism uniquely lifts to an $\mathscr{O}$-linear isomorphism between the Néron models over $\mathscr{O}_L$ (which is equivalent to giving an $\mathscr{O}$-linear isomorphism $E \simeq E'$ over L). Let $f_0 : E_0 \simeq E'_0$ is an $\mathscr{O}$-linear isomorphism over $\overline{\mathbf{F}}_p$. This identifies $\mathrm{End}(E_0)$ and $\mathrm{End}(E'_0)$, so it also identifies the saturations of $\mathscr{O}$ in both such maximal orders. Provided that the action of this saturation lifts (necessarily uniquely) to (the Néron models of) E and E', we may apply the uniqueness in the Deuring Lifting Theorem to conclude.

Our problem now concerns each pair separately. To be precise, we just have to check that for the saturation $\mathscr{O}' \subset F$ of $\mathscr{O}$ in $\mathrm{End}(E_0)$ inside $\mathrm{End}^0(E_0)$, the action of $\mathscr{O}'$ on E_0 lifts to an action

on (the Néron model of) E; such a lift is necessarily normalized via the embedding of $\mathscr{O}'$ into F since that property can be checked on the suborder $\mathscr{O}$. By the Serre–Tate deformation theorem and formal GAGA (as in the proof of the Deuring Lifting Theorem), to lift the action of $\mathscr{O}'$ to the Néron model over $\mathscr{O}_L$ it is equivalent to do so on the p-divisible group over $\mathscr{O}_L$. But the action on a p-divisible group only depends on $\mathscr{O}'$ through $\mathbf{Z}_p \otimes_{\mathbf{Z}} \mathscr{O}'$, and the maximality hypothesis on $\mathscr{O}$ at p ensures that the inclusion $\mathscr{O} \hookrightarrow \mathscr{O}'$ becomes an equality after p-adic completion. □

4.2. **Calculating Traces Using CM Elliptic Curves.** By lifting the supersingular elliptic curves to characteristic zero, we can express the trace of the Brandt matrix $B(m)$ in terms of counting CM elliptic curves.

Recall that the trace of $B(m)$ is the number of isogenies of degree m, up to automorphism, among the supersingular elliptic curves $E_1, \dots, E_n$ over $\overline{\mathbf{F}}_p$. Recall that R_i is the endomorphism ring of E_i over $\overline{\mathbf{F}}_p$. In the following argument, we will make a assumption that m is not a perfect square, so that such an isogeny $\alpha \in R_i$ generates a quadratic order $\mathbf{Z}[\alpha]$ in R_i. (The Eichler trace formula still holds when m is a perfect square.) Changing α by an automorphism of E_i (multiplying by a unit of R_i) changes the embedding by conjugation. Therefore the trace of $B(m)$ is the number of conjugacy classes of embeddings of quadratic orders $\mathbf{Z}[\alpha]$ in the R_i where α an element of norm m.

The field of fractions of $\mathbf{Z}[\alpha]$ must be an imaginary quadratic field F that is non-split at p. Indeed, the quaternion algebra D is ramified at p and ∞, so D_p and D_∞ are division algebras, and since F_p must embed in D_p we get the non-split assertion at p. If α has minimal polynomial $x^2 - tx + m$ then $\mathbf{Z}[\alpha]$ has discriminant $t^2 - 4m$, so $t^2 - 4m \leq 0$ (as F is imaginary quadratic).

Now let $\mathscr{O}_d$ be the unique imaginary quadratic order with discriminant d, and let $E(\mathscr{O}_d)$ denote the number of $R_i^\times$-conjugacy classes of embeddings of $\mathscr{O}_d$ into R_i for some i. We see that

$$\operatorname{tr}(B(m)) = \sum_{\substack{t \in \mathbf{Z} \\ t^2 - 4m \leq 0}} E(\mathscr{O}_{t^2-4m}).$$

Several observations are required before we attempt to lift. First, suppose that the embedding of $\mathscr{O}_d$ into R_i is saturated. Then the analysis before Theorem 4.1.1 shows that the lift of E_i over the field L has CM by the order $\mathscr{O}_d$ (as opposed to some larger order). Therefore, to lift and control the CM order we only need to consider saturated embeddings. Defining $E^*(\mathscr{O}_d)$ to be the number of saturated embeddings of $\mathscr{O}_d$ into some R_i (up to $R_i^\times$-conjugation), since every embedding extends to a unique saturated embedding of some larger order we see

$$\operatorname{tr}(B(m)) = \sum_{t \in \mathbf{Z}} \sum_{df^2 = t^2 - 4m} E^*(\mathscr{O}_d). \tag{4.2.1}$$

Secondly, observe that there are no saturated embeddings of an order whose conductor is a multiple of p. Indeed, if $\mathscr{O}$ were such an order then by considering the conductor we see that the localization $\mathscr{O}_p$ could not be the unique maximal order in $F_p \subset D_p$. This contradicts the embedding being globally saturated. Thus, we may and will apply Corollary 4.1.2.

Finally. consider the order $\mathscr{O}_d$ with $df^2 = t^4 - 4m$, and let $F = \operatorname{Frac}(\mathscr{O}_d)$. Take L to be as in Corollary 4.1.2 with a fixed map $\mathscr{O}_d \hookrightarrow F \hookrightarrow L$. Let $h(\mathscr{O}_d)$ be the size of $\operatorname{Pic}(\operatorname{Spec}(\mathscr{O}_d))$ (this agrees with the class number of the order defined in terms of the number of reduced binary quadratic forms of discriminant d up to equivalence). Recall that over the algebraic closure of L there are $h(\mathscr{O}_d)$ isomorphism classes of elliptic curves with $\operatorname{End}_L(E) = \mathscr{O}_d$. There is a unique way to identify $\mathscr{O}_d$ with the endomorphism ring so that the map is normalized. The curves and their endomorphism rings descend to some finite extension, so for sufficiently large L the number of pairs (E, ι) with ι normalized is $h(\mathscr{O}_d)$.

We now use Corollary 4.1.2 to calculate $E^*(\mathscr{O}_d)$. Note that the saturated embeddings come in pairs arising from complex conjugation of $\mathscr{O}_d$. Exactly one of each pair is normalized if p is inert in F, while both are if p is ramified. There are no saturated embeddings if p is split or if the conductor

is a multiple of p. Therefore defining ϵ_d to be 0 if p is split or the conductor is a multiple of p, 1 if p is inert, and $\frac{1}{2}$ if p is ramified, we see that

$$E^*(\mathscr{O}_d) = \epsilon_d h(\mathscr{O}_d).$$

Substituting into (4.2.1), we obtain

$$\mathrm{tr}(B(m)) = \sum_{t\in\mathbf{Z}} \sum_{df^2=t^2-4m} \epsilon_d h(\mathscr{O}_d) \tag{4.2.2}$$

We will rewrite this to recover Eichler's trace formula.

Definition 4.2.1. For an order $\mathscr{O}$ in an imaginary quadratic field of discriminant N, let $u(\mathscr{O})$ be the size of $\mathscr{O}^\times/\{\pm 1\}$. The *Hurwitz class number* $H(N)$ is defined as

$$H(N) = \sum_{df^2=N} \frac{h(\mathscr{O}_d)}{u(\mathscr{O}_d)}.$$

There is a modification of this concept that is notationally useful.

Definition 4.2.2. If p divides the conductor of $\mathscr{O}_d$, recursively define $H_p(d) = H_p(d/p^2)$. Otherwise set $H_p(d)$ to be 0 if p splits in $\mathscr{O}_d$, $H(d)$ if p is inert in $\mathscr{O}_d$, and $\frac{1}{2}H(d)$ if p is ramified but does not divide the conductor.

Then the inner sum in (4.2.2) is exactly $H_p(t^2-4m)$ because of the definition of ϵ_d, so we rewrite and obtain the Eichler trace formula

$$\mathrm{tr}B(m) = \sum_{\substack{t\in\mathbf{Z}\\ t^2-4m\le 0}} H_p(t^2-4m). \tag{4.2.3}$$

This statement is true of general m, but the proof given here works only for m not a perfect square.

Example 4.2.3. As before, take $p = 11$ and $m = 5$. Then the Eichler trace formula says that

$$\begin{aligned}
\mathrm{tr}B(5) &= H_{11}(-20) + 2H_{11}(-19) + 2H_{11}(-16) + 2H_{11}(-11) + 2H_{11}(-4)\\
&= H(-20) + 2H(-16) + 2\frac{1}{2}H(-11) + 2H(-4)\\
&= 2 + 2\cdot\frac{3}{2} + 2\cdot\frac{1}{2}\cdot 1 + 2\cdot\frac{1}{2} = 7.
\end{aligned}$$

This matches the earlier calculations.

Appendix A. Appendix

Brian Conrad provided the following proofs of several results for which we do not know of a suitable literature reference using modern techniques.

A.1. **Mostow's Result on Compactness.** We begin with a proof of Propostion 2.1.1. Recall $\underline{H}$ is a connected reductive group over a global field k that is k-anisotropic. We show that $\underline{H}(k)\backslash\underline{H}(\mathbf{A}_k)$ is compact.

Proof. The general case can be reduced to the adjoint semisimple case by bookkeeping arguments to handle central isogenies and quotients by central tori (the case of tori encodes Dirichlet's unit theorem). We now assume $\underline{H}$ is adjoint semisimple, the only case we need.

We first reduce to the case $F = \mathbf{Q}$ or $F = \mathbf{F}_p(t)$ (which will massively simplify some calculations) via a Weil restriction argument. To give some context for this, note that for any finite separable extension of fields k'/k and smooth connected affine k'-groups G' and $G := R_{k'/k}(G')$, the operation

$$T' \mapsto \mathrm{R}_{k'/k}(T')$$

defines a bijection between the set of maximal k'-tori in G' and the set of maximal k-tori in G (as one sees by extending scalars to k_s and using that $k' \otimes_k k_s$ is a direct product of copies of k_s). Since $\mathrm{Hom}_k(\mathrm{GL}_1, \mathrm{R}_{k'/k}(T')) = \mathrm{Hom}_{k'}(\mathrm{GL}_1, T')$, we see that T' is k'-anisotropic if and only if $\mathrm{R}_{k'/k}(T')$ is k-anisotropic. Thus, G' is k'-anisotropic if and only if G is k-anisotropic.

If k is a global field then topologically $G(\mathbf{A}_k) = G'(\mathbf{A}_{k'})$ carrying $G(k)$ onto $G'(k')$, so the coset spaces agree topologically too. Since every global field is finite separable over $\mathbf{Q}$ in characteristic 0 and over $\mathbf{F}_p(t)$ in characteristic $p > 0$, to verify the compactness property we can use Weil restriction to reduce to the case when the ground field is $\mathbf{Q}$ or $\mathbf{F}_p(t)$. (Note that this process destroys absolute simplicity if it held over the original global field, but it preserves the property of being semisimple of adjoint type.)

In characteristic 0, a nontrivial unipotent rational point generates a positive-dimensional unipotent group, which must contain a $\mathbf{G}_{\mathrm{a}}$ over the ground field. But it is a general fact that if a connected reductive group contains $\mathbf{G}_{\mathrm{a}}$ over the ground field then it contains GL_1 over the ground field (ultimately because maximal split connected unipotent subgroups are unipotent radicals of minimal parabolics over the ground field, and existence of a proper parabolic subgroup over the ground field amounts to isotropicity due to the Borel-Tits relative structure theory over fields). Hence, for an anisotropic connected reductive group over a field of characteristic 0, all rational points are *semisimple*. (This latter property is clear in the special case of $\underline{G}^{\mathrm{ad}}$!) For the rest of the argument, we focus on the number field case, so $F = \mathbf{Q}$; the technical advantage is that it allows us to access exponential maps. Let $\mathbf{A} = \mathbf{A}_{\mathbf{Q}}$.

Using the "algebraic" exponential map on nilpotent elements of the Lie algebra $\mathfrak{h}$ of the adjoint semisimple $\underline{H}$, it follows that there are no nontrivial nilpotent elements in $\mathfrak{h}$. Hence, by rationality of Jordan decomposition (valid in Lie algebras in characteristic 0), all elements of $\mathfrak{h}$ are semisimple. The adjoint representation $\mathrm{Ad}_{\underline{H}} : \underline{H} \to \mathrm{GL}(\mathfrak{h})$ has trivial kernel since $\underline{H}$ is of adjoint type, so $\mathrm{ad}_{\mathfrak{h}} = \mathrm{d}(\mathrm{Ad}_{\underline{H}})(e) : \mathfrak{h} \to \mathrm{End}(\mathfrak{h})$ is injective. Hence, any $X \in \mathfrak{h}$ has adjoint action with a nonzero eigenvalue (as X is semisimple).

Since $\mathrm{Ad}_{\underline{H}}$ is a closed immersion into $\mathrm{SL}(\mathfrak{h})$, it identifies $\underline{H}(\mathbf{A})$ as a closed subgroup of $\mathrm{SL}(\mathfrak{h}_{\mathbf{A}})$, which inherits the topology from $\mathrm{End}(\mathfrak{h}_{\mathbf{A}})$ since SL_n is closed in Mat_n for any $n \geq 1$. There is an important criterion of Mahler for proving compactness of adelic points modulo global points in the semisimple case by means of a faithful representation (such as $\mathrm{Ad}_{\underline{H}}$ in our setting). In the adelic setting over global fields this is [Oes, Prop. 1.2, Ch. IV] (which gives the statement and provides references for a proof), and we will recall the statement of this criterion below; the more classical setting of $\mathbf{R}$-points modulo an arithmetic subgroup amounts to reduction theory and Siegel sets using spaces of lattices, and as such is [Bo, Cor. 1.9].

The adelic Mahler criterion says that $\underline{H}(\mathbf{A})/\underline{H}(\mathbf{Q})$ is compact provided that whenever $\{h_n\}$ is a sequence in $\underline{H}(\mathbf{A})$ and $\{v_n\}$ is a sequence in $\mathfrak{h}$ such that $h_n v_n \to 0$ in $\mathfrak{h}_{\mathbf{A}}$ then $v_n = 0$ for $n \gg 0$. Consider the characteristic polynomial F_X of $\mathrm{ad}(X)$ on $\mathfrak{h}$, so for $d = \dim \underline{H}$ we have

$$F_X(T) = T^d + c_{d-1}(X)T^{n-1} + ... + c_0(X)$$

for regular functions $c_0, \ldots, c_{d-1}$ on the affine space $\mathfrak{h}$. Note that for all $h \in \underline{H}$, $F_{\mathrm{Ad}_{\underline{H}}(h)(X)} = F_X$ since $\mathrm{ad}(\mathrm{Ad}(h)(X))$ is the $\mathrm{Ad}(h)$-conjugate of $\mathrm{ad}(X)$ (as we check by using an inclusion of $\underline{H}$ into some GL_N and applying functoriality of ad and Ad). Hence, c_i is $\mathrm{Ad}_{\underline{H}}$-invariant for each i.

We saw above that if $X \in \mathfrak{h} - \{0\}$ then $\mathrm{ad}(X)$ is nonzero and semisimple, hence not nilpotent, so *some* $c_i(X) \in \mathbf{Q}$ is nonzero. Hence,

$$P(X) := \sum c_i(X)^2$$

is an $\mathrm{Ad}_{\underline{H}}$-invariant polynomial function on $\mathfrak{h}$ that is *nonzero* at all nonzero X in the $\mathbf{Q}$-vector space $\mathfrak{h}$. Here it is very useful that we are working over $\mathbf{Q}$.

We want to show $v_n = 0$ for $n \gg 0$. By continuity of polynomial functions, $P(h_n v_n) \to 0$ in $\mathbf{A}$. But the $\underline{H}$-action on $\mathfrak{h}$ is through $\mathrm{Ad}_{\underline{H}}$ and we've noted that P is $\mathrm{Ad}_{\underline{H}}$-invariant, so $P(v_n) \to 0$ in

A. Since $P(v_n)$ is in the *discrete* subring $\mathbf{Q} \subset \mathbf{A}$, $P(v_n) = 0$ for $n \gg 0$. We designed P so that it is non-zero at all $X \in \mathfrak{h} - \{0\}$, so $v_n = 0$ for $n \gg 0$. □

A.2. **Deuring's Results On Supersingular Elliptic Curves.** We now prove Proposition 3.2.1, which says that all supersingular elliptic curves over $\overline{\mathbf{F}}_p$ are isogenous and their endomorphism algebras are maximal orders in the quaternion algebra over $\mathbf{Q}$ ramified at p and ∞.

Proof. We will work with descents to a finite subfield $k \subset \overline{\mathbf{F}}_p$. Consider an elliptic curve E_0 over k that is a descent of a supersingular elliptic curve E over $\overline{\mathbf{F}}_p$. Let $q = \#k$, so the characteristic polynomial in $\mathbf{Z}[T]$ of the q-Frobenius endomorphism φ of E_0 is $\chi(T) := T^2 - aT + q$ with roots α, β that are algebraic integers whose archimedean absolute values are $\sqrt{q}$. The splitting field F of χ is either $\mathbf{Q}$ or quadratic. By replacing k with its extension of degree d inside $\overline{\mathbf{F}}_p$, we replace α and β with their dth powers.

We claim that by increasing k enough, we can arrange that $\alpha = \beta = p^n$ with $q = p^{2n}$ for some integer n. The key point is to show that F must have a *unique* p-adic place. This is obvious if $F = \mathbf{Q}$, and in the quadratic case if it fails then $F_p := \mathbf{Q}_p \otimes_{\mathbf{Q}} F$ is isomorphic to $\mathbf{Q}_p \times \mathbf{Q}_p$ as a $\mathbf{Q}_p$-algebra. Hence, via the two primitive idempotents, the *inclusion*

$$F_p \hookrightarrow \mathbf{Q}_p \otimes_{\mathbf{Q}} \mathrm{End}^0(E[p^\infty])$$

of $\mathbf{Q}_p$-algebras provides an isogeny decomposition of the p-divisible group of E as a direct product of two *non-zero* p-divisible groups, necessarily each of height 1 and connected (as E is *supersingular*, so the étale part of $E[p^\infty]$ vanishes). But by Dieudonné theory (see [CCO, 1.4.3] for a summary), the only connected p-divisible group of height 1 over $\overline{\mathbf{F}}_p$ is μ_{p^∞}, so $E[p^\infty]$ is isogenous to $\mu_{p^\infty} \times \mu_{p^\infty}$. Hence, $E[p]$ admits μ_p as a quotient, so the Cartier dual $E[p]^D$ admits $\mathbf{Z}/p\mathbf{Z}$ as a subgroup scheme. But the self-duality of E implies the self-duality of $E[p]$ [KM, 2.8.5], so we have a contradiction since $E[p]$ is infinitesimal (as E is supersingular).

Since F has a unique p-adic place, the p-adic absolute values of α and $\beta = q/\alpha$ coincide (treating the case $F = \mathbf{Q}$ separately via the archimedean information). Hence, α/β has absolute value 1 at *all* places of F, so it is root of unity. Replacing k with a suitable finite extension trivializes this root of unity and thereby brings us to the case that $\alpha = \beta$. In a complex embedding this must be a square root of q, so passing to a further quadratic extension of k if necessary finally brings us to the case $q = p^{2n}$ with $\alpha = \beta = p^n$ for some integer n.

By Tate's isogeny theorem for abelian varieties over finite fields ([Mu, App. I], [Ta1]), for any prime $\ell \neq p$ we have

$$\mathbf{Z}_\ell \otimes_{\mathbf{Z}} \mathrm{End}_k(E_0) \simeq \mathrm{End}_k(E_0[\ell^\infty]) = \mathrm{End}_{\mathrm{Gal}(\overline{k}/k)}(T_\ell(E_0)).$$

But the q-Frobenius in the Galois group acts on $T_\ell(E_0)$ acts through the effect of the q-Frobenius endomorphism φ that we have arranged to be $p^n \in \mathbf{Z}$, so the Galois-compatibility condition is vacuous and hence the right side is $\mathrm{Mat}_2(\mathbf{Z}_\ell)$. It follows that the associative algebra $\mathrm{End}_k(E_0)$ which is finite free as a $\mathbf{Z}$-module is non-commutative of rank 4 over $\mathbf{Z}$ with center equal to $\mathbf{Z}$, and moreover it *does not grow* as we increase k further, so $\mathrm{End}_k(E_0) = \mathrm{End}(E)$ and $\mathrm{End}^0(E)$ is a quaternion division algebra over $\mathbf{Q}$ that splits at all places away from p and ∞. Thus, $\mathrm{End}^0_k(E_0) = \mathrm{End}^0(E)$ is isomorphic to the unique quaternion division algebra D over $\mathbf{Q}$ with $\{p, \infty\}$ as its set of ramified places.

If E' is another supersingular elliptic curve over $\overline{\mathbf{F}}_p$, by increasing k some more if necessary we can arrange that E' also descends to an elliptic curve E'_0 over k whose q-Frobenius is p^n. By Honda–Tate theory, which classifies isogeny classes of simple abelian varieties over finite fields (see [CCO, 1.6.2] for a summary; this topic will be discussed in the spring), the isogeny class of an elliptic curve over a finite field is characterized by the Galois conjugacy class of the Frobenius endomorphism (viewed as an algebraic integer via its characteristic polynomial). Hence, E_0 and E'_0 are k-isogenous, so E and E' are isogenous.

It remains to prove that the order $\mathrm{End}(E) = \mathrm{End}_k(E_0)$ in $\mathrm{End}^0(E) = \mathrm{End}^0_k(E_0)$ is maximal, for which it is equivalent to check locally at all finite places of $\mathbf{Q}$. The case of places $\ell \neq p$ has been seen above (with $\mathbf{Z}_\ell \otimes_{\mathbf{Z}} \mathrm{End}(E) \simeq \mathrm{Mat}_2(\mathbf{Z}_\ell)$), so we just have to analyze $\mathbf{Z}_p \otimes_{\mathbf{Z}} \mathrm{End}(E) = \mathbf{Z}_p \otimes_{\mathbf{Z}} \mathrm{End}_k(E_0)$. Tate's isogeny theorem over finite fields is valid at $\ell = p$ too (see [CCO, A.1]), giving

$$\mathbf{Z}_p \otimes_{\mathbf{Z}} \mathrm{End}_k(E_0) \simeq \mathrm{End}_k(E_0[p^\infty]).$$

Thus, it suffices to show that the right side is a maximal order in its generic fiber division algebra over $\mathbf{Q}_p$. By Galois descent, the natural inclusion

$$j : \mathrm{End}_k(E_0[p^\infty]) \hookrightarrow \mathrm{End}(E[p^\infty])$$

identifies the left side with the subset of Galois-invariants on the right side. In particular, if the right side has generic fiber over $\mathbf{Q}_p$ with rank 4 then $j_{\mathbf{Q}_p}$ is an isomorphism and hence the Galois-action on $\mathrm{End}(E[p^\infty])$ is *trivial*, so j would be an *isomorphism*.

In contrast with varieties over fields, homomorphisms between geometric fibers of p-divisible groups over a field need *not* descend to a finite extension of the ground field! (One finds many étale counterexamples via Galois representations on $\mathbf{Z}_p$-lattices.) So it is not evident that j should become an isomorphism via increasing k to a finite extension. Nonetheless, we shall prove that this is what happens in our case.

More specifically, we claim that if Γ is *any* p-divisible group of height 2 and dimension 1 over $\overline{\mathbf{F}}_p$ then $\mathrm{End}(\Gamma)[1/p]$ is a quaternion division algebra over $\mathbf{Q}_p$ in which $\mathrm{End}(\Gamma)$ is moreover a maximal order. By the Serre–Tate equivalence between connected p-divisible groups and finite-dimensional commutative formal groups of finite height over complete local noetherian rings with characteristic p (see [Ta2, §2.2]), the study of such Γ is the "same" as that of a 1-dimensional commutative formal group over $\overline{\mathbf{F}}_p$ with height 2. Up to isomorphism there is *exactly one* such formal group with each height $h \geq 1$ over algebraically closed fields of characteristic p (e.g., μ_{p^∞} for $h = 1$); see [H, 21.9.1(i)] for a proof. Hence, it suffices to study *one* choice of Γ, as all choices are isomorphic. See [H, 20.2.14] (and preceding discussion there) for a proof working with a specific Γ. □

A.3. **Quaternions and Supersingular Elliptic Curves.** We provide a proof of Theorem 3.2.2.

Let E be a supersingular elliptic curve over $\overline{\mathbf{F}}_p$, and fix an isomorphism $\mathrm{End}^0(E) \simeq D$, so $\mathrm{End}(E)$ is identified with a maximal order R in D. Let E' be another supersingular elliptic curve over $\overline{\mathbf{F}}_p$, so E' is isogenous to E. Hence, for the right R-module $M := \mathrm{Hom}(E, E')$ we see that $M_{\mathbf{Q}}$ is 1-dimensional as a right vector space over $R_{\mathbf{Q}} = D$.

Lemma A.3.1. *The R-module M is projective* (*with rank* 1 *in the sense that* $\dim_D M_{\mathbf{Q}} = 1$).

Proof. It suffices to check that M_ℓ is free of rank 1 over R_ℓ for all primes ℓ. As we saw in the proof of Proposition 3.2.1, for a sufficiently large finite subfield $k \subset \overline{\mathbf{F}}_p$ with size $q = p^{2n}$ for some integer n there are elliptic curves E_0, E_0' over k that descend E, E' respectively such that the q-Frobenius endomorphisms of E_0 and E_0' are multiplication by p^n. These properties are preserved by further finite extension of k, and since M is $\mathbf{Z}$-finite we may increase k some more so that the natural inclusion $M_0 := \mathrm{Hom}_k(E_0, E_0') \hookrightarrow \mathrm{Hom}(E, E') = M$ is an equality, and likewise that $R_0 := \mathrm{End}_k(E_0) = \mathrm{End}(E) = R$.

By Tate's isogeny theorem, the natural map

$$\mathbf{Z}_\ell \otimes_{\mathbf{Z}} M_0 \to \mathrm{Hom}_k(E_0[\ell^\infty], E_0'[\ell^\infty])$$

is an isomorphism for all ℓ. Hence, it suffices to show that the right side is free of rank 1 as a left module over

$$\mathbf{Z}_\ell \otimes_{\mathbf{Z}} R_0 = \mathrm{End}_k(E_0[\ell^\infty]).$$

It is therefore sufficient to prove that $E_0'[\ell^\infty] \simeq E_0[\ell^\infty]$ for every prime ℓ. If $\ell \neq p$ then both ℓ-divisible groups have associated ℓ-adic Tate modules free of rank 2 on which the q-Frobenius in the Galois group acts through multiplication by p^n. This settles the case $\ell \neq p$.

Suppose instead that $\ell = p$. We know that $E[p^\infty] \simeq E'[p^\infty]$ since the associated commutative 1-dimensional formal groups of height 2 are isomorphic over the algebraically closed field $\overline{\mathbf{F}}_p$. Our task is to show that such an isomorphism is necessarily $\mathrm{Gal}(\overline{\mathbf{F}}_p/k)$-equivariant. Passing to the isogeny category, it suffices to show that all elements in $\mathrm{Hom}^0(E[p^\infty], E'[p^\infty])$ are Galois-invariant. But E'_0 is k-isogenous to E_0 by Tate's isogeny theorem, so we may replace $E'[p^\infty]$ with $E[p^\infty]$ to reduce to checking that all elements of $\mathrm{End}^0(E[p^\infty])$ are defined over k. The injection

$$\mathrm{End}^0_k(E_0[p^\infty]) \hookrightarrow \mathrm{End}^0(E[p^\infty])$$

is between $\mathbf{Q}_p$-algebras with the same dimension (namely, 4), so it is an equality. □

Rather generally, if A is an abelian variety over a field F, $R \subset \mathrm{End}(A)$ is a subring, and M is a finitely generated right R-module then the fppf abelian sheafification of $S \mapsto M \otimes_R A(S)$ on the category of F-schemes is represented by an abelian variety (here and below, we are using the standard convention for tensor products over a non-commutative ring); this abelian variety is denoted $M \otimes_R A$. Indeed, if we choose a finite presentation

$$R^n \to R^{n'} \to M \to 0$$

then it is easy to check that the cokernel of the evident map of abelian varities $A^n \to A^{n'}$ does the job. Beware that there is *no* natural left R-action on $M \otimes_R A$ when R is not commutative. The problem is that the "definition" $r.(m \otimes a) = m \otimes ra$ then is not well-posed, due to the equality $mr' \otimes a = m \otimes r'a$ and the non-commutativity of R.

If $M_{\mathbf{Q}}$ is free of rank d over $R_{\mathbf{Q}}$ then $\dim(M \otimes_R A) = d \cdot \dim(A)$; this follows from functoriality in M and the evident existence of a pair of R-linear maps $R^d \to M$ and $M \to R^d$ whose compositions in either order is multiplication by a common nonzero integer (providing an isogeny between $M \otimes_R A$ and A^d). We conclude that in the setup of Lemma 3.2.1, if N is a projective right R-module with rank 1 then $N \otimes_R E$ is an *elliptic curve* that is isogenous to E and hence is supersingular. We'll soon see that if we choose a D-basis of $N_{\mathbf{Q}}$ (or equivalently identify N with an R-submodule of D) then $\mathrm{End}^0(N \otimes_R E)$ is naturally identified with $\mathrm{End}^0(E) = D$ but the associated maximal order $\mathrm{End}(N \otimes_R E)$ is generally *not* R (even up to $D^\times$-conjugacy).

By Lemma A.3.1, the procedure $N \rightsquigarrow N \otimes_R E$ for varying N (but our fixed E) recovers *all* supersingular elliptic curves over $\overline{\mathbf{F}}_p$ (up to isomorphism) in a unique way:

Lemma A.3.2. *If E' is any supersingular elliptic curve over $\overline{\mathbf{F}}_p$ then the evaluation homomorphism*

$$\mathrm{Hom}(E, E') \otimes_R E \to E'$$

is an isomorphism. Conversely, if N is a projective right R-module with rank 1 then the map

$$N \to \mathrm{Hom}(E, N \otimes_R E)$$

defined by $n \mapsto (x \mapsto n \otimes x)$ is an R-linear isomorphism, where the functor $\mathrm{Hom}(E, \cdot)$ on commutative group schemes is valued in right R-modules via the left R-action on E.

In particular, if N is embedded into D as an R-submodule then $\mathrm{Hom}^0(E, N \otimes_R E)$ is naturally identified with D, so the element $1 \in D$ yields a preferred identification of $\mathrm{End}^0(N \otimes_R E)$ with $\mathrm{End}^0(E) = D$.

Proof. The map $N \to \mathrm{Hom}(E, N \otimes_R E)$ is R-linear precisely because $nr \otimes x = n \otimes rx$ for $n \in N$, $r \in R$, and functorial points x of E. The verification of the isomorphism assertions will require arguments with ℓ-divisible groups similarly to what we have already seen, except that there will be a small extra argument required at the level of finite group schemes.

Since a map between abelian varieties is an isomorphism if and only if it is an isomorphism on ℓ-divisible groups for *all* primes ℓ (an instructive exercise with group schemes), to verify the first isomorphism property it suffices to check on ℓ-divisible groups for all ℓ. Likewise, a map between finitely generated R-modules is an isomorphism if and only if it is so after completion at all primes

ℓ. By descent to a sufficiently large finite field, it suffices to treat the analogous assertions using elliptic curves and Hom-modules over a finite field k of size $q = p^{2n}$ such that the q-Frobenius endomorphism on the elliptic curves is p^n.

Tate's isogeny theorem over finite fields identifies $\mathrm{Hom}_k(E_0, E_0')_\ell$ with $\mathrm{Hom}_k(E_0[\ell^\infty], E_0'[\ell^\infty])$, and we have seen in earlier proofs that the ℓ-divisible groups of E_0 and E_0' are abstractly k-isomorphic for all ℓ (allowing $\ell = p$). Thus, the essential task for both isomorphism assertions is to unravel how the tensor operation $N \rightsquigarrow N \otimes_R E_0$ interacts with the passage to ℓ-divisible groups for all ℓ.

For any prime ℓ and projective right R-module N of rank $d \geq 1$, N_ℓ is free of rank d over R_ℓ. In particular, $N/\ell^n N$ is free of rank d over $R/\ell^n R$ for all ℓ and all $n \geq 1$. Thus, it suffices to show that for any prime ℓ, the natural map of group schemes

$$(N/\ell^n N) \otimes_{R/\ell^n R} E_0[\ell^n] \to (N \otimes_R E_0)[\ell^n]$$

is an isomorphism (where the left side is an evident Serre tensor construction via the fppf topology over a field, especially for $\ell = p$). Since N is a direct summand of R^m for some $m \geq 1$, by the evident functorality in N and compatibility with respect to direct products in N we reduce to the trivial case $N = R^m$. □

Observe that if we change the choice of isomorphism $\mathrm{End}^0(E) \simeq D$ then the maximal order $\mathrm{End}(E) \subset D$ is changed by $D^\times$-conjugation. Thus, $E \mapsto \mathrm{End}(E)$ is a well-defined map from the set of isomorphism classes of such E's onto the set of $D^\times$-conjugacy classes of maximal orders in D. Moreover, every maximal order $R' \subset D$ arises in this way. Indeed, the "lattice index" $[R : R'] = \{d \in D \,|\, d \cdot R \subset R'\}$ is a finitely generated right R-module that is projective of rank 1 because R_ℓ is $D_\ell^\times$-conjugate to R'_ℓ for all primes ℓ; it is also a left R'-submodule of D. Thus, $E' := [R : R'] \otimes_R E$ makes sense as a supersingular elliptic curve on which there is a natural left R'-action respecting a natural identification $\mathrm{End}^0(E') = \mathrm{End}^0(E) \simeq D$, so the resulting inclusion $R' \subset \mathrm{End}(E')$ of orders in D has to be an equality due to the maximality of R' in D.

We can now prove Theorem 3.2.2.

Proof. Let E be a supersingular elliptic curve over $\overline{\mathbf{F}}_p$ such that $\mathrm{End}(E)$ is in the $D^\times$-conjugacy class of R. (We have seen above that such an E exists.) We may and do fix an isomorphism $\mathrm{End}^0(E) \simeq D$ identifying $\mathrm{End}(E)$ with R.

By Lemma A.3.1 and Lemma A.3.2, every supersingular elliptic curve over $\overline{\mathbf{F}}_p$ has the form $M \otimes_R E$ for a projective right R-module M of rank 1, with $M \otimes_R E \simeq N \otimes_R E$ if and only if $M \simeq N$. (This is immediate from the fact that $M \simeq \mathrm{Hom}(E, M \otimes_R E)$ as right R-modules.) Thus, $E' \mapsto \mathrm{Hom}(E, E')$ is a bijection (depending on E) of Σ onto the set of isomorphism classes of rank-1 projective right R-modules.

Our problem now can be formulated purely in terms of algebra with R and D, having nothing to do with supersingular elliptic curves: for any maximal order R in D (so $D = R_{\mathbf{Q}}$), we claim that the double coset space

$$\underline{G}^{\mathrm{ad}}(\mathbf{Q}) \backslash \underline{G}^{\mathrm{ad}}(\mathbf{A}) / K = D^\times \backslash D_{\mathbf{A}^\infty}^\times / \widehat{R}^\times$$

(with $\widehat{R} := R \otimes_{\mathbf{Z}} \widehat{\mathbf{Z}}$) is in natural bijection with the set of isomorphism classes of projective right R-modules M of rank 1. (This is a non-commutative variant of the idelic description of ideal class groups of integer rings of number fields.)

Motivated by the dictionary between fractional ideals of a number field F and abstract invertible $\mathscr{O}_F$ modules with a generic trivialization, if we fix a D-basis of $M_{\mathbf{Q}}$ so as to identify $M_{\mathbf{Q}}$ with D then the R-submodule $M \subset D$ depends up to left $D^\times$-multiplication on exactly the isomorphism class of M as a right R-module. Thus, we need to identify the set of left $D^\times$-homothety classes of rank-1 projective right R-submodules of D and the above double coset space. We shall prove something finer: a bijection between $D_{\mathbf{A}^\infty}^\times / \widehat{R}^\times$ and the set of rank-1 projective right R-submodules of D, with left $D^\times$-homothey on such submodules going over to the left $D^\times$-action on $D_{\mathbf{A}^\infty}^\times / \widehat{R}^\times$.

For any $M \subset D$ we have $M[1/n] = R[1/n]$ inside D for sufficiently divisible nonzero $n \in \mathbf{Z}$, so the R_ℓ-freeness of M_ℓ implies that $M_\ell = m_\ell R_\ell$ for some $m = (m_\ell) \in D^\times_{\mathbf{A}^\infty}$. The element m is well-defined up to right multiplication against $\widehat{R}^\times$, and any change in the right D-module isomorphism $M_{\mathbf{Q}} \simeq D$ is obtained through *left* multiplication on D against an element of $D^\times$, so we have the desired double coset description. □

A.4. **Lifting Supersingular Curves.** This section gives proofs of Theorem 4.1.1. The notation is the same as in §4.1.

Proof. To construct the lift it suffices to do so after a preliminary descent of $(E, \mathscr{O})$ over a finite subfield k of $\overline{\mathbf{F}}_p$ (and then build the lift over the valuation ring of $L_0 := W(k)[1/p] \cdot F_p$). This is carried out in [CCO, 1.7.4.6] using deformation theory and p-divisible groups (where the ordinary case is also treated). The key ingredient is the Serre–Tate deformation theorem, which identifies the infinitesimal deformation theories of abelian schemes and their p-divisible groups over rings in which p is nilpotent. (Note that the p-divisible group of an abelian scheme over a complete local noetherian ring with residue characteristic p has associated formal group coinciding with that of the abelian scheme, so its algebra is canonically identified with that of the abelian scheme. Hence, the required compatibility condition on Lie algebras is also encoded in terms of the p-divisible group.) Strictly speaking, a solution to the lifting problem for the p-divisible group provides only an $\mathscr{O}$-linear formal elliptic curve deformation of E over $\mathrm{Spf}(\mathscr{O}_{L_0})$, but formal GAGA (for proper formal schemes and morphisms among them) ensures that this data can be uniquely algebraized over $\mathrm{Spec}(\mathscr{O}_{L_0})$.

This reduces the lifting problem for the monogenic order $\mathscr{O}$ acting on E to that of $\mathscr{O}$ acting on $E[p^\infty]$, or equivalently on the associated 1-dimensional commutative formal group of height 2. On the p-divisible group or formal group, such an action of $\mathscr{O}$ uniquely extends to an action of the $\mathbf{Z}_p$-order $\mathscr{O}_p := \mathbf{Z}_p \otimes_{\mathbf{Z}} \mathscr{O} \subset F_p$. Hence, our problem is exactly that of deformation in the category of formal $\mathscr{O}_p$-modules over $\mathscr{O}_p$-algebras (such as $\mathscr{O}_L$!), where we recall that in the definition of a formal module we impose the condition that the induced action of "coefficients" on the Lie algebra is the canonical one via the algebra structure on the base ring.

By these same methods, to prove the uniqueness up to unique isomorphism it suffices to do the same for the formal $\mathscr{O}_p$-module. The uniqueness of such an isomorphism is immediate from faithfulness of the "special fiber" functor on p-divisible groups over complete local noetherian rings with residue characteristic p (see [CCO, 1.4.4.3]). So our task is just to construct an isomorphism. The key point is to check that $\mathscr{O}_p$ is the valuation ring of F_p (rather than a general $\mathbf{Z}_p$-order), so we can then apply the deformation theory of formal modules over p-adic integer rings. To see that $\mathscr{O}_p$ is the valuation ring (even though globally $\mathscr{O}$ is generally just an order in $\mathscr{O}_F$), we recall that by supersingularity $R := \mathrm{End}(E)$ is a maximal order in $\mathrm{End}^0(E) = D$, and the saturation $\mathscr{O}$ is $F \cap R$. Hence, $\mathscr{O}_p = F_p \cap R_p$ inside D_p, and D_p is a quaternion *division algebra* over $\mathbf{Q}_p$, so the maximal order R_p in D_p is *unique* (not merely up to $D_p^\times$-conjugacy) and is characterized in terms of a "valuation" on D_p that restricts to a p-adic valuation on F_p. Thus, the intersection $F_p \cap R_p$ is the valuation ring of F_p as desired.

Since $\mathscr{O}_p$ is a quadratic p-adic integer ring and $E[p^\infty]$ has height 2, its associated formal group viewed as a formal $\mathscr{O}_p$-module has $\mathscr{O}_p$-height equal to $2/2 = 1$. Now it remains to recall (see [H, 22.4.4]) that if A is a complete discrete valuation ring with finite residue field $\mathbf{F}$ then the infintesimal deformation theory of a formal A-module Γ of A-height $h \geq 1$ and dimension 1 over an extension field of $\mathbf{F}$ is formally smooth of dimension $h - 1$ (and so it is formally étale when $h = 1$, as in the case of interest to us with $A = \mathscr{O}_p$). □

References

[Bo] A. Borel, *Introduction aux groupes arithmétiques*, Hermann, Paris, 1969.

[Br] C. Brislawn, *Traceable integral kernels on countably generated measure spaces*, Pacific J. Math. **150** (1991), no. 2, 229–240. MR 1123441 (92k:47042)

[Bu] K. Buzzard, *Computing modular forms on definite quaternion algebras.*

[CCO] C-L. Chai, B. Conrad, F. Oort, *Complex multiplication and lifting problems*, AMS Surveys **195** (2014).

[DL] M. Duflo and J-P. Labesse, *Sur la formule des traces de Selberg*, Ann. Sci. École Norm. Sup. (4) **4** (1971), 193–284. MR 0437462 (55 #10392)

[GJ] S. Gelbart and H. Jacquet, *Forms of* GL(2) *from the analytic point of view*, Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 1, Proc. Sympos. Pure Math., XXXIII, Amer. Math. Soc., Providence, R.I., 1979, pp. 213–251. MR 546600 (81e:10024)

[Gr] B. Gross, *Heights and the special values of L-series*, Number theory (Montreal, Que., 1985), CMS Conf. Proc., vol. 7, Amer. Math. Soc., Providence, RI, 1987, pp. 115–187. MR 894322 (89c:11082)

[GZ] B. Gross, D. Zagier, *On singular moduli*, J. Reine Angew. Math. **355** (1985), 191–220. MR 772491 (86j:11041)

[H] M. Hazewinkel, *Formal groups and applications*, Academic Press, New York, 1978.

[KM] N. Katz, B. Mazur, *Arithmetic moduli of elliptic curves*, Annals of Math Studies, 1985.

[K] R. Kottwitz, *Harmonic analysis on reductive p-adic groups and Lie algebras*, Harmonic analysis, the trace formula, and Shimura varieties, Clay Math. Proc., vol. 4, Amer. Math. Soc., Providence, RI, 2005, pp. 393–522. MR 2192014 (2006m:22016)

[Mu] D. Mumford, *Abelian varieties*, Oxford Univ. Press, 1970.

[Na] L. Nachbin, *The Haar integral*, Robert E. Krieger Publishing Co., Huntington, N.Y., 1976, Translated from the Portuguese by Lulu Bechtolsheim, Reprint of the 1965 edition. MR 0414833 (54 #2925)

[Oes] J. Oesterlé, *Nombres de Tamagawa et groupes unipotents en caractéristique p*, Inv. Math. **78**(1984), 13–88.

[RS] M. Reed, B. Simon, *Methods of modern mathematical physics. I. Functional analysis*, Academic Press, New York, 1972. MR 0493419 (58 #12429a)

[Se] J.-P. Serre, *Two letters on quaternions and modular forms (mod p)*, Israel J. Math. **95** (1996), 281–299, With introduction, appendix and references by R. Livné. MR 1418297 (98b:11049)

[Ta1] J. Tate, *Endomorphisms of abelian varieties over finite fields* Inventiones **2** (1966), pp. 134–144.

[Ta2] J. Tate, "p-divisible groups" in *Proceedings of a conference on local fields* (T. Springer, ed.), Springer–Verlag, 1967, 158–183.

[V] M-F. Vignéras, *Arithmétique des algèbres de quaternions*, Lecture Notes in Mathematics, vol. 800, Springer, Berlin, 1980. MR 580949 (82i:12016)