Algebra & Number Theory

Volume 14

2020

No. 5

p-adic distribution of CM points and Hecke orbits I: Convergence towards the Gauss point

Sebastián Herrero, Ricardo Menares and Juan Rivera-Letelier

msp

msp

ALGEBRA AND NUMBER THEORY 14:5 (2020)
https://doi.org/10.2140/ant.2020.14.1239

p-adic distribution of CM points and Hecke orbits I: Convergence towards the Gauss point

Sebastián Herrero, Ricardo Menares and Juan Rivera-Letelier

We study the asymptotic distribution of CM points on the moduli space of elliptic curves over $\mathbb{C}_p$, as the discriminant of the underlying endomorphism ring varies. In contrast with the complex case, we show that there is no uniform distribution. In this paper we characterize all the sequences of discriminants for which the corresponding CM points converge towards the Gauss point of the Berkovich affine line. We also give an analogous characterization for Hecke orbits. In the companion paper we characterize all the remaining limit measures of CM points and Hecke orbits.

1. Introduction

Given an algebraically closed field $\mathbb{K}$, denote by $Y(\mathbb{K})$ the moduli space of elliptic curves over $\mathbb{K}$. It is the space of all isomorphism classes of elliptic curves over $\mathbb{K}$, for isomorphisms defined over $\mathbb{K}$. For a class E in $Y(\mathbb{K})$, the j-invariant $j(E)$ of E is an element of $\mathbb{K}$ determining E completely. The map $j : Y(\mathbb{K}) \to \mathbb{K}$ so defined is a bijection. See for example [Silverman 2009] and [Lang 1973] for background on elliptic curves.

If $\mathbb{K}$ is of characteristic 0, then the endomorphism ring of an elliptic curve defined over $\mathbb{K}$ is isomorphic to $\mathbb{Z}$ or to an order in a quadratic imaginary extension of $\mathbb{Q}$. In the latter case, the order only depends on the class E in $Y(\mathbb{K})$ of the elliptic curve and E is said to have *complex multiplication* or to be a *CM point*. In this paper, the *discriminant of a CM point* is the discriminant of the corresponding order.* Moreover, a

MSC2010: primary 11G15; secondary 11F32, 11S82.
Keywords: equidistribution, elliptic curves, Hecke correspondences.

*This notion of discriminant is not to be confused with the discriminant of a Weierstrass model of an elliptic curve [Silverman 2009, Chapter III, Section 1].

discriminant is the discriminant of an order in a quadratic imaginary extension of $\mathbb{Q}$. An integer D is a discriminant if and only if $D < 0$ and $D \equiv 0, 1 \mod 4$.

For every discriminant D, the set

$$\Lambda_D := \{E \in Y(\mathbb{K}) : \text{CM point of discriminant } D\} \tag{1-1}$$

is finite and nonempty. So, we can define the probability measure $\bar{\delta}_D$ on $Y(\mathbb{K})$, by

$$\bar{\delta}_D := \frac{1}{\#\Lambda_D} \sum_{E \in \Lambda_D} \delta_E,$$

where δ_x denotes the Dirac measure on $Y(\mathbb{K})$ at x.

Throughout the rest of this paper we fix a prime number p and a completion $(\mathbb{C}_p, |\cdot|_p)$ of an algebraic closure of the field of p-adic numbers $\mathbb{Q}_p$. Our first goal is to study, for $\mathbb{K} = \mathbb{C}_p$, the asymptotic distribution of Λ_D as the discriminant D tends to $-\infty$. This is motivated by the following result in the case where $\mathbb{K}$ is the field of complex numbers $\mathbb{C}$. Recall that, if we consider the usual action of $\mathrm{SL}_2(\mathbb{Z})$ on the upper half-plane $\mathbb{H}$ by Möbius transformations, then $Y(\mathbb{C})$ can be naturally identified with the quotient space $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$. An appropriate multiple of the hyperbolic measure on $\mathbb{H}$ descends to a probability measure μ_{hyp} on $Y(\mathbb{C})$.

Theorem 1. *For every continuous and bounded function $\varphi\colon Y(\mathbb{C}) \to \mathbb{R}$, we have*

$$\frac{1}{\#\Lambda_D} \sum_{E \in \Lambda_D} \varphi(E) \to \int \varphi \, \mathrm{d}\mu_{\mathrm{hyp}},$$

as the discriminant D tends to $-\infty$. Equivalently, we have the weak convergence of measures

$$\bar{\delta}_D \to \mu_{\mathrm{hyp}},$$

as the discriminant D tends to $-\infty$.

The asymptotic distribution of CM points on $Y(\mathbb{C})$ was part of a family of problems studied by Linnik; see [Linnik 1968] and also [Michel and Venkatesh 2006]. By applying a certain "ergodic method", Linnik proved the result above for sequences of discriminants satisfying some congruence restrictions. In a breakthrough, Duke [1988] removed the congruence restrictions assumed by Linnik and proved Theorem 1 for fundamental discriminants. Duke's proof uses the theory of nonholomorphic modular forms of half-integral weight and bounds for their Fourier coefficients, building on work of Iwaniec [1987]. Finally, Clozel and Ullmo [2004] obtained Theorem 1 for arbitrary discriminants, by studying the action of Hecke correspondences on CM points and combining Duke's result together with the uniform distribution of Hecke orbits.

1A. ***Convergence of CM points towards the Gauss point.*** Our first goal is to describe the asymptotic distribution of CM points for the ground field $\mathbb{K} = \mathbb{C}_p$. However, it is easy to find sequences of discriminants $(D_n)_{n=1}^{\infty}$ for which the sequence of measures $(\bar{\delta}_{D_n})_{n=1}^{\infty}$ on $Y(\mathbb{C}_p)$ has no accumulation measure. A natural solution to this issue is to consider $Y(\mathbb{C}_p)$ as a subspace of the Berkovich affine line

$\mathbb{A}^1_{\mathrm{Berk}}$ over $\mathbb{C}_p$, using the j-invariant to identify $Y(\mathbb{C}_p)$ with the subspace $\mathbb{C}_p$ of $\mathbb{A}^1_{\mathrm{Berk}}$. In fact, every sequence of measures $(\bar{\delta}_{D_n})_{n=1}^{\infty}$ as above accumulates on at least one probability measure with respect to the weak topology on the space of Borel measures on $\mathbb{A}^1_{\mathrm{Berk}}$. See Section 2D for a brief review of the space $\mathbb{A}^1_{\mathrm{Berk}}$ and the weak topology on the space of measures on $\mathbb{A}^1_{\mathrm{Berk}}$.

In contrast with Theorem 1, for $\mathbb{K} = \mathbb{C}_p$ the measures $\bar{\delta}_D$ on $\mathbb{A}^1_{\mathrm{Berk}}$ do not converge to a limit as the discriminant D tends to $-\infty$. Our first main result is a characterization of all those sequences of discriminants $(D_n)_{n=1}^{\infty}$ tending to $-\infty$, such that the sequence of measures $(\bar{\delta}_{D_n})_{n=1}^{\infty}$ in $\mathbb{A}^1_{\mathrm{Berk}}$ converges to the Dirac measure at the "canonical" or "Gauss point" x_{can} of $\mathbb{A}^1_{\mathrm{Berk}}$. In the companion paper [Herrero et al. 2019] we show that in all the remaining cases the sequence $(\bar{\delta}_{D_n})_{n=1}^{\infty}$ accumulates on at least one probability measure supported on a compact subset of the supersingular locus of $Y(\mathbb{C}_p)$ and characterize all possible accumulation measures.

To state our first main result, we introduce some notation and terminology. Identify the residue field of $\mathbb{C}_p$ with an algebraic closure $\bar{\mathbb{F}}_p$ of the field with p elements $\mathbb{F}_p$. Recall that the endomorphism ring of an elliptic curve over $\bar{\mathbb{F}}_p$ is isomorphic to an order in either a quadratic imaginary extension of $\mathbb{Q}$ or a quaternion algebra over $\mathbb{Q}$. In the former case the corresponding elliptic curve class is *ordinary* and it is *supersingular* in the latter.

Denote by $\mathcal{O}_p$ the ring of integers of $\mathbb{C}_p$ and by $\pi : \mathcal{O}_p \to \bar{\mathbb{F}}_p$ the reduction map. An elliptic curve class E has *good reduction* if there is a representative Weierstrass equation with coefficients in $\mathcal{O}_p$ whose reduction is a smooth curve. Such reduction determines an elliptic curve defined over $\bar{\mathbb{F}}_p$, whose class $\tilde{E}$ only depends on E and is the *reduction of* E. Moreover, E has *ordinary* (resp. *supersingular*) *reduction* if $\tilde{E}$ is ordinary (resp. supersingular). An elliptic curve has good reduction precisely when $j(E)$ is in $\mathcal{O}_p$ and when this is not the case E has *bad reduction*. The moduli space $Y(\mathbb{C}_p)$ is thus partitioned into three pairwise disjoint sets: The *bad*, *ordinary* and *supersingular reduction loci*, denoted by $Y_{\mathrm{bad}}(\mathbb{C}_p)$, $Y_{\mathrm{ord}}(\mathbb{C}_p)$ and $Y_{\mathrm{sups}}(\mathbb{C}_p)$, respectively. Using $j : Y(\mathbb{C}_p) \to \mathbb{C}_p$ to identify $Y(\mathbb{C}_p)$ and $\mathbb{C}_p$, we thus have the partition

$$\mathcal{O}_p = Y_{\mathrm{ord}}(\mathbb{C}_p) \sqcup Y_{\mathrm{sups}}(\mathbb{C}_p).$$

Moreover, if we denote by $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$ the finite subset of $Y(\bar{\mathbb{F}}_p)$ of supersingular classes, then $Y_{\mathrm{sups}}(\mathbb{C}_p) = \pi^{-1}(Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p))$ is a finite union of residue discs of $\mathcal{O}_p$. Note that $Y_{\mathrm{ord}}(\mathbb{C}_p)$ is a union of infinitely many residue discs of $\mathcal{O}_p$.

Every CM point E has good reduction and the reduction type only depends on the discriminant D of E, as follows:

(i) If p splits in $\mathbb{Q}(\sqrt{D})$, then E has ordinary reduction.

(ii) If p ramifies or is inert in $\mathbb{Q}(\sqrt{D})$, then E has supersingular reduction.

See [Deuring 1941] or [Lang 1973, Chapter 13, Section 4, Theorem 12]. We call a discriminant D *p-ordinary* in the first case and *p-supersingular* in the second. Moreover, we define

$$|D|_{p\text{-sups}} := \begin{cases} 0 & \text{if } D \text{ is } p\text{-ordinary;} \\ |D|_p & \text{if } D \text{ is } p\text{-supersingular.} \end{cases}$$

Theorem A. *Let $(D_n)_{n=1}^{\infty}$ be a sequence of discriminants tending to $-\infty$. Then we have the weak convergence of measures*

$$\bar{\delta}_{D_n} \to \delta_{x_{\mathrm{can}}} \text{ as } n \to \infty \quad \text{if and only if} \quad |D_n|_{p\text{-sups}} \to 0 \text{ as } n \to \infty.$$

For readers unfamiliar with the Berkovich affine line, we give a concrete formulation of the convergence of measures in Theorem A in terms of $\mathbb{C}_p$ only, see Lemma 2.3(ii) in Section 2D.

We obtain Theorem A as a direct consequence of quantitative estimates in the cases where all the discriminants in $(D_n)_{n=1}^{\infty}$ are p-ordinary (Theorem 3.5 in Section 3B) or p-supersingular (Theorem 4.1 in Section 4). Note that in the former case Theorem A asserts that $\bar{\delta}_{D_n} \to \delta_{x_{\mathrm{can}}}$ weakly as $n \to \infty$. The following stronger statement is a direct consequence of our quantitative estimate in this case.

Corollary B (ordinary CM points are isolated). *Every disc of radius strictly less than one contained in $Y_{\mathrm{ord}}(\mathbb{C}_p)$ contains at most a finite number of CM points. In particular, the set of CM points in $Y_{\mathrm{ord}}(\mathbb{C}_p)$ is discrete.*

Corollary B seems to be well-known by the experts in the field, although we have not found this result explicitly stated in the literature. See Section 1C for comments and references.

1B. *Convergence of Hecke orbits towards the Gauss point.* To state our next main result, we first introduce Hecke correspondences. See Section 2B for background.

Given an algebraically closed field $\mathbb{K}$ of characteristic 0, a *divisor on* $Y(\mathbb{K})$ is an element of

$$\mathrm{Div}(Y(\mathbb{K})) := \bigoplus_{E \in Y(\mathbb{K})} \mathbb{Z}E,$$

the free abelian group spanned by the points of $Y(\mathbb{K})$. The *degree* and *support* of a divisor $\mathfrak{D} = \sum_{E \in Y(\mathbb{K})} n_E E$ in $\mathrm{Div}(Y(\mathbb{K}))$ are defined by

$$\deg(\mathfrak{D}) := \sum_{E \in Y(\mathbb{K})} n_E \quad \text{and} \quad \mathrm{supp}(\mathfrak{D}) := \{E \in Y(\mathbb{K}) : n_E \neq 0\},$$

respectively. If in addition $\deg(\mathfrak{D}) \geq 1$ and for every E in $Y(\mathbb{K})$ we have $n_E \geq 0$, then

$$\bar{\delta}_{\mathfrak{D}} := \frac{1}{\deg(\mathfrak{D})} \sum_{E \in Y(\mathbb{K})} n_E \delta_E$$

is a probability measure on $Y(\mathbb{K})$.

For n in $\mathbb{N} := \{1, 2, \ldots\}$ the n-th *Hecke correspondence* is the linear map

$$T_n \colon \mathrm{Div}(Y(\mathbb{K})) \to \mathrm{Div}(Y(\mathbb{K}))$$

defined for E in $Y(\mathbb{K})$, by

$$T_n(E) := \sum_{C \leq E \text{ of order } n} E/C,$$

where the sum runs over all subgroups C of E of order n. Note that $\operatorname{supp}(T_n(E))$ is the set of all E' in $Y(\mathbb{K})$ for which there is an isogeny $E \to E'$ of degree n. Moreover,

$$\deg(T_n(E)) = \sum_{d \mid n, d>0} d \geq n,$$

so $\deg(T_n(E)) \to \infty$ as $n \to \infty$.

In the case $\mathbb{K} = \mathbb{C}_p$, it is easy to see that for each E in $Y_{\text{bad}}(\mathbb{C}_p)$ (resp. $Y_{\text{ord}}(\mathbb{C}_p)$, $Y_{\text{sups}}(\mathbb{C}_p)$), we have that for every n in $\mathbb{N}$ the divisor $T_n(E)$ is supported on $Y_{\text{bad}}(\mathbb{C}_p)$ (resp. $Y_{\text{ord}}(\mathbb{C}_p)$, $Y_{\text{sups}}(\mathbb{C}_p)$).

Theorem C. *For every E in $Y_{\text{bad}}(\mathbb{C}_p) \cup Y_{\text{ord}}(\mathbb{C}_p)$, we have the weak convergence of measures*

$$\bar{\delta}_{T_n(E)} \to \delta_{x_{\text{can}}} \quad \text{as } n \to \infty.$$

Moreover, for E in $Y_{\text{sups}}(\mathbb{C}_p)$ and a sequence $(n_j)_{j=1}^{\infty}$ in $\mathbb{N}$ tending to ∞, we have the weak convergence of measures

$$\bar{\delta}_{T_{n_j}(E)} \to \delta_{x_{\text{can}}} \text{ as } j \to \infty \quad \text{if and only if} \quad |n_j|_p \to 0 \text{ as } j \to \infty.$$

When restricted to the case where E is in $Y_{\text{bad}}(\mathbb{C}_p)$, the above theorem is [Richard 2018, Théorème 1.2].

To the best of our knowledge, Theorem C gives the first example where equidistribution of orbits fails for correspondences of degree bigger than one, see Section 2B for a description of Hecke correspondences as algebraic correspondences. In the complex case, pluripotential theory has been used successfully to prove equidistribution for correspondences satisfying a mild "nonmodularity" condition, see for example [Dinh et al. 2020].

The uniform distribution of Hecke orbits on $Y(\mathbb{C})$ is a well-known result from the spectral theory of automorphic forms; see [Clozel and Ullmo 2004, Théorème 2.1], and also [Clozel et al. 2001; Eskin and Oh 2006] for extensions and [Linnik and Skubenko 1964] for related work.

Remark 1.1. In [Clozel et al. 2001; Eskin and Oh 2006], the starting point is an algebraic group G over $\mathbb{Q}$ and a congruence subgroup Γ of $G(\mathbb{Q})$, and the ambient space is $X = \Gamma \backslash G(\mathbb{R})$. In this context, there is a natural notion of Hecke correspondences on X. The aforementioned works establish the uniform distribution of every orbit of such Hecke correspondences under general hypotheses. In particular, the $\mathbb{Q}$-structure of G allows for p-adic variants of such results, see, e.g., [Clozel et al. 2001, Remark (1) in page 332]. In the particular case $G = \mathrm{SL}_2$ and $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, there is a natural isomorphism $Y(\mathbb{C}) \simeq \mathrm{SL}_2(\mathbb{Z}) \backslash \mathrm{SL}_2(\mathbb{R}) / \mathrm{SO}_2(\mathbb{R})$ and the natural projection from X to $Y(\mathbb{C})$ takes Hecke orbits as in [Clozel et al. 2001; Eskin and Oh 2006] to Hecke orbits on $Y(\mathbb{C})$ as defined in this paper. The uniform distribution of Hecke orbits on $Y(\mathbb{C})$ is thus a special case of [Clozel et al. 2001, Theorem 1.6], see also [Eskin and Oh 2006, Theorem 1.2]. However, this strategy breaks down for Hecke orbits on $Y(\mathbb{C}_p)$, because there is no analogous uniformization of $Y(\mathbb{C}_p)$ as a double quotient. Moreover, Theorem C shows that there is no uniform distribution of Hecke orbits on $Y(\mathbb{C}_p)$. Indeed, Theorem C and our results in the companion paper [Herrero et al. 2019] show that, in contrast with [Clozel et al. 2001; Clozel and

Ullmo 2004; Eskin and Oh 2006], the asymptotic distribution of $(T_{n_j}(E))_{j=1}^{\infty}$ on $Y(\mathbb{C}_p)$ depends on both the starting point E and the sequence of integers $(n_j)_{j=1}^{\infty}$.

1C. ***Notes and references.*** After the first version of this paper was written, we learned about the related work of Goren and Kassaei [2017]. For a prime number ℓ different from p, Goren and Kassaei [2017] studied the dynamics of the Hecke correspondence T_ℓ acting on the moduli space of elliptic curves with a marked torsion point of exact order N coprime to $p\ell$. So, on one hand [Goren and Kassaei 2017] is more general than this paper in that it considers modular curves with level structure. On the other hand, [loc. cit.] is more restrictive in that it only considers the dynamics of a single Hecke correspondence of prime index different from p, as opposed to the dynamics of the whole algebra of Hecke correspondences considered here. Note also that we use $\mathbb{C}_p$ as a ground field, which is natural to study equidistribution problems, whereas [loc. cit.] is restricted to algebraic extensions of $\mathbb{Q}_p$. In spite of the fact that both papers study the dynamics of similar maps, there is no significant intersection between the results of [loc. cit.] and those of this paper. See also [Herrero et al. 2019] for our additional results in the supersingular locus and the corresponding comparison with the results of [Goren and Kassaei 2017]. Finally, our results on the dynamics of the canonical branch $\boldsymbol{t}$ of T_p (defined on $Y_{\mathrm{ord}}(\mathbb{C}_p)$ in Section 3A) on ordinary CM points show that this map gives rise to a "$(p+1)$-volcano" in the sense of [loc. cit., Section 2.1], see Remark 3.6.

Corollary B seems well-known among experts in the field, although we have not found this result explicitly stated in the literature. Even for higher-dimensional abelian varieties it can be deduced from the explicit characterization of the Serre–Tate local coordinates of CM points as torsion points of the multiplicative group, see, e.g., [de Jong and Noot 1991, Proposition 3.5]. Our approach makes no use of these local coordinates, and is based on rigid analytic properties of the canonical branch $\boldsymbol{t}$ of T_p. For CM elliptic curves with ordinary reduction, the connection between these two approaches is well-known, see, e.g., [Dwork 1969, Section 7d)].

Since every CM point of $Y(\mathbb{C}_p)$ is in the bounded set $\mathcal{O}_p$, Theorem A yields the following stronger statement: For every continuous function $\varphi\colon Y(\mathbb{C}_p) \to \mathbb{R}$ and every sequence of discriminants $(D_n)_{n=1}^{\infty}$ tending to $-\infty$ and satisfying $|D_n|_{p\text{-sups}} \to 0$ as $n \to \infty$, we have

$$\frac{1}{\#\deg(\Lambda_{D_n})} \sum_{E \in \Lambda_{D_n}} \varphi(E) \to \int \varphi \, \mathrm{d}\delta_{x_{\mathrm{can}}} \quad \text{as } n \to \infty.$$

Although our formulation of Theorem 1 seems stronger than the one in [Clozel and Ullmo 2004, Théorème 2.4], it is easy to see that it is equivalent, see for example [Bilu 1997, Lemma 2.2].

1D. ***Strategy and organization.*** We now explain the strategy of the proof of Theorems A and C and simultaneously describe the organization of the paper.

After some preliminaries in Section 2, we proceed to the proof of Theorem A in Sections 3 and 4. Theorem A is a direct consequence of stronger quantitative estimates in two separate cases: The case where all the discriminants in $(D_n)_{n=1}^{\infty}$ are p-ordinary and the case where they are all p-supersingular.

The p-ordinary case is treated in Section 3. There are two main ingredients, both of which are related to the "canonical branch $\boldsymbol{t}$" of T_p that is defined in terms of the "canonical subgroup" in Section 3A; see also Appendix B. The first main tool is a simple formula, for every integer $m \geq 1$, of T_{p^m} on $Y_{\text{ord}}(\mathbb{C}_p)$ in terms of $\boldsymbol{t}$ (Proposition 3.4 in Section 3A). To establish this formula we use results of Tate and Deligne to show that $\boldsymbol{t}$ is rigid analytic. The second main tool is the interpretation of p-ordinary CM points as preperiodic points of $\boldsymbol{t}$ on $Y_{\text{ord}}(\mathbb{C}_p)$ (Theorem 3.5(i)), which is based on Deuring's work on the canonical subgroup. Our quantitative estimate in the p-ordinary case is stated as Theorem 3.5(ii) in Section 3B and its proof is given at the end of this section.

The p-supersingular case is technically more difficult. We use Katz–Lubin's extension of the theory of canonical subgroups to "not too supersingular" elliptic curves and "Katz' valuation". We recall these in Section 4A, where we also give an explicit formula relating Katz' valuation to the j-invariant (Proposition 4.3). We use Katz' valuation to give a concrete description of the action of Hecke correspondences on the supersingular locus in terms of a sequence of correspondences $(\tau_m)_{m=1}^{\infty}$ acting on the interval $[0, p/(p+1)]$ (Proposition 4.5 in Section 4B). To do this, we rely on results in [Katz 1973, Section 3] and, for $p = 2$ and 3, on certain congruences satisfied by certain Eisenstein series, see Proposition A.1 in Appendix A. Our quantitative estimate in the p-ordinary case is stated as Theorem 4.1 at the beginning of Section 4 and its proof is given at the end of this section.

In Appendix B we formulate some of our results on the canonical branch $\boldsymbol{t}$ of T_p, as a lift of the classical Eichler–Shimura congruence relation (Theorem B.1).

The proof of Theorem C splits in three complementary cases, according to the reduction type of E. In each case we obtain a stronger quantitative estimate. For the bad reduction case we use Tate's uniformization theory (Proposition 5.1 in Section 5A). Thanks to the multiplicative properties of Hecke correspondences (2-6), the ordinary reduction case (Proposition 5.2 in Section 5B) is reduced to two special cases: The asymptotic distribution of $(T_{p^m}(E))_{m=1}^{\infty}$ (Proposition 5.3) and, for a sequence $(n_j)_{j=1}^{\infty}$ of integers in $\mathbb{N}$ that are not divisible by p, the asymptotic distribution of $(T_{n_j}(E))_{j=1}^{\infty}$ (Proposition 5.4). The former case is obtained using the tools developed in Theorem 3.5 and the latter is reduced to the study of the action of Hecke correspondences on ordinary elliptic curves in $Y(\overline{\mathbb{F}}_p)$ and is elementary. Finally, the supersingular case (Proposition 5.6 in Section 5C) is obtained from the description of the action of Hecke correspondences on the supersingular locus in Section 4B and an explicit formula for the correspondences $(\tau_m)_{m=1}^{\infty}$ (Lemma 5.7).

2. Preliminaries

Recall that $\mathbb{N} = \{1, 2, \ldots\}$. Given n in $\mathbb{N}$, denote by

$$d(n) := \sum_{d>0, d \mid n} 1 \quad \text{and} \quad \sigma_1(n) := \sum_{d>0, d \mid n} d$$

the number and the sum of the positive divisors of n, respectively. We use several times the inequality

$$\sigma_1(n) \geq n, \tag{2-1}$$

and the fact that for every $\varepsilon > 0$ we have

$$d(n) = o(n^{\varepsilon}); \tag{2-2}$$

see for example [Apostol 1976, page 296].

For a set X and a subset A of X, we use $\mathbf{1}_A\colon X \to \{0, 1\}$ to denote the indicator function of A.

For a topological space X, denote by δ_x the *Dirac mass on X supported at x*. It is the Borel probability measure characterized by the property that for every Borel subset Y of X we have $\delta_x(Y) = 1$ if $x \in Y$ and $\delta_x(Y) = 0$ otherwise.

Normalize the norm $|\cdot|_p$ of $\mathbb{C}_p$ so that $|p|_p = 1/p$ and denote by $\operatorname{ord}_p\colon \mathbb{C}_p \to \mathbb{R} \cup \{+\infty\}$ the valuation defined by $\operatorname{ord}_p(0) = +\infty$ and for z in $\mathbb{C}_p^{\times}$ by $\operatorname{ord}_p(z) = -\log|z|_p/\log p$. Denote by $\mathcal{M}_p$ the maximal ideal of $\mathcal{O}_p$ and recall that we identify $\mathcal{O}_p/\mathcal{M}_p$ with $\bar{\mathbb{F}}_p$ and that $\pi\colon \mathcal{O}_p \to \bar{\mathbb{F}}_p$ denotes the reduction morphism. For ζ in $\bar{\mathbb{F}}_p$, denote by $\boldsymbol{D}(\zeta) := \pi^{-1}(\zeta)$ the residue disc corresponding to ζ.

2A. *Divisors.* A *divisor* on a set X[†] is a formal finite sum $\sum_{x\in X} n_x x$ in $\bigoplus_{x\in X} \mathbb{Z}x$. In the special case where for some x_0 in X we have $n_{x_0} = 1$ and $n_x = 0$ for every $x \neq x_0$, we use $[x_0]$ to denote this divisor. When there is no danger of confusion, sometimes we use x_0 to denote $[x_0]$.

Let $\mathfrak{D} = \sum_{x\in X} n_x[x]$ be a divisor on X. The *degree* and the *support* of $\mathfrak{D}$ are defined by

$$\deg(\mathfrak{D}) := \sum_{x\in X} n_x \quad \text{and} \quad \operatorname{supp}(\mathfrak{D}) := \{x \in X : n_x \neq 0\},$$

respectively. The divisor $\mathfrak{D}$ is *effective*, if for every x in X we have $n_x \geq 0$. For $A \subseteq X$, the *restriction of $\mathfrak{D}$ to A* is the divisor on X defined by

$$\mathfrak{D}|_A := \sum_{x\in A} n_x[x].$$

For a set X' and a map $f\colon X \to X'$, the *push-forward action of f on divisors* $f_*\colon \operatorname{Div}(X) \to \operatorname{Div}(X')$ is the linear extension of the action of f on points. In the particular case in which $X' = G$ is a commutative group, also define $f\colon \operatorname{Div}(X) \to G$ by

$$f(\mathfrak{D}) := \sum_{x\in X} n_x f(x) \in G.$$

If X is a topological space and $\mathfrak{D}$ is an effective divisor satisfying $\deg(\mathfrak{D}) \geq 1$, then $\bar{\delta}_{\mathfrak{D}} := \frac{1}{\deg(\mathfrak{D})} \sum_{x\in X} n_x \delta_x$ is a Borel measure on X. Note that in the case $G = \mathbb{R}$ and f is measurable, we have

$$\int f \, \mathrm{d}\bar{\delta}_{\mathfrak{D}} = \frac{f(\mathfrak{D})}{\deg(\mathfrak{D})}.$$

Since we are identifying $Y(\mathbb{C}_p)$ with $\mathbb{C}_p$ via j, we identify divisors on $Y(\mathbb{C}_p)$ and on $\mathbb{C}_p$ accordingly.

[†] We only use this definition in the case X is one of several types of one-dimensional objects. For such X, the notion of divisor introduced here can be seen as a natural extension of the usual notion of Weil divisor.

2B. *Hecke correspondences.* In this section we recall the construction and main properties of the Hecke correspondences. For details we refer the reader to [Shimura 1971, Sections 7.2 and 7.3] for the general theory, or to the survey [Diamond and Im 1995, Part II].

Let $\mathbb{K}$ be an algebraically closed field of characteristic 0. First, note that for every integer $n \geq 1$ and divisor $\mathfrak{D}$ in $\mathrm{Div}(Y(\mathbb{K}))$, we have

$$\deg(T_n(\mathfrak{D})) = \sigma_1(n) \deg(\mathfrak{D}).$$

Moreover, for $n = 1$ the correspondence T_1 is by definition the identity on $\mathrm{Div}(Y(\mathbb{K}))$.

We also consider the linear extension of Hecke correspondences to $\mathrm{Div}(Y(\mathbb{K})) \otimes \mathbb{Q}$.

For an integer $N \geq 1$, denote by $Y_0(N)$ the *modular curve of level N*. It is a quasiprojective variety defined over $\mathbb{Q}$. The points of $Y_0(N)$ over $\mathbb{K}$ parametrize the moduli space of equivalence classes of pairs (E, C), where E is an elliptic curve over $\mathbb{K}$ and C is a cyclic subgroup of E of order N. Here, two such pairs (E, C) and (E', C') are equivalent if there exists an isomorphism $\phi\colon E \to E'$ over $\mathbb{K}$ taking C to C'. In particular, when $N = 1$, for every algebraically closed field $\mathbb{K}$ we can parametrize $Y(\mathbb{K})$ by $Y_0(1)(\mathbb{K})$, and $Y_0(1)$ is isomorphic to the affine line $\mathbb{A}^1_{\mathbb{Q}}$.

For $N > 1$, denote by $\Phi_N(X, Y)$ the *modular polynomial of level N*, which is a symmetric polynomial in $\mathbb{Z}[X, Y]$ that is monic in both X and Y, see, e.g., [Lang 1973, Chapter 5, Sections 2 and 3]. This polynomial is characterized by the equality

$$\Phi_N(j(E), Y) = \prod_{C \leq E \text{ cyclic of order } N} (Y - j(E/C)) \quad \text{for every } E \text{ in } Y(\mathbb{K}). \tag{2-3}$$

This implies that a birational model for $Y_0(N)$ is provided by the plane algebraic curve

$$\Phi_N(X, Y) = 0. \tag{2-4}$$

For each prime q, let $\alpha_q, \beta_q\colon Y_0(q) \to Y_0(1)$ be the rational maps over $\mathbb{Q}$ given in terms of moduli spaces by

$$\alpha_q(E, C) := E \quad \text{and} \quad \beta_q(E, C) := E/C.$$

In terms of the model (2-4) with $N = q$, the rational maps α_q and β_q correspond to the projections on the X and Y coordinate, respectively. Denote by $(\alpha_q)_*$ and $(\beta_q)_*$ the push-forward action of α_q and β_q on divisors, respectively, as in Section 2A. Denote also by α_q^* the pull-back action of α_q on divisors, defined at x in $Y_0(1)(\mathbb{K})$ by

$$\alpha_q^*(x) := \sum_{\substack{y \in Y_0(q)(\mathbb{K}) \\ \alpha_q(y) = x}} \deg_{\alpha_q}(y)[y],$$

where $\deg_{\alpha_q}(y)$ is the local degree of α_q at y. This definition is extended by linearity to arbitrary divisors. The pull-back action β_q^* of β_q is defined in a similar way. Then the Hecke correspondence $T_q\colon \mathrm{Div}(Y(\mathbb{K})) \to \mathrm{Div}(Y(\mathbb{K}))$ is recovered as

$$T_q = (\alpha_q)_* \circ \beta_q^* = (\beta_q)_* \circ \alpha_q^*,$$

where the second equality follows from the first and from the symmetry of T_q.

For an arbitrary integer $n \geq 2$, the correspondence T_n can be recovered from different T_q, for q running over prime divisors of n, by using the identities

$$T_{q^r} = T_q \circ T_{q^{r-1}} - q \cdot T_{q^{r-2}} \quad \text{for } q \text{ prime and } r \geq 2; \tag{2-5}$$

$$T_\ell \circ T_m = T_{\ell m} \quad \text{for } \ell, m \geq 1 \text{ coprime.} \tag{2-6}$$

We conclude this section with the following lemma used in Sections 3A and 5B.

Lemma 2.1. *Let $n \geq 1$ be an integer. For E in $Y(\mathbb{C}_p)$, the divisor $T_n(E)$ varies continuously with respect to E in the following sense: For every commutative topological group G and every continuous function $f\colon Y(\mathbb{C}_p) \to G$, the function $T_n f\colon Y(\mathbb{C}_p) \to G$ given by*

$$T_n f(E) := f(T_n(E))$$

is continuous. In particular, for every open and closed subset $A \subseteq Y(\mathbb{C}_p)$, the integer valued map

$$E \mapsto \deg(T_n(E)|_A)$$

is locally constant.

Proof. We first treat the case where n equals a prime number q. Let $P_0(X), \ldots, P_q(X)$ be the polynomials in $\mathbb{Z}[X]$ such that

$$\Phi_q(X, Y) = P_0(X) + P_1(X)Y + \cdots + P_q(X)Y^q + Y^{q+1}.$$

Let $(E_m)_{m=1}^{\infty}$ be a sequence and E_0 be a point in $Y(\mathbb{C}_p)$, such that $j(E_m) \to j(E_0)$ when m tends to infinity. Then for every k in $\{0, 1, \ldots, q\}$, we have $P_k(j(E_m)) \to P_k(j(E_0))$ when m tends to infinity. It follows that the roots of the polynomial $\Phi_q(j(E_m), Y)$ converge to the roots of $\Phi_q(j(E_0), Y)$, in the following sense: For every m in $\{0, 1, 2, \ldots\}$ we can find $z_{m,0}, \ldots, z_{m,q}$ in $\mathbb{C}_p$, so that

$$\Phi_q(j(E_m), Y) = \prod_{k=0}^{q} (Y - z_{m,k}),$$

and so that for every k in $\{0, 1, \ldots, q\}$ we have $z_{m,k} \to z_{0,k}$ when m tends to infinity, see for example [Brink 2006, Theorem 2]. For each m in $\{0, 1, 2, \ldots\}$ and k in $\{0, 1, \ldots, q\}$, let $E_{m,k}$ be the curve in $Y(\mathbb{C}_p)$ with $j(E_{m,k}) = z_{m,k}$. By the definition of T_q and (2-3), we have for every $m \geq 0$

$$T_q(E_m) = \sum_{k=0}^{q} [E_{m,k}].$$

Since for every k in $\{0, 1, \ldots, q\}$ we have $j(E_{m,k}) \to j(E_{0,k})$ when m tends to infinity, we conclude that for every continuous function $f\colon Y(\mathbb{C}_p) \to G$ we have

$$T_q f(E_m) = \sum_{k=0}^{q} f(E_{k,m}) \to \sum_{k=0}^{q} f(E_{k,0}) = T_q f(E_0).$$

This proves that $T_q f$ is continuous.

We now treat the general case by using multiplicative induction, the relations (2-5) and (2-6), and the fact that for every pair of linear maps $L, \tilde{L}\colon \operatorname{Div}(Y(\mathbb{C}_p)) \to \operatorname{Div}(Y(\mathbb{C}_p))$, every pair of integers $m, \tilde{m}$, and every function $F\colon Y(\mathbb{C}_p) \to G$, one has

$$(L \circ \tilde{L})(F) = \tilde{L}(L(F)) \quad \text{and} \quad (mL + \tilde{m}\tilde{L})(F) = mL(F) + \tilde{m}\tilde{L}(F). \tag{2-7}$$

Denote by I the set of those integers $n \geq 1$ such that for every continuous function $f\colon Y(\mathbb{C}_p) \to G$, the function $T_n(f)$ is also continuous. Clearly I contains 1, since for every function f we have $T_1(f) = f$. By the proof given above, I contains all prime numbers. Let $n \geq 1$ be a given integer having each divisor in I, and let q be a prime number. Let $s \geq 0$ and $n_0 \geq 1$ be the integers such that $n = q^s n_0$, and such that q does not divide n_0. Then by the relations (2-5) and (2-6), and by (2-7), we have

$$T_{qn}(f) = T_{q^{s+1}n_0}(f) = T_{n_0}(T_{q^{s+1}}(f)),$$

and for $s \geq 1$

$$T_{q^{s+1}}(f) = T_{q^s}(T_q(f)) - qT_{q^{s-1}}(f).$$

Since n_0, q, q^s, and q^{s-1} if $s \geq 1$, are all in I, we conclude that $T_{qn}(f)$ is continuous, and that qn is in I. This completes the proof of the multiplicative induction step, and of the first part of the lemma.

The second part of the lemma is an easy consequence of the first. Indeed, let $A \subseteq Y(\mathbb{C}_p)$ be an open and closed subset. Then the function $\mathbf{1}_A$ is continuous and the first part implies that

$$E \mapsto T_n \mathbf{1}_A(E) = \mathbf{1}_A(T_n(E)) = \deg(T_n(E)|_A)$$

is also continuous. But $T_n \mathbf{1}_A$ has integer values, hence it must be locally constant. This completes the proof of the lemma. □

2C. *Hecke orbits of CM points and an estimate on class numbers.* In this section we first recall a special case of a formula of Zhang describing the effect of Hecke correspondences on CM points (Lemma 2.2), which is used in Sections 3, 4 and 5B. To do this, and for the rest of the paper, for every discriminant D we consider Λ_D as a divisor. We also use Siegel's classical lower bound on class numbers of quadratic imaginary extensions of $\mathbb{Q}$, to give the following estimate used in the proof of Theorem A: For every $\varepsilon > 0$ there is a constant $C > 0$ such that for every negative discriminant D, we have

$$h(D) := \deg(\Lambda_D) \geq C|D|^{1/2-\varepsilon}. \tag{2-8}$$

In this section we follow [Clozel and Ullmo 2004, Section 2.3], adding some details for the benefit of the reader.

We use d to denote a negative fundamental discriminant. For each discriminant D there is a unique negative fundamental discriminant d and integer $f \geq 1$ such that $D = df^2$. These are the *fundamental discriminant* and *conductor of* D, respectively. We denote by $\mathcal{O}_{d,f}$ the unique order of discriminant D in the quadratic imaginary extension $\mathbb{Q}(\sqrt{d})$ of $\mathbb{Q}$ and put

$$w_{d,f} := \#(\mathcal{O}_{d,f}^{\times}/\mathbb{Z}^{\times}) = (\#\mathcal{O}_{d,f}^{\times})/2.$$

The integer f is the index of $\mathcal{O}_{d,f}$ inside the ring of integers of $\mathbb{Q}(\sqrt{d})$. Note that $w_{-3,1} = 3$, $w_{-4,1} = 2$, and that in all the remaining cases $w_{d,f} = 1$.

Recall that the *Dirichlet convolution* of two functions $g, \tilde{g} \colon \mathbb{N} \to \mathbb{C}$, is defined by

$$(g * \tilde{g})(n) := \sum_{d \in \mathbb{N}, d \,|\, n} g(d)\tilde{g}\left(\frac{n}{d}\right).$$

Given a fundamental discriminant d, denote by $R_d \colon \mathbb{N} \to \mathbb{N} \cup \{0\}$ the function that to each n in $\mathbb{N}$ assigns the number of integral ideals of norm n in the ring of integers of $\mathbb{Q}(\sqrt{d})$. Moreover, denote by R_d^{-1} the inverse of R_d with respect to the Dirichlet convolution.

Lemma 2.2. *For every fundamental discriminant $d < 0$ and any pair of coprime integers $f \geq 1$ and $\tilde{f} \geq 1$, we have the relations*

$$T_f\left(\frac{\Lambda_{d\tilde{f}^2}}{w_{d,\tilde{f}}}\right) = \sum_{f_0 \in \mathbb{N}, f_0 \,|\, f} R_d\left(\frac{f}{f_0}\right)\frac{\Lambda_{d(f_0\tilde{f})^2}}{w_{d,f_0\tilde{f}}}; \tag{2-9}$$

$$\frac{\Lambda_{d(f\tilde{f})^2}}{w_{d,f\tilde{f}}} = \sum_{f_0 \in \mathbb{N}, f_0 \,|\, f} R_d^{-1}\left(\frac{f}{f_0}\right) T_{f_0}\left(\frac{\Lambda_{d\tilde{f}^2}}{w_{d,\tilde{f}}}\right). \tag{2-10}$$

If in addition f is not divisible by p, then we have

$$\Lambda_{d(pf)^2} = \begin{cases} T_p\left(\frac{\Lambda_{df^2}}{w_{d,f}}\right) - 2\frac{\Lambda_{df^2}}{w_{d,f}} & \text{if } p \text{ splits in } \mathbb{Q}(\sqrt{d}); \\ T_p\left(\frac{\Lambda_{df^2}}{w_{d,f}}\right) - \frac{\Lambda_{df^2}}{w_{d,f}} & \text{if } p \text{ ramifies in } \mathbb{Q}(\sqrt{d}); \\ T_p\left(\frac{\Lambda_{df^2}}{w_{d,f}}\right) & \text{if } p \text{ is inert in } \mathbb{Q}(\sqrt{d}), \end{cases} \tag{2-11}$$

and for every integer $m \geq 2$ we have

$$\Lambda_{d(p^m f)^2} = \begin{cases} T_{p^m}\left(\frac{\Lambda_{df^2}}{w_{d,f}}\right) - 2T_{p^{m-1}}\left(\frac{\Lambda_{df^2}}{w_{d,f}}\right) + T_{p^{m-2}}\left(\frac{\Lambda_{df^2}}{w_{d,f}}\right) & \text{if } p \text{ splits in } \mathbb{Q}(\sqrt{d}); \\ T_{p^m}\left(\frac{\Lambda_{df^2}}{w_{d,f}}\right) - T_{p^{m-1}}\left(\frac{\Lambda_{df^2}}{w_{d,f}}\right) & \text{if } p \text{ ramifies in } \mathbb{Q}(\sqrt{d}); \\ T_{p^m}\left(\frac{\Lambda_{df^2}}{w_{d,f}}\right) - T_{p^{m-2}}\left(\frac{\Lambda_{df^2}}{w_{d,f}}\right) & \text{if } p \text{ is inert in } \mathbb{Q}(\sqrt{d}). \end{cases} \tag{2-12}$$

To prove this lemma, we first record the following identity, which is also used in the proof (2-8) below and of Lemma 5.5 in Section 5B. Let ψ_d be the quadratic character associated to $K = \mathbb{Q}(\sqrt{d})$, which is given by the Kronecker symbol $\left(\frac{d}{\cdot}\right)$, and denote by $\mathbf{1} \colon \mathbb{N} \to \mathbb{C}$ the constant function equal to 1. Then we have the equality of functions

$$R_d = \psi_d * \mathbf{1}. \tag{2-13}$$

In fact, if we denote by $\zeta(s)$ the Riemann zeta function, by $\zeta_K(s)$ the Dedekind zeta function associated to K, and by $L(\psi_d, s)$ the Dedekind L-function associated to ψ_d, then the formula above is equivalent to the factorization $\zeta_K(s) = \zeta(s)L(\psi_d, s)$, whose proof can be found for example in [Cohen 2007, Proposition 10.5.5 on page 219], or [Lang 1994, Chapter XII, Section 1, Theorem 1].

Proof of Lemma 2.2. From the Möbius inversion formula we deduce that (2-9) and (2-10) are equivalent. Hence, it is enough to prove (2-10). We have the following formula of Zhang

$$T_f\left(\frac{\Lambda_d}{w_{d,1}}\right) = \sum_{f_0 \in \mathbb{N},\, f_0 \mid f} R_d\left(\frac{f}{f_0}\right)\frac{\Lambda_{df_0^2}}{w_{d,f_0}}, \tag{2-14}$$

see for example [Clozel and Ullmo 2004, Lemme 2.6] or [Zhang 2001, Proposition 4.2.1]. Applying the Möbius inversion formula, one obtains

$$\frac{\Lambda_{df^2}}{w_{d,f}} = \sum_{f_0 \in \mathbb{N},\, f_0 \mid f} R_d^{-1}\left(\frac{f}{f_0}\right) T_{f_0}\left(\frac{\Lambda_d}{w_{d,1}}\right). \tag{2-15}$$

On the other hand, note that if f and $\tilde{f}$ in $\mathbb{N}$ are coprime, then by (2-6) and (2-15), we obtain (2-10).

Finally, (2-11) and (2-12) are a direct consequence of (2-9), (2-13) and the fact that $\psi_d(p) = 1$ (resp. $0, -1$) if p splits (resp. ramifies, is inert) in $\mathbb{Q}(\sqrt{d})$. □

To prove (2-8), recall from the theory of complex multiplication that for a fundamental discriminant d the number $h(d)$ equals the class number of the quadratic extension $\mathbb{Q}(\sqrt{d})$ of $\mathbb{Q}$, see for example [Cox 2013, Corollary 10.20]. A celebrated result by Siegel states that for every $\varepsilon > 0$ there exists a constant $C > 0$ such that for every fundamental discriminant $d < 0$ we have

$$h(d) \geq C|d|^{1/2-\varepsilon}, \tag{2-16}$$

see for example [Siegel 1935], or [Lang 1994, Chapter XVI, Section 4, Theorem 4]. On the other hand, by [Lang 1973, Chapter 8, Section 1, Theorem 7] for every integer $f \geq 2$ we have

$$h(df^2) = \frac{w_{d,f}}{w_d} h(d) f \prod_{q \mid f,\ \text{prime}} \left(\frac{q - \psi_d(q)}{q}\right). \tag{2-17}$$

Given $\varepsilon > 0$, there are C' in $]0, 1[$ and N in $\mathbb{N}$ such that $(q-1)/q \geq q^{-\varepsilon}$ for every $q > N$ and $(q-1)/q \geq C' q^{-\varepsilon}$ for every $2 \leq q \leq N$. Hence, for every integer $f \geq 2$ we have

$$\prod_{q \mid f,\ \text{prime}} \left(\frac{q - \psi_d(q)}{q}\right) \geq \prod_{q \mid f,\ \text{prime}} \left(\frac{q-1}{q}\right) \geq (C')^N \prod_{q \mid f,\ \text{prime}} q^{-\varepsilon} \geq (C')^N f^{-\varepsilon}.$$

Combined with (2-16) and (2-17), this completes the proof of (2-8).

2D. ***The Berkovich affine line over $\mathbb{C}_p$ and the Gauss point.*** We refer the reader to [Berkovich 1990] for the general theory of Berkovich spaces, and to [Baker and Rumely 2010, Chapter 1] for the special case of the Berkovich affine line over $\mathbb{C}_p$, which is the only Berkovich space of relevance in this paper.

The *Berkovich affine line over* $\mathbb{C}_p$, which we denote by $\mathbb{A}^1_{\text{Berk}}$, is a topological space defined as follows: As a set, $\mathbb{A}^1_{\text{Berk}}$ is the collection of all multiplicative seminorms on the polynomial ring $\mathbb{C}_p[X]$ that take values in $\mathbb{R}_0^+$ and that extend the p-adic norm $|\cdot|_p$ on $\mathbb{C}_p$. Hence, a point $x \in \mathbb{A}^1_{\text{Berk}}$ is given by a map

$x\colon \mathbb{C}_p[X] \to \mathbb{R}_0^+$ satisfying for every a in $\mathbb{C}_p$ and for all f and g in $\mathbb{C}_p[X]$,

$$x(a) = |a|_p, \quad x(f+g) \le x(f) + x(g) \quad \text{and} \quad x(fg) = x(f)x(g).$$

The topology of $\mathbb{A}^1_{\mathrm{Berk}}$ is the weakest topology such that for every $f \in \mathbb{C}_p[X]$, the function $\mathbb{A}^1_{\mathrm{Berk}} \to \mathbb{C}_p$ given by $x \mapsto x(f)$ is continuous. The topological space $\mathbb{A}^1_{\mathrm{Berk}}$ is Hausdorff, locally compact, metrizable and path-connected. It contains $\mathbb{C}_p$ as a dense subspace via the map $\iota\colon \mathbb{C}_p \to \mathbb{A}^1_{\mathrm{Berk}}$ given, for $z \in \mathbb{C}_p$ and $f \in \mathbb{C}_p[X]$, by $\iota(z)(f) := |f(z)|_p$. We identify divisors on $\mathbb{C}_p$ and on $\iota(\mathbb{C}_p)$ accordingly.

The *canonical point* or *Gauss point* x_{can} of $\mathbb{A}^1_{\mathrm{Berk}}$ is the Gauss norm

$$\sum_{n=0}^{N} a_n X^n \mapsto \sup\left\{\left|\sum_{n=0}^{N} a_n z^n\right|_p : z \in \mathcal{O}_p\right\} = \max\{|a_n|_p : n \in \{0, \dots, N\}\}.$$

Given $a \in \mathbb{C}_p$ and $r > 0$, define

$$\begin{aligned} \boldsymbol{D}(a, r) &:= \{x \in \mathbb{C}_p : |x - a|_p < r\}; \\ \boldsymbol{D}^\infty(a, r) &:= \{x \in \mathbb{C}_p : |x - a|_p > r\}; \\ \mathcal{D}(a, r) &:= \{x \in \mathbb{A}^1_{\mathrm{Berk}} : x(X - a) < r\}; \\ \mathcal{D}^\infty(a, r) &:= \{x \in \mathbb{A}^1_{\mathrm{Berk}} : x(X - a) > r\}. \end{aligned}$$

A basis of neighborhoods of x_{can} in $\mathbb{A}^1_{\mathrm{Berk}}$ is given by the collection of sets

$$\mathcal{A}(A; R) := \mathcal{D}(0, R) \cap \bigcap_{a \in A} \mathcal{D}^\infty(a, R^{-1}), \tag{2-18}$$

where $R > 1$ and A is a finite subset of $\mathcal{O}_p$.

We conclude this section with the following result. Recall that a sequence of Borel probability measures $(\mu_n)_{n\in\mathbb{N}}$ on a topological space X *converges weakly to a Borel measure* μ *on* X, if for every continuous and bounded function $f\colon X \to \mathbb{R}$ we have

$$\lim_{n\to\infty} \int f \, \mathrm{d}\mu_n = \int f \, \mathrm{d}\mu;$$

see, e.g., [Billingsley 1968, Section 1.1].

Lemma 2.3. *Let $(\mathfrak{D}_n)_{n\in\mathbb{N}}$ be a sequence of effective divisors on $\mathbb{C}_p$ such that for every n we have* $\deg(\mathfrak{D}_n) \ge 1$. *Then, the following are equivalent*:

(i) $\bar{\delta}_{\iota(\mathfrak{D}_n)} \to \delta_{x_{\mathrm{can}}}$ *weakly as $n \to \infty$.*

(ii) *For every $R > 1$ and every a in $\mathcal{O}_p$, we have for $\boldsymbol{D} = \boldsymbol{D}(a, R^{-1})$ and $\boldsymbol{D} = \boldsymbol{D}^\infty(a, R)$,*

$$\lim_{n\to\infty} \frac{\deg(\mathfrak{D}_n|_{\boldsymbol{D}})}{\deg(\mathfrak{D}_n)} = \lim_{n\to\infty} \bar{\delta}_{\mathfrak{D}_n}(\boldsymbol{D}) = 0.$$

For the reader's convenience we provide a self-contained proof of this lemma, which applies to the Berkovich affine line over an arbitrary complete and algebraically closed field. Using that $\mathbb{A}^1_{\mathrm{Berk}}$ is metrizable, the lemma can also be obtained as a direct consequence of the following observations: (i) is equivalent to the assertion that for every neighborhood $\mathcal{U}$ of x_{can} in $\mathbb{A}^1_{\text{Berk}}$ we have

$$\lim_{n\to\infty} \bar{\delta}_{\mathfrak{D}_n}(\mathcal{U}) = 1.$$

This last statement is equivalent to the contrapositive of (ii).

Proof of Lemma 2.3. Assume that (i) holds and let $R > 1$ and a in $\mathcal{O}_p$ be given. Note that the first equality in (ii) is a direct consequence of the definitions. To prove the second equality, take a continuous function $\phi\colon \mathbb{R}_0^+ \to [0, 1]$ satisfying $\phi(1) = 0$ and $\phi(t) = 1$ for $0 \le t \le R^{-1}$ and for $t \ge R$. Let $\alpha\colon \mathbb{A}^1_{\text{Berk}} \to \mathbb{R}$ be the continuous function given by $\alpha(x) = x(X - a)$ and put $F := \phi \circ \alpha$. By construction we have

$$F(x_{\text{can}}) = \phi(1) = 0 \quad \text{and} \quad F(x) = 1 \text{ for all } x \in \mathcal{D}(a, R^{-1}) \cup \mathcal{D}^\infty(a, R).$$

Using that for $z \in \mathbb{C}_p$ we have

$$z \in \boldsymbol{D}(a, R^{-1}) \Leftrightarrow \iota(z) \in \mathcal{D}(a, R^{-1}) \quad \text{and} \quad z \in \boldsymbol{D}^\infty(a, R) \Leftrightarrow \iota(z) \in \mathcal{D}^\infty(a, R), \tag{2-19}$$

we get

$$0 \le \bar{\delta}_{\mathfrak{D}_n}(\boldsymbol{D}(a, R^{-1}) \cup \boldsymbol{D}^\infty(a, R)) = \bar{\delta}_{\iota(\mathfrak{D}_n)}(\mathcal{D}(a, R^{-1}) \cup \mathcal{D}^\infty(a, R)) \le \int F \, \mathrm{d}\bar{\delta}_{\iota(\mathfrak{D}_n)}.$$

Since F is continuous and bounded, our hypothesis (i) implies that

$$\bar{\delta}_{\mathfrak{D}_n}(\boldsymbol{D}(a, R^{-1})) \to 0 \quad \text{and} \quad \bar{\delta}_{\mathfrak{D}_n}(\boldsymbol{D}^\infty(a, R)) \to 0 \text{ as } n \to \infty.$$

This completes the proof of the implication (i) $\Rightarrow$ (ii).

Now, assume that (ii) holds, let $F\colon \mathbb{A}^1_{\text{Berk}} \to \mathbb{R}$ be a continuous and bounded function and let $\varepsilon > 0$ be given. Since the sets (2-18) form a basis of neighborhoods of x_{can}, there are $R > 1$ and a finite subset A of $\mathcal{O}_p$ such that

$$|F(x) - F(x_{\text{can}})| < \varepsilon \quad \text{for all } x \in \mathcal{A}(A; R). \tag{2-20}$$

Let R' in $]1, R[$ be fixed. From the definition of $\mathcal{A} := \mathcal{A}(A; R)$, we have

$$\mathcal{A}' := \mathbb{A}^1_{\text{Berk}} \setminus \mathcal{A} \subseteq \mathcal{D}^\infty(0, R') \cup \bigcup_{a\in A} \mathcal{D}(a, (R')^{-1}).$$

Using (2-19) and (ii) with R replaced by R' and with a in $A \cup \{0\}$, we obtain

$$\begin{aligned}
\deg(\iota(\mathfrak{D}_n)|_{\mathcal{A}'}) &\le \deg(\iota(\mathfrak{D}_n)|_{\mathcal{D}^\infty(0,R')}) + \sum_{a\in A} \deg(\iota(\mathfrak{D}_n)|_{\mathcal{D}(a,(R')^{-1})}) \\
&= \deg(\mathfrak{D}_n|_{\boldsymbol{D}^\infty(0,R')}) + \sum_{a\in A} \deg(\mathfrak{D}_n|_{\boldsymbol{D}(a,(R')^{-1})}) \\
&= o(\deg(\iota(\mathfrak{D}_n))).
\end{aligned}$$

Together with our choice of $\mathcal{A}(A; R)$, this implies

$$\begin{aligned}\left|\int F \,\mathrm{d}\bar{\delta}_{\iota(\mathfrak{D}_n)} - F(x_{\mathrm{can}})\right| &\\ &\leq \left|\frac{F(\iota(\mathfrak{D}_n)|_{\mathcal{A}}) - F(x_{\mathrm{can}})\deg(\iota(\mathfrak{D}_n)|_{\mathcal{A}})}{\deg(\mathfrak{D}_n)}\right| + \left|\frac{F(\iota(\mathfrak{D}_n)|_{\mathcal{A}'}) - F(x_{\mathrm{can}})\deg(\iota(\mathfrak{D}_n)|_{\mathcal{A}'})}{\deg(\mathfrak{D}_n)}\right| \\ &\leq \varepsilon + 2\Big(\sup_{x\in\mathbb{A}^1_{\mathrm{Berk}}} |F(x)|\Big)\frac{\deg(\iota(\mathfrak{D}_n)|_{\mathcal{A}'})}{\deg(\iota(\mathfrak{D}_n))},\end{aligned}$$

and therefore

$$\limsup_{n\to\infty}\left|\int F \,\mathrm{d}\bar{\delta}_{\iota(\mathfrak{D}_n)} - F(x_{\mathrm{can}})\right| \leq \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, this completes the proof of the implication (ii) $\Rightarrow$ (i) and of the lemma. □

3. CM points in the ordinary reduction locus

The purpose of this section is to give a strengthened version of Theorem A in the case where all the discriminants in the sequence $(D_n)_{n=1}^{\infty}$ are p-ordinary (Theorem 3.5(ii) in Section 3B). An important tool is "the canonical branch $\boldsymbol{t}$" of T_p on $Y_{\mathrm{ord}}(\mathbb{C}_p)$, which is defined using the canonical subgroup in Section 3A. We use it to give, for every integer $m \geq 1$, a simple formula of T_{p^m} (Proposition 3.4 in Section 3A). Moreover, we show that p-ordinary CM points correspond precisely to the preperiodic points of $\boldsymbol{t}$ on $Y_{\mathrm{ord}}(\mathbb{C}_p)$ (Theorem 3.5(i)). Once these are established, Theorem 3.5(ii) follows from dynamical properties of $\boldsymbol{t}$ on $Y_{\mathrm{ord}}(\mathbb{C}_p)$ (Lemma 3.7). In Appendix B we extend and further study the canonical branch $\boldsymbol{t}$ of T_p.

We use properties of reduction morphisms that are stated in most of the classical literature only for elliptic curves over discrete valued fields. To extend the application of these results to elliptic curves over $\mathbb{C}_p$ we use the continuity of the Hecke correspondences (Lemma 2.1 in Section 2B). To this purpose, we introduce the following notation: $\mathbb{Q}_p^{\mathrm{unr}}$ is the maximal unramified extension of $\mathbb{Q}_p$ inside $\overline{\mathbb{Q}}_p$, and $\mathbb{C}_p^{\mathrm{unr}}$ is its completion. Then, $\mathbb{C}_p^{\mathrm{unr}}$ is an infinite degree extension of $\mathbb{Q}_p$ with the same valuation group and with residue field $\bar{\mathbb{F}}_p$. The algebraic closure $\overline{\mathbb{C}_p^{\mathrm{unr}}}$ of $\mathbb{C}_p^{\mathrm{unr}}$ inside $\mathbb{C}_p$ is dense in $\mathbb{C}_p$. Since $\overline{\mathbb{C}_p^{\mathrm{unr}}}$ can be written as the union of finite extensions of $\mathbb{C}_p^{\mathrm{unr}}$, it follows that every elliptic curve in $Y(\overline{\mathbb{C}_p^{\mathrm{unr}}})$ can be defined over a complete discrete valued field with residue field $\bar{\mathbb{F}}_p$. The same holds for finite subgroups and isogenies between elliptic curves over $\overline{\mathbb{C}_p^{\mathrm{unr}}}$.

In what follows, we use $Y_{\mathrm{ord}}(\overline{\mathbb{C}_p^{\mathrm{unr}}}) := Y_{\mathrm{ord}}(\mathbb{C}_p) \cap Y(\overline{\mathbb{C}_p^{\mathrm{unr}}})$.

3A. ***The canonical branch of T_p on $Y_{\mathrm{ord}}(\mathbb{C}_p)$.*** In this section we define a branch of the Hecke correspondence T_p on $Y_{\mathrm{ord}}(\mathbb{C}_p)$ that we use to give a simple description, for every integer $m \geq 1$, of T_{p^m} that is crucial in what follows (Proposition 3.4). See also Appendix B. We start recalling the following result describing the endomorphism ring of the reduction of a CM point in the ordinary locus.

Proposition 3.1 [Lang 1973, Chapter 13, Section 4, Theorem 12]. *Let $d < 0$ be a fundamental discriminant and let $f \geq 1$ and $m \geq 0$ be integers such that f is not divisible by p. Then, for an elliptic curve E defined over a discrete valued subfield of $\mathbb{C}_p$ having ordinary reduction,* $\operatorname{End}(E) \simeq \mathcal{O}_{d,p^m f}$ *implies that the reduction $\tilde{E}$ of E satisfies* $\operatorname{End}(\tilde{E}) \simeq \mathcal{O}_{d,f}$. *In particular, if* $\operatorname{End}(E)$ *is an order in a quadratic imaginary extension of $\mathbb{Q}$ whose conductor is not divisible by p, then the reduction map* $\operatorname{End}(E) \to \operatorname{End}(\tilde{E})$ *is an isomorphism.*

To define the canonical branch of T_p on $Y_{\text{ord}}(\mathbb{C}_p)$, we use the *canonical subgroup* of an elliptic curve E in $Y_{\text{ord}}(\overline{\mathbb{C}_p^{\text{unr}}})$, which is defined as the unique subgroup of order p of E in the kernel of the reduction morphism $E \to \tilde{E}$. Equivalently, $H(E)$ is the kernel of the reduction morphism $E[p] \to \tilde{E}[p]$. For an elliptic curve $e \in Y(\overline{\mathbb{F}}_p)$ denote by $\operatorname{Frob}\colon e \to e^{(p)}$ the Frobenius morphism, which is the isogeny given in affine coordinates by $(x, y) \mapsto (x^p, y^p)$.

Theorem 3.2. (i) *For E in $Y_{\text{ord}}(\overline{\mathbb{C}_p^{\text{unr}}})$ the reduction of $E/H(E)$ equals $\tilde{E}^{(p)}$ and every isogeny $\varphi\colon E \to E/H(E)$ whose kernel is equal to $H(E)$ reduces to the Frobenius morphism*

$$\operatorname{Frob}\colon \tilde{E} \to \tilde{E}^{(p)}.$$

Moreover, the kernel of the isogeny dual to φ is different from the canonical subgroup of $E/H(E)$.

(ii) *For each ordinary elliptic curve $e \in Y(\overline{\mathbb{F}}_p)$ there exists a unique elliptic curve $e^{\uparrow} \in Y(\overline{\mathbb{C}_p^{\text{unr}}})$ reducing to e for which the reduction map induces a ring isomorphism*

$$\operatorname{End}(e^{\uparrow}) \simeq \operatorname{End}(e).$$

(iii) *Given two ordinary elliptic curves $e_1, e_2 \in Y(\overline{\mathbb{F}}_p)$, the reduction map induces a group isomorphism*

$$\operatorname{Hom}(e_1^{\uparrow}, e_2^{\uparrow}) \simeq \operatorname{Hom}(e_1, e_2).$$

In particular, the Frobenius morphism $\operatorname{Frob}\colon e \to e^{(p)}$ *lifts to an isogeny $e^{\uparrow} \to (e^{(p)})^{\uparrow}$ with kernel $H(e^{\uparrow})$, and $e^{\uparrow}/H(e^{\uparrow}) = (e^{(p)})^{\uparrow}$.*

Proof. Item (i) follows from the definition of canonical subgroup and properties of reduction morphisms; see, e.g., [Diamond and Shurman 2005, Proof of Lemma 8.7.1]. Item (ii) is usually known as “Deuring’s lifting theorem”, see for example [Deuring 1941] or [Lang 1973, Chapter 13, Section 5, Theorem 14]. Item (iii) is another known consequence of Deuring’s work. To prove surjectivity, first note that every isogeny in $\operatorname{Hom}(e_1, e_2)$ can be written as a composition of Frobenius morphisms, of duals of Frobenius morphisms, and of an isogeny whose degree is not divisible by p. In view of items (i) and (ii), and of Proposition 3.1, we can restrict to the case of an isogeny of degree n not divisible by p. This case is a direct consequence of item (ii), and the fact that the reduction morphism $E \to \tilde{E}$ induces a bijective map $E[n] \to \tilde{E}[n]$, see for example [Silverman 2009, Chapter VII, Proposition 3.1(b)]. □

The following result is due to Tate in the case $p = 2$ and to Deligne in the general case. To state it, define

$$\begin{aligned} \boldsymbol{t} \colon Y_{\mathrm{ord}}(\overline{\mathbb{C}_p^{\mathrm{unr}}}) &\to Y_{\mathrm{ord}}(\overline{\mathbb{C}_p^{\mathrm{unr}}}) \\ E &\mapsto \boldsymbol{t}(E) := E/H(E), \end{aligned} \tag{3-1}$$

and for e in $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$ put

$$\delta_e := \begin{cases} 1 & \text{if } p \geq 5,\ j(e) \neq 0, 1728; \\ 3 & \text{if } p \geq 5,\ j(e) = 0; \\ 2 & \text{if } p \geq 5,\ j(e) = 1728; \\ 6 & \text{if } p = 3,\ j(e) = 0 = 1728; \\ 12 & \text{if } p = 2,\ j(e) = 0 = 1728. \end{cases} \tag{3-2}$$

Note that in all the cases $\delta_e = (\#\,\mathrm{Aut}(e))/2$; see, e.g., [Silverman 1994, Chapter III, Theorem 10.1].

Theorem 3.3. *For each e in $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$ choose β_e in $\boldsymbol{D}(j(e)) \cap \mathbb{Q}_p^{\mathrm{unr}}$, so that $\pi(\beta_e) = j(e)$, and put $\delta'_e := \delta_e$ if $\beta_e = 0$ and $p \neq 3$ or if $\beta_e = 1728$ and $p \neq 2$, and $\delta'_e := 1$ otherwise. Then, the map $\boldsymbol{t}$ admits an expansion of the form*

$$\boldsymbol{t}(z) = z^p + pk(z) + \sum_{e \in Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)} \sum_{n=1}^{\infty} \frac{A_n^{(e)}}{(z - \beta_e)^n}, \tag{3-3}$$

where $k(z)$ is a polynomial of degree $p - 1$ in z with coefficients in $\mathbb{Z}$, and for each $n \geq 1$ the coefficient $A_n^{(e)}$ belongs to $\mathbb{Q}_p(\{\beta_e : e \in Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)\})$ and

$$\mathrm{ord}_p(A_n^{(e)}) \geq \delta'_e \left(\frac{1}{p+1} + n \frac{p}{p+1} \right). \tag{3-4}$$

In particular, $\boldsymbol{t}(z)$ extends to a rigid analytic function $Y_{\mathrm{ord}}(\mathbb{C}_p) \to Y_{\mathrm{ord}}(\mathbb{C}_p)$ of degree p that we also denote by $\boldsymbol{t}$.

For $p \geq 5$, this result is proved in [Dwork 1969, Chapter 7]. In the case $\delta'_e > 1$, (3-4) can be obtained from the method of proof described in [loc. cit.], or from the estimate in [loc. cit., page 80] combined with the fact that $\mathrm{ord}_p(A_n^{(e)})$ is an integer and that $\beta_e = 0$ implies $p \equiv 2 \bmod 3$. For $p = 2$ and 3, this result is stated in [loc. cit., page 89] with a weaker version of (3-4). We provide the details of the proof when $p = 2$ and 3; see Proposition B.2 in Appendix B.

The theorem above implies that $\boldsymbol{t}$ extends to a rigid analytic map from $Y_{\mathrm{ord}}(\mathbb{C}_p)$ to itself. We denote this extension also by $\boldsymbol{t}$ and call it the *canonical branch of T_p on $Y_{\mathrm{ord}}(\mathbb{C}_p)$*.

For $z \in Y_{\mathrm{ord}}(\mathbb{C}_p)$, let $\boldsymbol{t}^*(z)$ be the divisor on $Y_{\mathrm{ord}}(\mathbb{C}_p)$ given by

$$\boldsymbol{t}^*(z) := \sum_{\substack{w \in Y_{\mathrm{ord}}(\mathbb{C}_p) \\ \boldsymbol{t}(w) = z}} \deg_{\boldsymbol{t}}(w)[w],$$

where $\deg_{\boldsymbol{t}}(w)$ is the local degree of $\boldsymbol{t}$ at w. Note that by Theorem 3.3 the rigid analytic map $\boldsymbol{t}\colon Y_{\mathrm{ord}}(\mathbb{C}_p) \to Y_{\mathrm{ord}}(\mathbb{C}_p)$ is of degree p, so for z in $Y_{\mathrm{ord}}(\mathbb{C}_p)$ we have

$$\deg(\boldsymbol{t}^*(z)) = p \quad \text{and} \quad \boldsymbol{t}_*(\boldsymbol{t}^*(z)) = p[z].$$

As usual, for an integer $i \geq 1$ we denote by $\boldsymbol{t}^i$ the i-th fold composition of $\boldsymbol{t}$ with itself. We also use $\boldsymbol{t}^0$ to denote the identity on $Y_{\mathrm{ord}}(\mathbb{C}_p)$.

Proposition 3.4. *For every E in $Y_{\mathrm{ord}}(\mathbb{C}_p)$ and every integer $m \geq 1$, we have*

$$T_{p^m}(E) = \sum_{i=0}^{m} (\boldsymbol{t}^*)^{m-i}([\boldsymbol{t}^i(E)]). \tag{3-5}$$

When $m = 1$, the relation (3-5) reads

$$T_p(E) = \boldsymbol{t}^*(E) + [\boldsymbol{t}(E)]. \tag{3-6}$$

See Theorem B.1 in Appendix B for an extension.

Proof. The relation (3-5) for $m \geq 2$ follows from (3-6) by induction using the recursive formula (2-5). To prove (3-6), first note that for E in $Y_{\mathrm{ord}}(\mathbb{C}_p)$ satisfying $\deg_{\boldsymbol{t}}(E) \geq 2$ we have $\boldsymbol{t}'(E) = 0$. Therefore there are at most a finite number of such E in the affinoid $Y_{\mathrm{ord}}(\mathbb{C}_p)$; see for example [Fresnel and van der Put 2004, Proposition 3.3.6]. It follows that for every E in $Y_{\mathrm{ord}}(\mathbb{C}_p)$ outside a finite set of exceptions, we have $\#\operatorname{supp}(\boldsymbol{t}^*(E)) = p$. Thus, the set D of all those E in $Y_{\mathrm{ord}}(\overline{\mathbb{C}_p^{\mathrm{unr}}})$ with this property is dense in $Y_{\mathrm{ord}}(\mathbb{C}_p)$. To prove (3-6) for E in D, use the definition of $T_p(E)$ and $\boldsymbol{t}(E)$, and Theorem 3.2(i), to obtain

$$T_p(E) = [\boldsymbol{t}(E)] + \sum_{\substack{C \leq E, \#C = p \\ C \neq H(E)}} [E/C] = [\boldsymbol{t}(E)] + \boldsymbol{t}^*(E).$$

To prove (3-6) for an arbitrary E in $Y_{\mathrm{ord}}(\mathbb{C}_p)$, first note that by Lemma 2.1 for every open and closed subset A of $Y_{\mathrm{ord}}(\mathbb{C}_p)$ the function

$$E \mapsto \mathbf{1}_A(T_n(E) - \boldsymbol{t}^*(E) - [\boldsymbol{t}(E)]) = \deg((T_n(E) - \boldsymbol{t}^*(E) - [\boldsymbol{t}(E)])|_A)$$

is continuous. Since it is equal to 0 on the dense subset D of $Y_{\mathrm{ord}}(\mathbb{C}_p)$, we conclude that it is constant equal to 0. Since this holds for every open and closed subset A of $Y_{\mathrm{ord}}(\mathbb{C}_p)$, this proves (3-6) and completes the proof of the lemma. □

3B. ***CM points as preperiodic points.*** The purpose of this section is to prove the following result. In the case where all the discriminants in the sequence $(D_n)_{n=1}^{\infty}$ are p-ordinary, Theorem A is a direct consequence of item (ii) of this result together with (2-8) and Lemma 2.3.

Given a set X and a map $T\colon X \to X$, a point x in X is *periodic* if for some integer $r \geq 1$ we have $T^r(x) = x$. Then the integer r is a *period of* x and the smallest such integer is the *minimal period of* x. Moreover, a point y is *preperiodic* if it is not periodic and if for some integer $m \geq 1$ the point $T^m(y)$ is periodic. We call the least such integer m the *preperiod of* y.

Theorem 3.5. *Let ζ in $\bar{\mathbb{F}}_p$ be the j-invariant of an ordinary elliptic curve and denote by r the minimal period of ζ under the Frobenius map $z \mapsto z^p$. Then there is a unique periodic point E_0 of $\boldsymbol{t}$ in $\boldsymbol{D}(\zeta)$. The minimal period of E_0 is r. Moreover, E_0 is a CM point and, if we denote by D_0 the discriminant of the endomorphism ring of E_0, then the conductor of D_0 is not divisible by p and the following properties hold*:

(i) *Given a discriminant D, the set $\operatorname{supp}(\Lambda_D|_{\boldsymbol{D}(\zeta)})$ is nonempty if and only if for some integer $m \geq 0$ we have $D = D_0 p^{2m}$. Moreover,*

$$\operatorname{supp}(\Lambda_{D_0}|_{\boldsymbol{D}(\zeta)}) = \{E_0\}$$

and for each integer $m \geq 1$ the set $\operatorname{supp}(\Lambda_{D_0 p^{2m}}|_{\boldsymbol{D}(\zeta)})$ is equal to the set of all the preperiodic points of $\boldsymbol{t}$ in $\boldsymbol{D}(\zeta)$ of preperiod m, and is contained in $\boldsymbol{t}^{-m}(\boldsymbol{t}^m(E_0))$. In particular, CM points in $Y_{\mathrm{ord}}(\mathbb{C}_p)$ correspond precisely to the periodic and preperiodic points of $\boldsymbol{t}$ in $Y_{\mathrm{ord}}(\mathbb{C}_p)$.

(ii) *For every disc $\boldsymbol{B}$ of radius strictly less than 1 contained in $\boldsymbol{D}(\zeta)$ there is a constant $C > 0$ such that for every discriminant $D < 0$, we have*

$$\deg(\Lambda_D|_{\boldsymbol{B}}) \leq C.$$

Remark 3.6. The natural directed graph associated to the dynamics of $\boldsymbol{t}$ on the set of ordinary CM points is a "$(p+1)$-volcano" in the sense of [Goren and Kassaei 2017, Section 2.1]. This follows from Theorem 3.5(i) and the fact that $\boldsymbol{t}$ is of degree p on $Y_{\mathrm{ord}}(\mathbb{C}_p)$ by Theorem 3.3. Note in particular that the "rim" is the directed subgraph associated to the dynamics of $\boldsymbol{t}$ on the set of its periodic points in $Y_{\mathrm{ord}}(\mathbb{C}_p)$. Moreover, on the set of preperiodic points of $\boldsymbol{t}$ in $Y_{\mathrm{ord}}(\mathbb{C}_p)$, the preperiod corresponds to the function "b" of [Goren and Kassaei 2017].

To prove Theorem 3.5, we describe the dynamics of $\boldsymbol{t}$ on $Y_{\mathrm{ord}}(\mathbb{C}_p)$ in Lemma 3.7 below. This description is mostly based on the fact that

$$\boldsymbol{t}(z) \equiv z^p \mod p\mathcal{O}_p, \tag{3-7}$$

see Theorem 3.3. We deduce from general considerations that each residue disc $\boldsymbol{D} \subseteq Y_{\mathrm{ord}}(\mathbb{C}_p)$ contains a unique periodic point z_0 of $\boldsymbol{t}$, that this point satisfies $|\boldsymbol{t}'(z_0)| < 1$, and that every point in $\boldsymbol{D}$ is asymptotic to z_0.[‡] The fact that no periodic point of $\boldsymbol{t}$ in $Y_{\mathrm{ord}}(\mathbb{C}_p)$ is a ramification point is used in a crucial way in the proof of the estimate (5-5) of Proposition 5.3 in Section 5B.

Lemma 3.7 (dynamics of $\boldsymbol{t}$ on $Y_{\mathrm{ord}}(\mathbb{C}_p)$). *Let e be an ordinary elliptic curve defined over $\bar{\mathbb{F}}_p$ and let $r \geq 1$ be the minimal period of $j(e)$ under the Frobenius map. Then, $e^{\uparrow}$ is the unique elliptic curve in $\boldsymbol{D}(j(e))$ that is periodic for $\boldsymbol{t}$. The minimal period of $e^{\uparrow}$ for $\boldsymbol{t}$ is r and $e^{\uparrow}$ is also characterized as the unique elliptic curve in $\boldsymbol{D}(j(e)) \cap \overline{\mathbb{C}_p^{\mathrm{unr}}}$ whose endomorphism ring is an order in an quadratic imaginary extension of $\mathbb{Q}$ of conductor not divisible by p. Moreover, if for every integer $i \geq 0$ we put $z_i := \boldsymbol{t}^i(e^{\uparrow})$, then the following properties hold*:

[‡]This is somewhat similar to the case of a rational map having good reduction equal to the Frobenius map, see for example [Rivera-Letelier 2003, Sections 3.1 and 4.5].

(i) *For each integer* $i \geq 0$ *we have* $0 < |\boldsymbol{t}'(z_i)|_p < 1$.

(ii) *There is* ρ *in* $]0, 1[$ *such that for every integer* $i \geq 0$ *and all* z *and* z' *in* $\boldsymbol{D}(z_i, \rho)$, *we have*

$$\deg_{\boldsymbol{t}}(z) = 1 \quad \textit{and} \quad |\boldsymbol{t}(z) - \boldsymbol{t}(z')|_p = |\boldsymbol{t}'(z_i)|_p \cdot |z - z'|_p.$$

In particular, $\boldsymbol{t}$ *is injective on* $\boldsymbol{D}(z_i, \rho)$.

(iii) *For every* $c \in]0, 1[$ *there exists* κ_c *in* $]0, 1[$ *such that for every integer* $i \geq 0$, *every* z *in* $\boldsymbol{D}(z_i, 1)$ *satisfying* $|z - z_i|_p \leq c$ *and every integer* $m \geq 1$, *we have*

$$|\boldsymbol{t}^m(z) - z_{i+m}|_p \leq \kappa_c^m |z - z_i|_p.$$

(iv) *For all* $i \geq 0$ *and* z *in* $\boldsymbol{D}(z_i, 1)$, *the sequence*

$$(|\boldsymbol{t}^m(z) - z_{i+m})|_p)_{m=0}^{\infty}$$

is nonincreasing and converges to 0.

Proof. We start by proving (i). Suppose for a contradiction that z_i is a ramification point of $\boldsymbol{t}$. Without loss of generality, assume that $i = 0$ and put $E := e^{\uparrow}$ and $E^p := (e^{(p)})^{\uparrow}$. By Proposition 3.4 with $m = 1$ there are distinct subgroups C and C' of E^p of order p such that

$$E^p/C = E^p/C' = E, C \neq H(E^p) \quad \text{and} \quad C' \neq H(E^p).$$

Let ψ (resp. ψ') be an isogeny $E^p \to E$ with kernel C (resp. C') and denote by $\hat{\psi}$ (resp. $\hat{\psi}'$) its dual isogeny. Then the kernel of $\hat{\psi}$ and of $\hat{\psi}'$ are both equal to $H(E)$. It follows that there is σ in $\mathrm{Aut}(E^p)$ such that $\sigma \circ \hat{\psi} = \hat{\psi}'$; see, e.g., [Silverman 2009, Chapter III, Corollary 4.11]. Since $\sigma \neq \pm 1$, we have $j(E^p) \in \{0, 1728\}$ and therefore $r = 1$, $\boldsymbol{t}(z_0) = z_0$ and $E^p = E$. In particular, C and C' are subgroups of E and $\psi, \psi' \in \mathrm{End}(E)$. The kernel of each of the reduced isogenies $\tilde{\psi}$ and $\tilde{\psi}'$ is equal to $e[p](\bar{\mathbb{F}}_p)$, so there is $\tilde{\alpha}$ in $\mathrm{Aut}(e)$ such that $\tilde{\alpha} \circ \tilde{\psi} = \tilde{\psi}'$. Since the reduction map $\mathrm{End}(E) \to \mathrm{End}(e)$ is an isomorphism by Theorem 3.2(ii), we can find an automorphism $\alpha \in \mathrm{Aut}(E)$ satisfying $\alpha \circ \psi = \psi'$. This implies that the kernel C of ψ is equal to the kernel C' of ψ', and we obtain a contradiction. This completes the proof that z_i is not a ramification point of $\boldsymbol{t}$ and therefore that $\boldsymbol{t}'(z_i) \neq 0$.

To prove that $|\boldsymbol{t}'(z_i)| < 1$ note that by Theorem 3.3, we can write

$$\boldsymbol{t}(w + z_i) - z_{i+1} = \boldsymbol{t}(w + z_i) - \boldsymbol{t}(z_i) = \sum_{n=1}^{\infty} B_n^{(i)} w^n, \tag{3-8}$$

where the coefficients $B_n^{(i)}$ belong to $\mathcal{O}_p$ and satisfy $|B_n^{(i)}|_p \leq \frac{1}{p}$ for $n \neq p$. Since $\boldsymbol{t}'(z_i) = B_1^{(i)}$, this completes the proof of (i).

To prove the assertions at the beginning of the lemma, for each integer $i \geq 0$ denote by $e^{(p^i)}$ the image of e by the i-th iterate of the Frobenius morphism. Then by Theorem 3.2(iii) we have

$$z_i = \boldsymbol{t}^i(e^{\uparrow}) = (e^{(p^i)})^{\uparrow} \in \pi^{-1}(j(e)^{p^i}).$$

It follows that z_0 is periodic of minimal period r for $\boldsymbol{t}$. To prove uniqueness, note that by (3-8) for every integer $i \geq 0$ and distinct z and z' in $D(z_i, 1)$ we have

$$|\boldsymbol{t}(z) - \boldsymbol{t}(z')|_p < |z - z'|_p. \tag{3-9}$$

Thus, there can be at most one periodic point of $\boldsymbol{t}$ in $\boldsymbol{D}(z_0, 1)$. Finally, combining Theorem 3.2(ii) and Proposition 3.1 we obtain that $e^{\uparrow}$ is the unique elliptic curve reducing to e and whose endomorphism ring is an order of conductor not divisible by p. This completes the proof of the assertions at the beginning of the proposition, so it only remains to prove (ii), (iii) and (iv).

To prove (ii), note that by (i) there is ρ in $]0, 1[$ so that for every i in $\{0, \dots, r-1\}$, we have

$$\max\{|B_n^{(i)}|_p \rho^{n-1} : n \geq 2\} \leq |B_1^{(i)}|_p.$$

Then by the ultrametric inequality for every integer $i \geq 0$ and $z \in \boldsymbol{D}(z_i, \rho)$ we have $|\boldsymbol{t}'(z)|_p = |B_1^{(i)}|_p$, which is different from 0 by (i). In particular, $\deg_{\boldsymbol{t}}(z_i) = 1$. Moreover, for z' in $\boldsymbol{D}(z_i, \rho)$ we have by the ultrametric inequality

$$|\boldsymbol{t}(z) - \boldsymbol{t}(z')|_p = |B_1^{(i)}|_p |z - z'|_p.$$

This completes the proof of (ii).

Item (iii) is a direct consequence of (3-8) with

$$\kappa_c := \max\{|B_n^{(i)}| c^{n-1} : n \geq 1, i \in \{0, \dots, r-1\}\},$$

noting that for every integer $n \geq 1$ and all integers $i, i' \geq 0$ such that $i - i'$ is divisible by r, we have $B_n^{(i')} = B_n^{(i)}$.

To prove item (iv), note that the fact that the sequence is nonincreasing follows from (3-9) and the fact that it converges to 0 form (iii) with $c = |z - z_i|$. This completes the proof the lemma. □

Proof of Theorem 3.5. The first assertions are given by Lemma 3.7.

To prove (i), note that Proposition 3.1 implies that if a discriminant $D < 0$ is such that $\operatorname{supp}(\Lambda_D|_{\boldsymbol{D}(\zeta)})$ is nonempty, then there is an integer $m \geq 0$ such that $D = D_0 p^{2m}$. On the other hand, Lemma 3.7 implies $\operatorname{supp}(\Lambda_{D_0}|_{\boldsymbol{D}(\zeta)}) = \{E_0\}$. Fix an integer $m \geq 1$ and note that by Lemma 3.7 for every integer $j \geq 1$ the point $E_j := \boldsymbol{t}^j(E_0)$ is the unique periodic point of $\boldsymbol{t}$ in $\boldsymbol{D}(\zeta^{p^j})$. So, if E is a preperiodic point of $\boldsymbol{t}$ in $\boldsymbol{D}(\zeta)$ of preperiod m, then $\boldsymbol{t}^m(E) = E_m$. This implies that the set of all preperiodic points of $\boldsymbol{t}$ in $\boldsymbol{D}(\zeta)$ of preperiod m is contained in $\boldsymbol{t}^{-m}(E_m)$ and is equal to

$$\boldsymbol{t}^{-m}(E_m) \setminus \boldsymbol{t}^{-(m-1)}(E_{m-1}) = \boldsymbol{t}^{-(m-1)}(\boldsymbol{t}^{-1}(E_m) \setminus \{E_{m-1}\}).$$

Since the degree of $\boldsymbol{t}$ is p and by Lemma 3.7(i) we have $\boldsymbol{t}'(E_{m-1}) \neq 0$, the set $\boldsymbol{t}^{-1}(E_m) \setminus \{E_{m-1}\}$ is nonempty and equal to $\operatorname{supp}(\boldsymbol{t}^*([E_m]) - [E_{m-1}])$. We thus conclude that the set of preperiodic points of $\boldsymbol{t}$ in $\boldsymbol{D}(\zeta)$ of preperiod m is equal to $\boldsymbol{t}^{-(m-1)}(\operatorname{supp}(\boldsymbol{t}^*([E_m]) - [E_{m-1}]))$ and it is nonempty. Thus, to complete the proof of (i) it is sufficient to show that the set of preperiodic points of $\boldsymbol{t}$ in $\boldsymbol{D}(\zeta)$ of preperiod

m is equal to $\mathrm{supp}(\Lambda_{D_0 p^{2m}}|_{\boldsymbol{D}(\zeta)})$. Note that by (2-11) and Proposition 3.4 we have

$$\mathrm{supp}(\boldsymbol{t}_*(\Lambda_{D_0})) \subseteq \mathrm{supp}(T_p(\Lambda_{D_0})) = \mathrm{supp}(\Lambda_{D_0}) \cup \mathrm{supp}(\Lambda_{D_0 p^2}).$$

By Lemma 3.7 the set $\mathrm{supp}(\Lambda_{D_0})$, hence $\mathrm{supp}(\boldsymbol{t}_*(\Lambda_{D_0}))$, is formed by periodic points of $\boldsymbol{t}$ while points in $\mathrm{supp}(\Lambda_{D_0 p^2})$ are not periodic. This implies

$$\boldsymbol{t}_*(\Lambda_{D_0}) = \Lambda_{D_0}. \tag{3-10}$$

Let d and f_0 be the fundamental discriminant and conductor of D_0, respectively. Since p splits in $\mathbb{Q}(\sqrt{d})$ we deduce that for every integer $k \geq 0$ we have $R_d(p^k) = k+1$. By (2-9), Proposition 3.4 and (3-10) we get

$$\mathrm{supp}((\boldsymbol{t}^*)^m(\Lambda_{D_0})) = \bigcup_{k=0}^{m} \mathrm{supp}(\Lambda_{D_0 p^{2k}}).$$

This implies the equality

$$\mathrm{supp}((\boldsymbol{t}^*)^m(\Lambda_{D_0})) \setminus \mathrm{supp}((\boldsymbol{t}^*)^{m-1}(\Lambda_{D_0})) = \mathrm{supp}(\Lambda_{D_0 p^{2m}}). \tag{3-11}$$

By Lemma 3.7 and (3-10) the set $\mathrm{supp}(\Lambda_{D_0}) \cap (\boldsymbol{D}(\zeta) \cup \boldsymbol{D}(\zeta^p) \cup \cdots \cup \boldsymbol{D}(\zeta^{p^{r-1}}))$ equals the set of periodic points of $\boldsymbol{t}$ in $\boldsymbol{D}(\zeta) \cup \boldsymbol{D}(\zeta^p) \cup \cdots \cup \boldsymbol{D}(\zeta^{p^{r-1}})$. By (3-11) we conclude that the set $\mathrm{supp}(\Lambda_{D_0 p^{2m}}|_{\boldsymbol{D}(\zeta)})$ equals the set of preperiodic points of $\boldsymbol{t}$ in $\boldsymbol{D}(\zeta)$ of preperiod m. This completes the proof of (i).

To prove (ii), let c in $]0, 1[$ be such that $\boldsymbol{B} \subseteq \boldsymbol{D}(z_0, c)$, let ρ and κ_c be given by Lemma 3.7 and let $M \geq 1$ be an integer such that $c\kappa_c^{rM} < \rho$. Let $D < 0$ be a discriminant and z in $\mathrm{supp}(\Lambda_D) \cap \boldsymbol{B}$ be given. By (i) there is an integer $m \geq 0$ such that $\boldsymbol{t}^{rm}(z) = E_0$. Assume by contradiction that the least integer m with this property satisfies $m > M$. Then by Lemma 3.7 and our choice of M we have

$$|\boldsymbol{t}^{rM}(z) - E_0|_p \leq c\kappa_c^{rM} < \rho.$$

On the other hand, $\boldsymbol{t}^{r(m-M)}$ is injective on $\boldsymbol{D}(z_0, \rho)$ by Lemma 3.7(ii) and it maps $\boldsymbol{t}^{rM}(z)$ and E_0 to E_0, so $\boldsymbol{t}^{rM}(z) = E_0$. This contradicts the minimality of m and proves that for every z in $\mathrm{supp}(\Lambda_D) \cap \boldsymbol{B}$ we have $\boldsymbol{t}^{rM}(z) = E_0$. Equivalently,

$$\mathrm{supp}(\Lambda_D|_{\boldsymbol{B}}) \subseteq \bigcup_{i=1}^{M} \boldsymbol{t}^{-ir}(E_0).$$

Since this last set is finite and independent of D, this proves (ii) and completes the proof of the theorem. □

4. CM points in the supersingular reduction locus

The goal of this section is to prove the following result on the asymptotic distribution of CM points in the supersingular reduction locus. From this result and Theorem 3.5(ii), we deduce Theorem A at the end of this section.

Theorem 4.1. *For every e in $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$ fix an arbitrary γ_e in $\boldsymbol{D}(j(e))$ and for r in $]0,1[$, put*

$$\boldsymbol{B}(r) := \bigcup_{e \in Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)} \boldsymbol{D}(\gamma_e, r).$$

Then the following properties hold:

(i) *For every r in $]0,1[$ there exists $m>0$ such that for every discriminant $D<0$ satisfying* $\operatorname{ord}_p(D) \geq m$, *we have* $\deg(\Lambda_D|_{\boldsymbol{B}(r)}) = 0$.

(ii) *For every $m>0$ there exists r in $]0,1[$ such that for every p-supersingular discriminant $D<0$ satisfying* $\operatorname{ord}_p(D) \leq m$, *we have* $\operatorname{supp}(\Lambda_D) \subseteq \boldsymbol{B}(r)$.

We present the proof of Theorem 4.1 in Section 4C below. In Section 4A we recall the definition of Katz' valuation. For that purpose, we briefly review Katz' theory of algebraic modular forms and the interpretation of the Eisenstein series E_{p-1} as an algebraic modular form over $\mathbb{Q} \cap \mathbb{Z}_p$. In Section 4B we use Katz–Lubin's extension of the theory of canonical subgroups to not too supersingular elliptic curves to give a description of the action of Hecke correspondences on the supersingular locus (Section 4B). For $p=2$ and 3, we also rely on certain congruences satisfied by certain Eisenstein series (Proposition A.1 in Appendix A). This description is used in the proof of Theorem 4.1 and also in Section 5C on Hecke orbits in the supersingular locus.

4A. ***Katz' valuation.*** In this section we define Katz' valuation, which is based on Katz' theory of algebraic modular forms, and give an explicit formula relating it to the j-invariant (Proposition 4.3).

For the reader's convenience we start with a short review of Katz' theory of algebraic modular forms. For details see [Katz 1973, Chapter 1]. Let $k \in \mathbb{Z}$ be an integer and let R_0 be a ring (commutative and with identity). Denote by R_0-Alg the category of R_0-algebras. Given an R_0-algebra R, define an elliptic curve E over R as a proper, smooth morphism of schemes $E \to \operatorname{Spec}(R)$, whose geometric fibers are connected curves of genus one, together with a section $\operatorname{Spec}(R) \to E$, and denote by $\Omega^1_{E/R}$ the invertible sheaf of differential forms of degree 1 of E over R. By replacing $\operatorname{Spec}(R)$ by an appropriate affine subset we can assume that $\Omega^1_{E/R}$ admits a nowhere vanishing global section. In this paper we assume, for simplicity, that this is always the case and denote by $\Omega^1_{E/R}(E)'$ the (nonempty) set of nowhere vanishing global sections of $\Omega^1_{E/R}$. An algebraic modular form F of weight k and level one over R_0 is a family of maps

$$F_R \colon \{(E,\omega) : E \text{ elliptic curve over } R, \omega \in \Omega^1_{E/R}(E)'\} \to R \quad (R \in R_0\text{-Alg}),$$

satisfying the following properties:

(i) $F_R(E,\omega)$ depends only on the isomorphism class of the pair (E,ω). More precisely, for every isomorphism of elliptic curves $\varphi \colon E \to E'$ over R, we have $F_R(E', \varphi_*\omega) = F_R(E,\omega)$. Here, $\varphi_*\omega$ denotes the push-forward of ω by φ.

(ii) $F_R(E, \lambda\omega) = \lambda^{-k} F_R(E,\omega)$ for every $\lambda \in R^\times$.

(iii) F_R is compatible with base change. Namely, for every R_0-algebra morphism $g\colon R \to R'$, for the base change $(E, \omega)_{R'}$ of (E, ω) to R' by g we have $F_{R'}((E, \omega)_{R'}) = g(F_R(E, \omega))$.

Taking into account property (iii), from now on we simply write F instead of F_R. Moreover, let R_1 be an R_0-algebra. Then, property (iii) ensures that F induces an algebraic modular form F_1 over R_1. We say that F_1 is the *base change* of F to R_1. We also say that F is a *lifting* of F_1 to R.

Let q be a formal variable and denote by $\mathrm{Tate}(q)$ the *Tate curve*, which is an elliptic curve over the field of fractions $\mathbb{Z}((q))$ of the ring of formal power series $\mathbb{Z}[\![q]\!]$; see [Katz 1973, Appendix 1]. The j-invariant of $\mathrm{Tate}(q)$ has the form

$$j(\mathrm{Tate}(q)) = \tfrac{1}{q} + 744 + \sum_{n=1}^{\infty} c_n q^n, \quad c_n \in \mathbb{Z}. \tag{4-1}$$

The *q-expansion* of an algebraic modular form F over R_0 as above is defined as the element $F(q) \in \mathbb{Z}((q)) \otimes_{\mathbb{Z}} R_0$ obtained by evaluating F at the pair $(\mathrm{Tate}(q), \omega_{\mathrm{can}})$ consisting of the Tate curve together with its canonical differential ω_{can}, both considered over $\mathbb{Z}((q)) \otimes_{\mathbb{Z}} R_0$. Moreover, F is said to be *holomorphic at infinity* if $F(q) \in \mathbb{Z}[\![q]\!] \otimes_{\mathbb{Z}} R_0$.

Now, we state a version of the *q-expansion principle*, which is a particular case of [Katz 1973, Corollary 1.9.1].

Theorem 4.2. *Let R_0 be a ring and let $K \supseteq R_0$ be a R_0-algebra. Let $k \in \mathbb{Z}$ be an integer and let F be an algebraic modular form over K of weight k, level one and holomorphic at infinity. Assume that $F(q) \in \mathbb{Z}((q)) \otimes_{\mathbb{Z}} R_0$. Then, F is the base change of a unique algebraic modular form over R_0 of weight k.*

There is a natural link between the previous theory and the classical theory of modular forms. We refer to [Katz 1973, Section A1.1] for details. For each classical holomorphic modular form of weight k and level one $f\colon \mathbb{H} \to \mathbb{C}$, there exists a unique algebraic modular form F over $\mathbb{C}$ associated to f that is holomorphic at infinity. The Fourier expansion at infinity of f and the q-expansion of F are related by

$$f(\tau) = \sum_{n=0}^{\infty} a_n e^{2\pi i n\tau} \quad \text{if and only if} \quad F(q) = \sum_{n=0}^{\infty} a_n q^n.$$

For an even integer $k \geq 4$, let E_k be the *normalized Eisenstein series*

$$\mathrm{E}_k(\tau) = 1 - \frac{2k}{B_k} \sum_{n=1}^{\infty} \sigma_{k-1}(n) e^{2\pi i n\tau}, \quad \tau \in \mathbb{H}.$$

Here, the symbol B_k denotes the k-th Bernoulli number and $\sigma_{k-1}(n) := \sum_{d \mid n, d>0} d^{k-1}$. The complex function E_k is a classical holomorphic modular form of weight k and level one. Then, this function induces an algebraic modular form over $\mathbb{C}$, which we also denote by E_k, having the q-expansion with rational coefficients

$$\mathrm{E}_k(q) = 1 - \frac{2k}{B_k} \sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n. \tag{4-2}$$

When $p \geq 5$ and $k = p-1$, the von Staudt–Clausen theorem ensures that $\operatorname{ord}_p\left((2k)B_k^{-1}\right) = 1$. In particular, the coefficients of the Fourier expansion of E_{p-1} lie in $\mathbb{Z}_{(p)} := \mathbb{Q} \cap \mathbb{Z}_p$. Hence, by Theorem 4.2, we can consider E_{p-1} as an algebraic modular form of weight $p-1$ over $\mathbb{Z}_{(p)}$. On the other hand, the same reasoning and a direct examination of the Fourier expansions of E_4 and E_6 allow us to consider these Eisenstein series as algebraic modular forms of weight four and six over $\mathbb{Z}$.

For E in $Y_{\mathrm{sups}}(\mathbb{C}_p)$, which we regard as an elliptic curve over $\mathcal{O}_p$, choose ω in $\Omega^1_{E/\mathcal{O}_p}(E)'$ and define *Katz' valuation*

$$v_p(E) := \begin{cases} \operatorname{ord}_p(\mathrm{E}_{p-1}(E,\omega)) & \text{if } p \geq 5; \\ \frac{1}{3}\cdot \operatorname{ord}_3(\mathrm{E}_6(E,\omega)) & \text{if } p = 3; \\ \frac{1}{4}\cdot \operatorname{ord}_2(\mathrm{E}_4(E,\omega)) & \text{if } p = 2. \end{cases}$$

Since for every λ in $\mathcal{O}_p^\times$ we have $\mathrm{E}_k(E,\lambda\omega) = \lambda^{-k}\,\mathrm{E}_k(E,\omega)$, this definition does not depend on the particular choice of ω. The above definition is motivated by the following considerations. The Hasse invariant A_{p-1} is the unique algebraic modular form of weight $p-1$ over $\mathbb{F}_p$ with q-expansion $A_{p-1}(q) = 1$; see [Katz 1973, Chapter 2]. When $p \geq 5$, the base change to $\mathbb{F}_p$ of the form E_{p-1} equals A_{p-1}. On the other hand, when p equals 2 or 3 it is not possible to lift A_{p-1} to an algebraic modular form of level one, holomorphic at infinity, over $\mathbb{Z}_{(p)}$. However, the base change of E_4 (resp. E_6) to $\mathbb{F}_2$ (resp. to $\mathbb{F}_3$) is A_1^4 (resp. A_2^3). See Appendix A for details.

Since the Hasse invariant vanishes at supersingular elliptic curves, for every E in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ we have that $0 < v_p(E) \leq \infty$. An elliptic curve E in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ is *not too supersingular* if $v_p(E) < p/(p+1)$, and it is *too supersingular* otherwise.

The following result gives an explicit relation between $v_p(E)$ and $j(E)$. For e in $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$, we use the number δ_e defined by (3-2) in Section 3A.

Proposition 4.3. *For each e in $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$, denote by j_e the j-invariant of the unique zero of* E_{p-1} (*resp.* $\mathrm{E}_4, \mathrm{E}_6$) *in $\boldsymbol{D}(e)$ if $p \geq 5$* (*resp.* $p = 2, 3$). *Then, for every E in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ we have*

$$v_p(E) = \sum_{e \in Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)} \frac{1}{\delta_e} \operatorname{ord}_p(j(E) - \mathrm{j}_e).$$

Moreover, if $p \geq 5$ and $\mathrm{j}_e \equiv 0$ (*resp.* $\mathrm{j}_e \equiv 1728$) $\bmod \mathcal{M}_p$, *then* $\mathrm{j}_e = 0$ (*resp.* $\mathrm{j}_e = 1728$). *In the case* $p = 2$ (*resp.* $p = 3$), $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$ *has a unique element e and* $\mathrm{j}_e = 0$ (*resp.* $\mathrm{j}_e = 1728$).

It follows from the proof of this proposition that for every e in $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$ the number j_e is algebraic over $\mathbb{Q}$ and is in the quadratic unramified extension of $\mathbb{Q}_p$. We note that in the case $\mathrm{j}_e \not\equiv 0, 1728 \mod \mathcal{M}_p$, the elliptic curve class whose j-invariant is j_e is not CM,[§] but it is "fake CM" in the sense of [Coleman and McMurdy 2006]; see Remark 4.4 below.

[§] In fact, j_e need not be an algebraic integer: For $p = 13$ (resp. 17, 19, 23) there is a unique e in $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$ whose j-invariant is different from 0 and 1728, and we have $\mathrm{j}_e = 2^7\cdot 3^3\cdot 5^3/691$ (resp. $2^{10}\cdot 3^3\cdot 5^3/3617$, $2^8\cdot 3^3\cdot 5^3\cdot 11/43867$, $2^8\cdot 3^3\cdot 5^3\cdot 41/(131\cdot 593)$).

Proof of Proposition 4.3. Assume $p \geq 5$, so $p-1 \not\equiv 2, 8 \mod 12$. We can thus write $p-1$ uniquely in the form $p-1 = 12m + 4\delta + 6\varepsilon$ with $m \geq 0$ integer and $\delta, \varepsilon \in \{0, 1\}$. The modular discriminant

$$\Delta(\tau) = e^{2\pi i \tau} \prod_{n=1}^{\infty} (1 - e^{2\pi i n \tau})^{24}, \qquad \tau \in \mathbb{H},$$

is a classical holomorphic modular form of weight 12 and level one; see, e.g., [Diamond and Shurman 2005, Sections 1.1 and 1.2]. The infinite product above shows that the Fourier coefficients of Δ are rational integers. Hence, Theorem 4.2 ensures that Δ can be considered as an algebraic modular form over $\mathbb{Z}$. At the level of classical modular forms, we have the identity

$$\mathrm{E}_{p-1} = \Delta^m \, \mathrm{E}_4^{\delta} \, \mathrm{E}_6^{\varepsilon} \, P(j),$$

where $P(X)$ is a monic polynomial over $\mathbb{Z}_{(p)}$ of degree m such that $P_{\mathrm{sups}}(X) := X^{\delta}(X - 1728)^{\varepsilon} P(X)$ reduces modulo p to the supersingular polynomial, i.e., the monic separable polynomial over $\mathbb{F}_p$ whose roots are the j-invariants of the supersingular elliptic curves over $\bar{\mathbb{F}}_p$; see, e.g., [Kaneko and Zagier 1998, Theorem 1]. Using the classical identities $\mathrm{E}_4^3 = \Delta j$ and $\mathrm{E}_6^2 = \Delta(j - 1728)$ we get

$$\mathrm{E}_{p-1}^{12} = \Delta^{p-1} j^{4\delta} (j - 1728)^{6\varepsilon} P(j)^{12}.$$

Theorem 4.2 ensures that the above identity also holds at the level of algebraic modular forms over $\mathbb{Z}_{(p)}$. Write

$$P_{\mathrm{sups}}(X) = \prod_{e \in Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)} (X - \mathrm{j}_e),$$

where $\mathrm{j}_e \in \boldsymbol{D}\big(j(e)\big)$ for each $e \in Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$. Now, for every pair (E, ω) over $\mathcal{O}_p$ having good reduction we have $\Delta(E, \omega) \in \mathcal{O}_p^{\times}$, hence

$$|\mathrm{E}_{p-1}(E, \omega)|_p^{12} = |j(E)|_p^{4\delta} |j(E) - 1728|_p^{6\varepsilon} \prod_{\substack{e \in Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p) \\ \mathrm{j}_e \not\equiv 0, 1728}} |j(E) - \mathrm{j}_e|_p^{12}.$$

Since $p \geq 5$, we have that $j = 0$ (resp. $j = 1728$) is supersingular at p if and only if $p \equiv 2 \bmod 3$ (resp. $p \equiv 3 \bmod 4$) [Silverman 2009, Chapter V, Examples 4.4 and 4.5]. This implies the result when $p \geq 5$. The cases $p = 2$ and 3 follow similarly from the formulas

$$|\mathrm{E}_4(E, \omega)|_2^3 = |j(E)|_2 \quad \text{and} \quad |\mathrm{E}_6(E, \omega)|_3^2 = |j - 1728|_3,$$

respectively. This completes the proof of the proposition. □

Remark 4.4. Let e in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ be such that $\mathrm{j}_e \not\equiv 0, 1728 \mod \mathcal{M}_p$, and let E_e be the elliptic curve class in $Y(\mathbb{C}_p)$ such that $j(E_e) = \mathrm{j}_e$. Then E_e is not CM, but it is "fake CM" in the sense of [Coleman and McMurdy 2006]. In particular, j_e is not a singular modulus over $\mathbb{C}_p$. To show that E_e is not CM, choose a field isomorphism $\mathbb{C}_p \simeq \mathbb{C}$ and τ_e in $\mathbb{H}$ such that $E_e(\mathbb{C}) \simeq \mathbb{C}/(\mathbb{Z} + \tau_e \mathbb{Z})$. It is sufficient to show that τ_e is transcendental over $\mathbb{Q}$; see, e.g., [Lang 1973, Chapter 1, Section 5]. The complex number τ_e must be a zero of the holomorphic function $\tau \mapsto \mathrm{E}_{p-1}(\tau)$. Since $j(\tau_e) = \mathrm{j}_e$ is different from 0 and 1728, it follows that τ_e is not equivalent to $\rho = \frac{1}{2}(1+\sqrt{-3})$ or $i = \sqrt{-1}$ under the action of the modular group $\mathrm{SL}_2(\mathbb{Z})$ by Möbius transformations on $\mathbb{H}$. Then [Kohnen 2003, Theorem 1] implies that τ_e is transcendental over $\mathbb{Q}$.

To see that E_e is fake CM, note first that, since the reduction modulo p of $P_{\mathrm{sups}}(X)$ is separable and splits completely over $\mathbb{F}_{p^2}$, by Hensel's lemma all roots of $P_{\mathrm{sups}}(X)$ are in the ring of integers $\mathcal{O}$ of the unramified quadratic extension of $\mathbb{Q}_p$. As j_e is a root of $P_{\mathrm{sups}}(X)$, this implies that E_e represents an elliptic curve over $\mathcal{O}$. Let $[p]_e$ and ϕ be the multiplication by p and the p^2-power Frobenius endomorphism on the supersingular curve e, respectively. Then there exists σ in $\mathrm{Aut}(e)$ satisfying $\sigma \circ [p]_e = \phi$; see [Silverman 2009, Chapter II, Corollary 2.12]. Since $j(e) = \pi(\mathrm{j}_e)$ is different from 0 and 1728, we have $\sigma = \pm 1$ and $\pm[p]_e = \phi$. Choose $\pi_0 = \pm p$ as a uniformizer of $\mathcal{O}$. The multiplication by π_0 map on the formal group $\mathcal{F}_{E_e}$ of E_e defines an endomorphism $f(X)$ of $\mathcal{F}_{E_e}$, satisfying

$$f(X) \equiv \pi_0 X \bmod X^2 \quad \text{and} \quad f(X) \equiv X^{p^2} \bmod \pi_0.$$

It follows that $\mathcal{F}_{E_e}$ is a Lubin–Tate formal group over $\mathcal{O}$; see [Hazewinkel 1978, Section 8], and compare with [Coleman and McMurdy 2006, Remark 3.4]. In particular $\mathrm{End}(\mathcal{F}_{E_e}) \simeq \mathcal{O}$ and therefore E_e is fake CM; see [Hazewinkel 1978, Theorem 8.1.5 and Proposition 23.2.6].

4B. ***Katz' kite.*** The goal of this section is to give the following description of the action of Hecke correspondences on the supersingular locus.

Proposition 4.5. *Let $\hat{v}_p \colon Y_{\mathrm{sups}}(\mathbb{C}_p) \to [0, p/(p+1)]$ be the map defined by*

$$\hat{v}_p := \min\left\{v_p, \frac{p}{p+1}\right\}.$$

Moreover, denote by $\mathrm{\tau}_0$ the identity on $\mathrm{Div}([0, p/(p+1)])$, *let $\mathrm{\tau}_1$ be the piecewise-affine correspondence on* $[0, p/(p+1)]$ *defined by*

$$\mathrm{\tau}_1(x) := \begin{cases} [px] + p[x/p] & \text{if } x \in [0, 1/(p+1)]; \\ [1-x] + p[x/p] & \text{if } x \in]1/(p+1), p/(p+1)], \end{cases}$$

and for each integer $m \geq 2$ define the correspondence $\mathrm{\tau}_m$ on $[0, p/(p+1)]$ *recursively, by*

$$\mathrm{\tau}_m := \mathrm{\tau}_1 \circ \mathrm{\tau}_{m-1} - p\mathrm{\tau}_{m-2}.$$

Then for every integer $m \geq 0$ and every integer $n_0 \geq 1$ not divisible by p, we have

$$(\hat{v}_p)_* \circ T_{p^m n_0}|_{Y_{\mathrm{sups}}(\mathbb{C}_p)} = \sigma_1(n_0) \cdot \mathrm{\tau}_m \circ (\hat{v}_p)_*.$$

See Figure 1 for the graph of the correspondence $\mathrm{\tau}_1$ and Lemma 5.7 in Section 5C for a formula of $\mathrm{\tau}_m$ for every $m \geq 0$.

The proof of Proposition 4.5 is given after a couple of lemmas. The following is a reformulation, in our setting, of a theorem of Katz–Lubin on the existence of canonical subgroups for elliptic curves that are not too supersingular; see [Katz 1973, Theorems 3.1 and 3.10.7] and also [Buzzard 2003, Theorem 3.3].

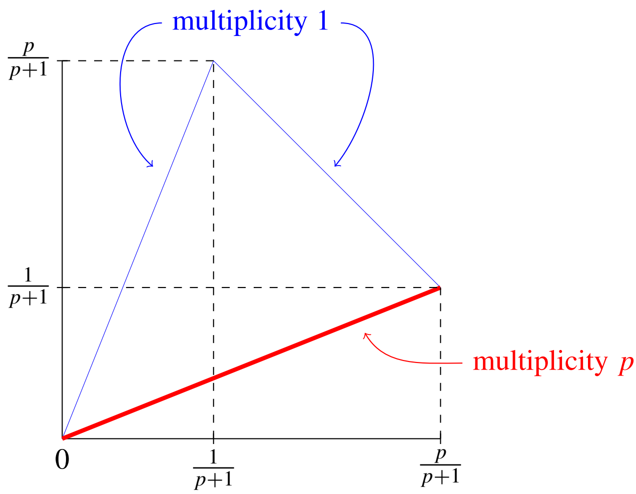

Figure 1. Graph of the correspondence τ_1 representing the action of T_p in terms of the projection $\hat{v}_p$.

Lemma 4.6. *For every elliptic curve E in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ that is not too supersingular there is a unique subgroup $H(E)$ of E of order p satisfying*

$$\hat{v}_p(E/H(E)) = \begin{cases} p v_p(E) & \text{if } v_p(E) \in \left]0, 1/(p+1)\right]; \\ 1 - v_p(E) & \text{if } v_p(E) \in \left]1/(p+1), p/(p+1)\right[. \end{cases} \tag{4-3}$$

Furthermore, $H(E)$ is also uniquely characterized by the property that for every subgroup C of E of order p that is different from $H(E)$, we have

$$v_p(E/C) = p^{-1} v_p(E). \tag{4-4}$$

In addition, the map

$$\begin{aligned} \boldsymbol{t} \colon \left\{E \in Y_{\mathrm{sups}}(\mathbb{C}_p) : v_p(E) < \tfrac{p}{p+1}\right\} &\to Y_{\mathrm{sups}}(\mathbb{C}_p) \\ E &\mapsto \boldsymbol{t}(E) := E/H(E) \end{aligned}$$

satisfies the following properties:

(i) *Let E be in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ and let C be a subgroup of E of order p. In the case $v_p(E) < p/(p+1)$, assume in addition that $C \neq H(E)$. Then*

$$v_p(E/C) = p^{-1} \hat{v}_p(E) \quad \textit{and} \quad \boldsymbol{t}(E/C) = E.$$

(ii) *For E in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ satisfying $1/(p+1) < v_p(E) < p/(p+1)$, we have $\boldsymbol{t}^2(E) = E$.*

Proof. For E in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ that is not too supersingular, note that the uniqueness statements about $H(E)$ follow from the fact that (4-3) and (4-4) imply that $H(E)$ is the unique subgroup C of E of order p satisfying $v_p(E/C) \neq p^{-1} v_p(E)$.

Assume $p \geq 5$ and let E be an elliptic curve in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ that is not too supersingular, so that $v_p(E) < p/(p+1)$. Let ω be a differential form in $\Omega^1_{E/\mathcal{O}_p}(E)'$ and put $r_E := \mathrm{E}_{p-1}(E, \omega) \in \mathcal{O}_p$. Since $\overline{\mathbb{C}_p^{\mathrm{unr}}}$ and $\mathbb{C}_p$ have the same valuation group we can find $r \in \overline{\mathbb{C}_p^{\mathrm{unr}}}$ satisfying $\mathrm{ord}_p(r) = \mathrm{ord}_p(r_E)$. Then r lies in the ring of integers R_0 of some finite extension of $\mathbb{C}_p^{\mathrm{unr}}$, and R_0 is a complete discrete valuation ring of residue characteristic p and generic characteristic zero. The triple $(E, \omega, r r_E^{-1})$ defines a r-situation in the sense of [Katz 1973, Theorem 3.1] (see also [loc. cit., Section 2.2]) and therefore there is a canonical subgroup $H(E)$ of E of order p. Then [loc. cit., Theorem 3.10.7(2, 3)] implies (4-3) and (ii), see also the proof of [Buzzard 2003, Theorem 3.3(iii)], and (4-4) and (i) are given by [Katz 1973, Theorem 3.10.7(5)]. Finally, note that for E in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ satisfying $v_p(E) \geq p/(p+1)$, the assertion (i) follows from [loc. cit., Theorem 3.10.7(4)]. This completes the proof of the proposition in the case $p \geq 5$.

It remains to prove the proposition in the cases $p = 2$ and $p = 3$. We only give the proof in the case $p = 2$, the case $p = 3$ being analogous. Let E_1 be an algebraic modular form of weight one and level n_1, with $3 \leq n_1 \leq 11$ odd, holomorphic at infinity and defined over $\mathbb{Z}[1/n_1]$ whose reduction modulo 2 is A_1; see Appendix A for details on level structures. Let E in $Y_{\mathrm{sups}}(\mathbb{C}_2)$ be an elliptic curve that is not too supersingular, let ω be a differential form in $\Omega^1_{E/\mathcal{O}_2}(E)'$ and α_{n_1} a level n_1 structure on E over $\mathcal{O}_2$. By Proposition A.1 and our hypothesis $v_2(E) < \frac{2}{3}$, we have

$$\mathrm{ord}_2(\mathrm{E}_1(E, \omega, \alpha_{n_1})) = v_2(E) < \tfrac{2}{3}.$$

Then, [Katz 1973, Theorem 3.1] gives the existence of $H(E)$ which might depend on the choice of α_{n_1}. The fact that $H(E)$ depends only on E follows from the characterization in [loc. cit., Theorem 3.10.7(1)] of the canonical subgroup as the subgroup of order 2 containing the unique point corresponding to the solution with valuation $1 - v_2(E)$ of the equation $[2](X) = 0$ in the formal group of E (here $[2]$ denotes the multiplication by 2 map and X is a certain normalized parameter for the formal group). Then (4-3), (4-4), (i) and (ii) follow from [loc. cit., Theorem 3.10.7] as in the case $p \geq 5$ above. This completes the proof of the lemma. □

Lemma 4.7. *Let E in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ be such that*

$$v_p(E) < \begin{cases} 1 & \text{if } p \geq 5; \\ (2p-1)/(2p) & \text{if } p = 2 \text{ or } 3. \end{cases}$$

Then for every subgroup C of E of order not divisible by p, we have $v_p(E/C) = v_p(E)$.

Proof. For E_0 in $Y(\mathbb{C}_p)$ and ζ in $\mathbb{Z}_p$, denote by $[\zeta]_{E_0}$ the multiplication by ζ map in the formal group of E_0.

Put $E' := E/C$ and denote by $\phi\colon E \to E'$ an isogeny with kernel C. Let X (resp. Y) be a parameter of the formal group of E (resp. E'), such that for any $(p-1)$-th root of unity $\zeta \in \mathbb{Z}_p$ we have $[\zeta]_E(X) = \zeta X$ (resp. $[\zeta]_{E'}(Y) = \zeta Y$); see [Katz 1973, Lemma 3.6.2(2)]. Let ω be a differential form in $\Omega^1_{E/\mathcal{O}_p}(E)'$ whose expansion in the parameter X is of the form

$$\omega = \left(1 + \sum_{n=1}^{\infty} a_n X^n\right) dX,$$

where $a_n \in \mathcal{O}_p$ for all $n \geq 1$. Then, by [loc. cit., Proposition 3.6.6] we have

$$[p]_E(X) = pX + aX^p + \sum_{m\geq 2} c_m X^{m(p-1)+1},$$

where $c_m \in \mathcal{O}_p$ for all $m \geq 2$ and $a \in \mathcal{O}_p$ satisfies

$$a \equiv A_{p-1}((E,\omega)_{\mathcal{O}_p/p\mathcal{O}_p}) \bmod p\mathcal{O}_p, \tag{4-5}$$

where $(E,\omega)_{\mathcal{O}_p/p\mathcal{O}_p}$ denotes the base change of (E,ω) to $\mathcal{O}_p/p\mathcal{O}_p$. Similarly,

$$[p]_{E'}(Y) = pY + a'Y^p + \sum_{m\geq 2} c'_m Y^{m(p-1)+1},$$

where $c'_m \in \mathcal{O}_p$ for all $m \geq 2$ and $a' \in \mathcal{O}_p$ satisfies, for some differential form ω' of $\Omega^1_{E'/\mathcal{O}_p}(E)'$,

$$a' \equiv A_{p-1}((E',\omega')_{\mathcal{O}_p/p\mathcal{O}_p}) \bmod p\mathcal{O}_p. \tag{4-6}$$

Since the order of $\operatorname{Ker}(\phi) = C$ is not divisible by p, the isogeny ϕ induces an isomorphism of formal groups of the form

$$\phi(X) = \sum_{n=1}^{\infty} t_n X^n,$$

where $t_n \in \mathcal{O}_p$ for all $n \geq 1$. Since $\phi(X)$ is invertible, we must have $t_1 \in \mathcal{O}_p^\times$. By the identity $[p]_{E'} \circ \phi = \phi \circ [p]_E$ we get

$$p(t_1X + t_2X^2 + t_3X^3 + \cdots) + a'(t_1X + t_2X^2 + t_3X^3 + \cdots)^p + \cdots$$
$$= t_1(pX + aX^p + \cdots) + t_2(pX + aX^p + \cdots)^2 + \cdots$$

Comparing the coefficients of X^p, we get

$$pt_p + a't_1^p = t_1 a + t_p p^p.$$

Using that $t_1 \in \mathcal{O}_p^\times$ we obtain

$$\operatorname{ord}_p(a') = \operatorname{ord}_p(a't_1^{p-1}) = \operatorname{ord}_p(a + t_1^{-1}t_p(p^p - p)). \tag{4-7}$$

In the case $p \geq 5$, (4-5) implies $\operatorname{ord}_p(a - \mathrm{E}_{p-1}(E,\omega)) \geq 1$, so by our hypothesis $v_p(E) < 1$ we have $\operatorname{ord}_p(a) = v_p(E) < 1$. Combined with (4-7), this implies $\operatorname{ord}_p(a') = \operatorname{ord}_p(a) = v_p(E) < 1$. Finally, by (4-6) we have $\operatorname{ord}_p(a' - \mathrm{E}_{p-1}(E',\omega')) \geq 1$, so $v_p(E') = \operatorname{ord}_p(a') = v_p(E)$. This proves the lemma in the case $p \geq 5$. For the case $p = 2$ or 3, (4-5), (4-6), (4-7), our hypothesis $v_p(E) < (2p-1)/(2p)$ and Proposition A.1 imply in a similar way

$$\operatorname{ord}_p(a) = v_p(E) < \frac{2p-1}{2p}, \quad \operatorname{ord}_p(a') = \operatorname{ord}_p(a) \quad \text{and} \quad v_p(E') = \operatorname{ord}_p(a').$$

This completes the proof of the lemma. □

Proof of Proposition 4.5. By the multiplicative property of Hecke correspondences (2-6) and Lemma 4.7, it is sufficient to consider the case $n_0 = 1$. Moreover, in view of (2-5) and the recursive definition of τ_m for $m \geq 2$, it is sufficient to consider the case $m = 1$. For E in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ satisfying $\hat{v}_p(E) < p/(p+1)$, this is given by (4-3) and (4-4) in Lemma 4.6, together with the fact that $\deg(T_p(E)) = p+1$. Finally, for E in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ satisfying $\hat{v}_p(E) = p/(p+1)$ the desired statement follows from Lemma 4.6(i). This completes the proof of the proposition. □

4C. ***Proof of Theorem 4.1.*** The proof of Theorem 4.1 is below, after a couple of lemmas.

Lemma 4.8. *Let $D < 0$ be a discriminant and let E and E' be in* $\operatorname{supp}(\Lambda_D)$. *Then, for every integer $m \geq 1$ there exists an isogeny $E \to E'$ of degree coprime to m.*

Proof. Denote by d and f the fundamental discriminant and conductor of D, respectively, and fix a field isomorphism $\mathbb{C}_p \simeq \mathbb{C}$. Since E and E' are CM with ring of endomorphisms isomorphic to $\mathcal{O}_{d,f}$, we can find proper fractional $\mathcal{O}_{d,f}$-ideals $\mathfrak{a}$ and $\mathfrak{a}'$ in $\mathbb{Q}(\sqrt{D})$ for which we have the complex uniformizations $E(\mathbb{C}) \simeq \mathbb{C}/\mathfrak{a}$ and $E'(\mathbb{C}) \simeq \mathbb{C}/\mathfrak{a}'$. Then there is a natural identification

$$\iota\colon \operatorname{Hom}(E, E') \to \mathfrak{a}'\mathfrak{a}^{-1} = \{\lambda \in \mathbb{C} : \lambda\mathfrak{a} \subseteq \mathfrak{a}'\}.$$

Without loss of generality, assume $\mathfrak{a}' \subset \mathfrak{a}$, and choose $\mathbb{Z}$-generators α and β of the ideal $\mathfrak{a}'\mathfrak{a}^{-1}$ of $\mathcal{O}_{d,f}$. Then

$$f(x, y) := (\alpha x - \beta y)\overline{(\alpha x - \beta y)}/[\mathcal{O}_{d,f} : \mathfrak{a}'\mathfrak{a}^{-1}]$$

is a positive definite primitive binary quadratic form with integer coefficients and discriminant d [Cox 2013, Theorem 7.7 and Exercise 7.17]. Moreover, there are integers x_0 and y_0 such that $f(x_0, y_0)$ is coprime to m [loc. cit., Lemma 2.25]. If we denote by ϕ_0 the isogeny in $\operatorname{Hom}(E, E')$ satisfying $\lambda_0 := \iota(\phi_0) = \alpha x_0 - \beta y_0$, then

$$\begin{aligned}
\deg(\phi_0) &= \#\operatorname{Ker}(\phi_0)\\
&= [\mathfrak{a}' : \lambda_0\mathfrak{a}]\\
&= [\mathfrak{a}'\mathfrak{a}^{-1} : \lambda_0\mathcal{O}_{d,f}]\\
&= [\mathcal{O}_{d,f} : \lambda_0\mathcal{O}_{d,f}]/[\mathcal{O}_{d,f} : \mathfrak{a}'\mathfrak{a}^{-1}]\\
&= \lambda_0\overline{\lambda_0}/[\mathcal{O}_{d,f} : \mathfrak{a}'\mathfrak{a}^{-1}]\\
&= f(x_0, y_0).
\end{aligned}$$

This proves that $\deg(\phi_0)$ is coprime to m, and completes the proof of the lemma. □

The following lemma is analogous to [Coleman and McMurdy 2006, Lemma 4.8], which concerns $p \geq 3$ in the context of certain modular curves of level bigger than 1. See also [Gross 1986, Proposition 5.3].

Lemma 4.9. *Let D be a p-supersingular discriminant and $m \geq 0$ the largest integer such that p^m divides the conductor of D. Then for every E in* $\operatorname{supp}(\Lambda_D)$ *we have*

$$\hat{v}_p(E) = \begin{cases} \frac{1}{2} \cdot p^{-m} & \text{if } p \text{ ramifies in } \mathbb{Q}(\sqrt{D}); \\ p/(p+1) \cdot p^{-m} & \text{if } p \text{ is inert in } \mathbb{Q}(\sqrt{D}).^{\P} \end{cases}$$

Proof. Let d be the fundamental discriminant of D and $f \geq 1$ the integer such that the conductor of D is equal to $p^m f$, so $D = d(fp^m)^2$ and f is not divisible by p. By Lemmas 4.7 and 4.8 with $m = p$, we deduce that for E in $\operatorname{supp}(\Lambda_D)$ the number $\hat{v}_p(D) := \hat{v}_p(E)$ is independent of E. By Zhang's formula (2-9) with $\tilde{f} = p^m$ it follows that there exists an isogeny of degree f from some elliptic curve in $\operatorname{supp}(\Lambda_{dp^{2m}})$ to an elliptic curve in $\operatorname{supp}(\Lambda_D)$. We conclude from Lemma 4.7 that $\hat{v}_p(D) = \hat{v}_p(dp^{2m})$. Thus, it is enough to prove the lemma in the case where $f = 1$.

We start with $m = 0$ and $m = 1$. By (2-11) with $f = 1$ and Proposition 4.5 with $m = 1$ and $n_0 = 1$, we have

$$\operatorname{supp}(\tau_1(\hat{v}_p(d))) = \begin{cases} \{\hat{v}_p(d), \hat{v}_p(dp^2)\} & \text{if } p \text{ ramifies in } \mathbb{Q}(\sqrt{d}); \\ \{\hat{v}_p(dp^2)\} & \text{if } p \text{ is inert in } \mathbb{Q}(\sqrt{d}). \end{cases}$$

From the definition of τ_1 we have that $p/(p+1)$ is the only value of x in $]0, p/(p+1)]$ such that $\tau_1(x)$ is supported on a single point. We conclude that if p is inert in $\mathbb{Q}(\sqrt{d})$, then $\hat{v}_p(d) = p/(p+1)$ and therefore $\hat{v}_p(dp^2) = 1/(p+1)$. On the other hand, $\frac{1}{2}$ is the only value of x in $]0, p/(p+1)]$ satisfying $x \in \operatorname{supp}(\tau_1(x))$. So, if p ramifies in $\mathbb{Q}(\sqrt{d})$, then $\hat{v}_p(d) = \frac{1}{2}$ and therefore $\hat{v}_p(dp^2) = \frac{1}{2}p^{-1}$. This completes the proof of the lemma when $m = 0$ and $m = 1$. Assume $m \geq 2$ and note that by (2-12) with $f = 1$ and by Proposition 4.5 with $n_0 = 1$,

$$\{\hat{v}_p(dp^{2m})\} = \begin{cases} \operatorname{supp}\big((\tau_m - \tau_{m-1})\big(\tfrac{1}{2}\big)\big) & \text{if } p \text{ ramifies in } \mathbb{Q}(\sqrt{d}); \\ \operatorname{supp}((\tau_m - \tau_{m-2})(p/(p+1))) & \text{if } p \text{ is inert in } \mathbb{Q}(\sqrt{d}). \end{cases}$$

From the definition of τ_m, we see that the right-hand side contains $\frac{1}{2} \cdot p^{-m}$ if p ramifies in $\mathbb{Q}(\sqrt{d})$ and $p/(p+1) \cdot p^{-m}$ if p is inert in $\mathbb{Q}(\sqrt{d})$. This proves $\hat{v}_p(dp^{2m}) = \frac{1}{2} \cdot p^{-m}$ in the former case and $\hat{v}_p(dp^{2m}) = p/(p+1) \cdot p^{-m}$ in the latter, and completes the proof of the lemma. □

Proof of Theorem 4.1. To prove (i), note that by Proposition 4.3 there is $m > 0$ so that $\hat{v}_p(\boldsymbol{B}(r)) \subseteq]p/(p+1) \cdot p^{-m}, p/(p+1)]$. Then by Lemma 4.9 for every p-supersingular discriminant $D < 0$ satisfying $\operatorname{ord}_p(D) \geq 2m + 3$ we have $\operatorname{supp}((\hat{v}_p)_*(\Lambda_D)) \cap \hat{v}_p(\boldsymbol{B}(r)) = \varnothing$, and therefore $\deg(\Lambda_D|_{\boldsymbol{B}(r)}) = 0$. On the other hand, if D is a p-ordinary discriminant, then $\operatorname{supp}(\Lambda_D) \subset Y_{\text{ord}}(\mathbb{C}_p)$ is disjoint from $\boldsymbol{B}(r)$, and therefore $\deg(\Lambda_D|_{\boldsymbol{B}(r)}) = 0$. This completes the proof of (i).

To prove (ii), note that by Proposition 4.3 there is r in $]0, 1[$ so that

$$\hat{v}_p^{-1}\left(\left[\tfrac{1}{2} \cdot p^{-m}, \tfrac{p}{p+1}\right]\right) \subseteq \boldsymbol{B}(r).$$

¶When $D = -3$ (resp. $D = -4$) is p-supersingular we have $j(E) = 1728$ (resp. 0) and $v_p(E) = \infty$, so in this formula we cannot replace the projection $\hat{v}_p$ by the valuation v_p. Compare with [Coleman and McMurdy 2006, Lemma 4.8].

Then by Lemma 4.9 for every p-supersingular discriminant $D < 0$ satisfying $\mathrm{ord}_p(D) \le m$ we have $\mathrm{supp}((\hat{v}_p)_*(\Lambda_D)) \subseteq \left[\frac{1}{2} \cdot p^{-m}, p/(p+1)\right]$ and therefore $\mathrm{supp}(\Lambda_D) \subseteq \boldsymbol{B}(r)$. This completes the proof of (ii) and of the theorem. □

Proof of Theorem A. In the case where all the discriminants in the sequence $(D_n)_{n=1}^{\infty}$ are p-ordinary (resp. p-supersingular), Theorem A is a direct consequence of Theorem 3.5(ii) (resp. Theorem 4.1), together with (2-8) and Lemma 2.3. The general case follows from these two special cases. □

5. Hecke orbits

The goal of this section is to prove Theorem C on the asymptotic distribution of Hecke orbits. The proof is divided into three complementary cases, according to whether the starting elliptic curve class has bad, ordinary or supersingular reduction. These are stated as Propositions 5.1, 5.2 and 5.6 in Sections 5A, 5B and 5C, respectively. In each case we prove a stronger quantitative statement.

5A. ***Hecke orbits in the bad reduction locus.*** In this section we prove a stronger version of the part of Theorem C concerning the bad reduction locus, which is stated as Proposition 5.1 below. We start by recalling some well-known results on the uniformization of p-adic elliptic curves with multiplicative reduction. See [Tate 1995] for the case of elliptic curves over complete discrete valued field, and [Roquette 1970] for the case of complete valued fields (see also [Silverman 1994, Chapter V, Theorem 3.1 and Remark 3.1.2]).

Let z be in $\boldsymbol{D}(0, 1)^* := \{z' \in \mathbb{C}_p : 0 < |z'|_p < 1\}$. We obtain, by the specialization $q = z$ in the Tate curve, an elliptic curve $\mathrm{Tate}(z)$ over $\mathbb{C}_p$ whose j-invariant satisfies

$$|j(\mathrm{Tate}(z))|_p = |z|_p^{-1} > 1, \tag{5-1}$$

see (4-1). This defines a bijective map

$$\begin{aligned} \boldsymbol{D}(0, 1)^* &\to Y_{\mathrm{bad}}(\mathbb{C}_p) \\ z &\mapsto \mathrm{Tate}(z). \end{aligned}$$

Moreover, for each $z \in \boldsymbol{D}(0, 1)^*$ there exists an explicit uniformization by $\mathbb{C}_p^\times$ of the set of $\mathbb{C}_p$-points of $\mathrm{Tate}(z)$. This uniformization induces an isomorphism of analytic groups $\varphi_z \colon \mathbb{C}_p^\times / z^{\mathbb{Z}} \to \mathrm{Tate}(z)(\mathbb{C}_p)$, see [Tate 1995, Theorem 1] for details. This allows us to give, for each integer $n \ge 1$, the following description of $T_n(\mathrm{Tate}(z))$. Note that for each positive divisor k of n and each $\ell \in \boldsymbol{D}(0, 1)^*$ satisfying $\ell^k = z^{n/k}$, the set

$$C_{n,\ell} := \{a \in \mathbb{C}_p^\times : a^{n/k} \in \ell^{\mathbb{Z}}\}/z^{\mathbb{Z}} \tag{5-2}$$

is a subgroup of order n of $\mathbb{C}_p^\times / z^{\mathbb{Z}}$. It is the kernel of the morphism of analytic groups $\mathbb{C}_p^\times / z^{\mathbb{Z}} \to \mathbb{C}_p^\times / \ell^{\mathbb{Z}}$ induced by the map $a \mapsto a^{n/k}$. Precomposing this morphism with φ_z^{-1} and then composing with φ_ℓ, we obtain an isogeny $\mathrm{Tate}(z) \to \mathrm{Tate}(\ell)$ of degree n whose kernel is $\varphi_z(C_{n,\ell})$. Since every subgroup of order

n of $\mathbb{C}_p^\times/z^{\mathbb{Z}}$ is of the form (5-2), we deduce that

$$T_n(\mathrm{Tate}(z)) = \sum_{\substack{k>0,\, k \mid n \\ \ell^k = z^{n/k}}} \mathrm{Tate}(\ell). \tag{5-3}$$

In the case where E is in $Y_{\mathrm{bad}}(\mathbb{C}_p)$, Theorem C is a direct consequence of the following result together with (2-1), (2-2) and Lemma 2.3.

Proposition 5.1. *Let z in $\boldsymbol{D}(0,1)^*$ and $R > 1$ be given. Then, for every $\varepsilon > 0$ there exists $C > 0$ such that for every integer $n \geq 1$ we have*

$$\deg(T_n(\mathrm{Tate}(z))|_{\boldsymbol{D}^\infty(0,R)}) \leq C n^{1/2} d(n).$$

Proof. Set $C := \sqrt{-\log(|z|_p)/\log(R)}$ and let $n \geq 1$ be an integer. By (5-1), for a positive divisor k of n and $\ell \in \boldsymbol{D}(0,1)^*$ with $\ell^k = z^{n/k}$, we have

$$|\mathrm{Tate}(\ell)|_p = |\ell|_p^{-1} = |z|_p^{-n/k^2}.$$

Noting that $|z|_p^{-n/k^2} > R$ is equivalent to $k < C n^{1/2}$, from (5-3) we deduce

$$\deg(T_n(\mathrm{Tate}(z))|_{\boldsymbol{D}^\infty(0,R)}) = \sum_{\substack{k>0, k \mid n \\ 0<k<C\sqrt{n}}} k < C n^{1/2} d(n).$$

This completes the proof of the proposition. □

5B. ***Hecke orbits in the ordinary reduction locus.*** The goal of this section is to prove the following result describing, for an elliptic curve E in $Y_{\mathrm{ord}}(\mathbb{C}_p)$, the asymptotic distribution of the Hecke orbit $(T_n(E))_{n=1}^\infty$. In the case where E is in $Y_{\mathrm{ord}}(\mathbb{C}_p)$, Theorem C with $n = p^m n_0$ is a direct consequence of this result together with (2-1) and Lemma 2.3.

Proposition 5.2. *Let $\boldsymbol{D}$ be a residue disc contained in $Y_{\mathrm{ord}}(\mathbb{C}_p)$ and let $\boldsymbol{B}$ be a disc of radius strictly less than 1 contained in $Y_{\mathrm{ord}}(\mathbb{C}_p)$. Then for every $\varepsilon > 0$ there is a constant $C > 0$ such that for every E in $\boldsymbol{D}$ and all integers $m \geq 0$ and $n_0 \geq 1$ such that n_0 is not divisible by p, we have*

$$\deg(T_{p^m n_0}(E)|_{\boldsymbol{B}}) \leq C(m+1)n_0^{\varepsilon}.$$

To prove Proposition 5.2 we use the multiplicative property of the Hecke correspondences, see (2-6) in Section 2B. We first treat the case $n_0 = 1$ (Propositions 5.3) and the case $m = 0$ (Propositions 5.4) separately. The proof of Proposition 5.2 is given at the end of this section.

Proposition 5.3. *Let ζ in $\bar{\mathbb{F}}_p$ be the j-invariant of an ordinary elliptic curve, denote by r the minimal period of ζ under the Frobenius map $z \mapsto z^p$ and put $\boldsymbol{O} := \bigcup_{i=0}^{r-1} \boldsymbol{D}(\zeta^{p^i})$. Then for every E in $\boldsymbol{D}(\zeta)$ and every integer $m \geq 1$, we have*

$$\mathrm{supp}(T_{p^m}(E)) \subseteq \boldsymbol{O}. \tag{5-4}$$

Moreover, for every disc $\boldsymbol{B}$ of radius strictly less than 1 contained in $\boldsymbol{O}$ there is a constant $C_1 > 0$ such that for every E in $\boldsymbol{O}$ and every integer $m \geq 1$, we have

$$\deg(T_{p^m}(E)|_{\boldsymbol{B}}) \leq C_1 m. \tag{5-5}$$

Proof. The inclusion (5-4) is a direct consequence of Proposition 3.4 and (3-7). To prove (5-5), let e be an ordinary elliptic curve with j-invariant ζ, for every integer $i \geq 0$ put $z_i := \boldsymbol{t}^i(e^{\uparrow})$ and for every integer $i \leq -1$ let i' be the unique integer in $\{0, \dots, r-1\}$ such that $i - i'$ is divisible by r and put $z_i := z_{i'}$. Note that for all nonnegative integers a, b, every integer i and every point z in $\boldsymbol{D}(z_i, 1)$, the set $\boldsymbol{t}^{-a}(\boldsymbol{t}^b(z))$ is contained in $\boldsymbol{D}(z_{i+b-a}, 1)$. Let c in $]0, 1[$ be such that $\boldsymbol{B}$ is contained in $\boldsymbol{B}(c) := \bigcup_{i=0}^{r-1} \boldsymbol{D}(z_i, c)$, let ρ and κ_c be given by Lemma 3.7 and let $i_1 \geq 0$ be a sufficiently large integer so that $c\kappa_c^{i_1} < \rho$.

Fix E in $\bigcup_{i=0}^{r-1} \boldsymbol{D}(z_i, 1)$ and let $m \geq 1$ be a given integer. Without loss of generality we assume $E \in \boldsymbol{D}(z_0, 1)$. We treat the cases $m < i_1$ and $m \geq i_1$ separately. If $m < i_1$, then we have

$$\deg(T_{p^m}(E)|_{\boldsymbol{B}(c)}) \leq \deg(T_{p^m}(E)) = \frac{p^{m+1}-1}{p-1} \leq p^{i_1} m.$$

Now, assume $m \geq i_1$. If for every i in $\{0, \dots, m\}$ the set $\boldsymbol{t}^{-(m-i)}(\boldsymbol{t}^i(E))$ is disjoint from $\boldsymbol{D}(z_{2i-m}, c)$, then

$$\deg(T_{p^m}(E)|_{\boldsymbol{B}(c)}) = \sum_{i=0}^{m} \deg((\boldsymbol{t}^*)^{(m-i)}([\boldsymbol{t}^i(E)])|_{\boldsymbol{D}(z_{2i-m}, c)}) = 0.$$

So we assume this is not the case and denote by i_0 the least integer i in $\{0, \dots, m\}$ such that $\boldsymbol{t}^{-(m-i)}(\boldsymbol{t}^i(E))$ contains a point E_0 in $\boldsymbol{D}(z_{2i-m}, c)$. Note that by Lemma 3.7(iii) the point $E_1 := \boldsymbol{t}^{i_1}(E_0)$ satisfies

$$|E_1 - z_{2i_0-m+i_1}|_p \leq c\kappa_c^{i_1} < \rho,$$

so it is in $\boldsymbol{D}(z_{2i_0-m+i_1}, \rho)$.

If $m \leq i_0 + i_1$, then we have

$$\deg(T_{p^m}(E)|_{\boldsymbol{B}(c)}) = \sum_{i=i_0}^{m} \deg((\boldsymbol{t}^*)^{m-i}([\boldsymbol{t}^i(E)])) \leq \sum_{i=i_0}^{m} p^{m-i} = \frac{p^{m-i_0+1}-1}{p-1} \leq p^{i_1}(m+1).$$

Suppose $m > i_0 + i_1$, and let i be an integer satisfying $i_0 \leq i \leq m - i_1$. Noting that for every E' in $\boldsymbol{t}^{-(m-i)}(\boldsymbol{t}^i(E))$ we have

$$\deg_{\boldsymbol{t}^{m-i}}(E') = \deg_{\boldsymbol{t}^{m-i-i_1}}(\boldsymbol{t}^{i_1}(E')) \deg_{\boldsymbol{t}^{i_1}}(E'),$$

we obtain

$$(\boldsymbol{t}^*)^{m-i}([\boldsymbol{t}^i(E)]) = \sum_{E'' \in \boldsymbol{t}^{-(m-i-i_1)}(\boldsymbol{t}^i(E))} \deg_{\boldsymbol{t}^{m-i-i_1}}(E'')(\boldsymbol{t}^*)^{i_1}([E'']). \tag{5-6}$$

On the other hand, for every z in $\boldsymbol{t}^{-(m-i)}(\boldsymbol{t}^i(E))$ contained in $\boldsymbol{D}(z_{2i-m}, c)$, we have by Lemma 3.7(iii) and our choice of i_1,

$$|\boldsymbol{t}^{i_1}(z) - z_{2i-m+i_1}|_p \leq c\kappa_c^{i_1} < \rho,$$

so $\boldsymbol{t}^{i_1}(z) \in \boldsymbol{D}(z_{2i-m+i_1}, \rho)$. Since for such z we have

$$\boldsymbol{t}^{m-i-i_1}(\boldsymbol{t}^{i_1}(z)) = \boldsymbol{t}^i(E) = \boldsymbol{t}^{m-i-i_1}(\boldsymbol{t}^{2i-2i_0}(E_1))$$

and by Lemma 3.7(ii) the map $\boldsymbol{t}^{m-i-i_1}$ is injective on $\boldsymbol{D}(z_{2i-m+i_1}, \rho)$, we conclude that $\boldsymbol{t}^{i_1}(z) = \boldsymbol{t}^{2i-2i_0}(E_1)$. Since we also have

$$\deg_{\boldsymbol{t}^{m-i-i_1}}(\boldsymbol{t}^{2i-2i_0}(E_1)) = 1$$

by Lemma 3.7(ii), when we restrict (5-6) to $\boldsymbol{D}(z_{2i-m}, c)$ we obtain

$$(\boldsymbol{t}^*)^{m-i}([\boldsymbol{t}^i(E)])|_{\boldsymbol{D}(z_{2i-m},c)} = (\boldsymbol{t}^*)^{i_1}([\boldsymbol{t}^{2i-2i_0}(E_1)])|_{\boldsymbol{D}(z_{2i-m},c)},$$

and therefore

$$\deg((\boldsymbol{t}^*)^{m-i}([\boldsymbol{t}^i(E)])|_{\boldsymbol{D}(z_{2i-m},c)}) \le \deg((\boldsymbol{t}^*)^{i_1}([\boldsymbol{t}^{2i-2i_0}(E_1)])) = p^{i_1}.$$

Together with Proposition 3.4 and our definition of i_0, this implies

$$\begin{aligned}\deg(T_{p^m}(E)|_{\boldsymbol{B}(c)}) &\le \sum_{i=i_0}^{m-i_1-1} \deg((\boldsymbol{t}^*)^{m-i}([\boldsymbol{t}^i(E)])|_{\boldsymbol{D}(z_{2i-m},c)}) + \sum_{i=m-i_1}^{m} \deg((\boldsymbol{t}^*)^{m-i}([\boldsymbol{t}^i(E)])) \\ &\le p^{i_1}(m-i_0-i_1) + \sum_{i=m-i_1}^{m} p^{m-i} \\ &\le p^{i_1}(m+1).\end{aligned}$$

This completes the proof of Proposition 5.3 with $C_1 = 2p^{i_1}$. □

Proposition 5.4. *Let $\boldsymbol{D}$ and $\boldsymbol{D}'$ be residue discs contained in $Y_{\mathrm{ord}}(\mathbb{C}_p)$. Then for every $\varepsilon > 0$ there is a constant $C_2 > 0$ such that for every E in $\boldsymbol{D}$ and every integer $n \ge 1$ that is not divisible by p, we have*

$$\deg(T_n(E)|_{\boldsymbol{D}'}) \le C_2 n^{\varepsilon}.$$

To prove this proposition we first establish an intermediate estimate.

Lemma 5.5. *Let e and e' be ordinary elliptic curves over $\bar{\mathbb{F}}_p$, and for each integer $n \ge 1$ denote by* $\mathrm{Hom}_n(e, e')$ *the set of isogenies from e to e' of degree n. Then, for every $\varepsilon > 0$ we have*

$$\#\,\mathrm{Hom}_n(e, e') = o(n^{\varepsilon}). \tag{5-7}$$

Proof. Assume there is a nonzero element ϕ_0 in $\mathrm{Hom}(e', e)$, for otherwise there is nothing to prove. Then, the map $\iota\colon \mathrm{Hom}(e, e') \to \mathrm{End}(e)$ given by $\iota(\phi) = \phi_0 \circ \phi$ is an injection, and $\deg(\iota(\phi)) = \deg(\phi_0)\deg(\phi)$. It is thus enough to prove (5-7) when $e' = e$.

Since e is ordinary, the ring $\mathrm{End}(e)$ is isomorphic to an order inside a quadratic imaginary extension K of $\mathbb{Q}$. Moreover, the isomorphism can be taken such that the degree of an isogeny is the same as the field norm of the corresponding element in K; see, e.g., [Silverman 2009, Chapter V, Theorem 3.1]. Let d be the discriminant of K. Then $\mathcal{O}_{d,1}$ is the ring of integers of K, and hence it is enough to show

$$\#\{x \in \mathcal{O}_{d,1} : x\bar{x} = n\} = o(n^{\varepsilon}).$$

Since the group of units $\mathcal{O}_{d,1}^{\times}$ is finite, this estimate follows from (2-2) and (2-13). □

Proof of Proposition 5.4. Let e be the ordinary elliptic curve over $\bar{\mathbb{F}}_p$ so that $\boldsymbol{D}' = \boldsymbol{D}(j(e))$. In view of Lemma 5.5, it is sufficient to show that for every E in $\boldsymbol{D}$ and every integer $n \geq 1$ that is not divisible by p we have

$$\deg(T_n(E)|_{\boldsymbol{D}'}) \leq \#\operatorname{Hom}_n(\tilde{E}, e). \tag{5-8}$$

Since the function $E \mapsto \deg(T_n(E)|_{\boldsymbol{D}'})$ is locally constant by Lemma 2.1, it is sufficient to establish this inequality in the case where E is in $Y_{\mathrm{ord}}(\overline{\mathbb{C}_p^{\mathrm{unr}}})$.

To prove (5-8), recall that the reduction morphism $E \to \tilde{E}$ induces a bijective map $E[n] \to \tilde{E}[n]$; see for example [Silverman 2009, Chapter VII, Proposition 3.1(b)]. In addition, note that for a subgroup C of E of order n such that $j(E/C)$ is in $\boldsymbol{D}'$, there is an isogeny $\tilde{E} \to e$ whose kernel is equal to the reduction of C. This defines an injective map

$$\{C \leq E : \#C = n,\ j(E/C) \in \boldsymbol{D}'\} \to \operatorname{Hom}_n(\tilde{E}, e),$$

proving (5-8) and completing the proof of the proposition. □

Proof of Proposition 5.2. Let ζ in $\bar{\mathbb{F}}_p$ be such that $\boldsymbol{B} \subseteq \boldsymbol{D}(\zeta)$, let $r \geq 1$ be the minimal period of ζ under the Frobenius map and put $\boldsymbol{O} := \bigcup_{i=0}^{r-1} \boldsymbol{D}(\zeta^{p^i})$. Let C_1 be given by Proposition 5.3 and let C_2 be the maximum value of the constants given by Proposition 5.4 with $\mathbf{D} = \mathbf{D}(\zeta), \dots, \mathbf{D}(\zeta^{p^{r-1}})$.

Let E in $\boldsymbol{D}$ be given. By (5-4), for every E' in $\operatorname{supp}(T_{n_0}(E))$ that is not in $\boldsymbol{O}$ we have

$$\deg(T_{p^m}(E')|_{\boldsymbol{B}}) \leq \deg(T_{p^m}(E')|_{\boldsymbol{O}}) = 0.$$

On the other hand, for every E' in $\operatorname{supp}(T_{n_0}(E))$ that is in $\boldsymbol{O}$, we have by Proposition 5.3

$$\deg(T_{p^m}(E')|_{\boldsymbol{B}}) \leq C_1 m + 1.$$

Together with (2-6) and Proposition 5.4 with $\boldsymbol{D}' = \boldsymbol{D}(\zeta), \dots, \boldsymbol{D}(\zeta^{p^{r-1}})$, this implies

$$\deg(T_{p^m n_0}(E)|_{\boldsymbol{B}}) \leq (C_1 m + 1)\deg(T_{n_0}(E)|_{\boldsymbol{O}}) \leq r C_2 (C_1 + 1)(m+1) n_0^{\varepsilon}.$$

This proves the theorem with $C = r C_2 (C_1 + 1)$. □

5C. *Hecke orbits in the supersingular reduction locus.* The purpose of this section is to prove the following result on Hecke orbits inside the supersingular reduction locus. In the case where E is in $Y_{\mathrm{sups}}(\mathbb{C}_p)$, Theorem C with $n = p^m n_0$ is a direct consequence of this result together with (2-1) and Lemma 2.3.

Proposition 5.6. *For every e in $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$ fix an arbitrary γ_e in $\boldsymbol{D}(j(e))$ and for every $r > 0$, put*

$$\boldsymbol{B}(r) := \bigcup_{e \in Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)} \boldsymbol{D}(\gamma_e, r).$$

Then the following properties hold:

(i) *For every r in $]0,1[$ there is a constant $C>0$ such that for every E in $Y_{\mathrm{sups}}(\mathbb{C}_p)$, every integer $m\geq 0$ and every integer $n_0\geq 1$ that is not divisible by p, we have*

$$\deg(T_{p^m n_0}(E)|_{\boldsymbol{B}(r)})\leq C\sigma_1(n_0).$$

(ii) *For every r_0 in $]0,1[$ and every integer $m_0\geq 0$, there is r in $]0,1[$ such that for every m in $\{0,\dots,m_0\}$ and integer $n_0\geq 1$ not divisible by p, we have for every E in $\boldsymbol{B}(r_0)$*

$$\operatorname{supp}(T_{p^m n_0}(E))\subseteq \boldsymbol{B}(r).$$

The proof of this result is based on the following lemma, giving for each integer $m\geq 0$ a formula for the correspondence τ_m defined in Proposition 4.5. To state this lemma, for each integer $k\geq 0$ put

$$x_k:=\frac{p}{p+1}\cdot p^{-k}\quad\text{and}\quad I_k:=[x_{k+1},x_k],$$

and note that $\bigcup_{k=0}^{\infty} I_k=]0,p/(p+1)]$. Moreover, for all integers $k,k'\geq 0$ denote by

$$A_{k,k'}^{(+1)}:I_k\to I_{k'}\quad\text{and}\quad A_{k,k'}^{(-1)}:I_k\to I_{k'}$$

the unique affine bijection preserving or reversing the orientation, respectively. Note that for every $k\geq 0$ we have $1-A_{k,0}^{(+1)}=A_{k,0}^{(-1)}$ and that for every $k'\geq 1$ we have

$$pA_{k,k'}^{(\pm 1)}=A_{k,k'-1}^{(\pm 1)}. \tag{5-9}$$

Lemma 5.7. *For each integer $m\geq 0$ denote by τ_m the correspondence acting on $[0,p/(p+1)]$ defined in Proposition 4.5. Then for all integers $k,m\geq 0$, we have*

$$\tau_m|_{I_k}=\begin{cases}\sum_{i=0}^{m} p^i(A_{k,2i-(m-k)}^{(+1)})_* & \text{if } m\leq k;\\ \sum_{i=0}^{m-k-1} p^i(A_{k,i}^{((-1)^{m-k-i})})_*+\sum_{i=m-k}^{m} p^i(A_{k,2i-(m-k)}^{(+1)})_* & \text{if } m\geq k+1.\end{cases}$$

Proof. Fix $k\geq 0$. We proceed by induction on m. The case $m=0$ is trivial and the case $m=1$ is a direct consequence of the definition given in Proposition 4.5. Let $m\geq 2$ be given and suppose that the lemma holds with m replaced by $m-1$ and by $m-2$. If $m\leq k$, then by (5-9)

$$\tau_1(\tau_{m-1}|_{I_k})=\sum_{i=0}^{m-1} p^i(A_{k,2i-(m-k)}^{(+1)})_*+\sum_{i=0}^{m-1} p^{i+1}(A_{k,2i-(m-k)+2}^{(+1)})_*=p\tau_{m-2}|_{I_k}+\sum_{i=0}^{m} p^i(A_{k,2i-(m-k)}^{(+1)})_*,$$

which proves the induction step in the case $m\leq k$. In the case $m=k+1$, using $1-A_{k,0}^{(+1)}=A_{k,0}^{(-1)}$ we have

$$\tau_1(\tau_k|_{I_k})=(A_{k,0}^{(-1)})_*+\sum_{i=1}^{k} p^i(A_{k,2i-1}^{(+1)})_*+\sum_{i=0}^{k} p^{i+1}(A_{k,2i+1}^{(+1)})_*=p\tau_{k-1}|_{I_k}+(A_{k,0}^{(-1)})_*+\sum_{i=1}^{k+1} p^i(A_{k,2i-1}^{(+1)})_*.$$

This proves the induction step in the case $m = k+1$. If $m = k+2$, then

$$\tau_1(\tau_{k+1}|_{I_k}) = (A_{k,0}^{(+1)})_* + p(A_{k,1}^{(-1)})_* + \sum_{i=1}^{k+1} p^i (A_{k,2i-2}^{(+1)})_* + \sum_{i=1}^{k+1} p^{i+1} (A_{k,2i}^{(+1)})_*$$

$$= (A_{k,0}^{(+1)})_* + p(A_{k,1}^{(-1)})_* + p\tau_k|_{I_k} + \sum_{i=2}^{k+2} p^i (A_{k,2i-2}^{(+1)})_*.$$

This proves the induction step in the case $m = k+2$. Finally, if $m \geq k+3$, then $\tau_1(\tau_{m-1}|_{I_k})$ is equal to

$$(A_{k,0}^{((-1)^{m-k})})_* + \sum_{j=1}^{m-k-2} p^j (A_{k,j-1}^{((-1)^{m-k-j-1})})_* + \sum_{j=0}^{m-k-2} p^{j+1} (A_{k,j+1}^{((-1)^{m-k-j-1})})_* + \sum_{i=m-k-1}^{m-1} p^i (A_{k,2i-(m-k)}^{(+1)})_* + \sum_{i=m-k-1}^{m-1} p^{i+1} (A_{k,2i-(m-k-2)}^{(+1)})_*$$

$$= \sum_{\ell=0}^{m-k-1} p^\ell (A_{k,\ell}^{((-1)^{m-k-\ell})})_* + p \sum_{s=0}^{m-k-3} p^s (A_{k,s}^{((-1)^{m-k-s-2})})_* + \sum_{i=m-k}^{m} p^i (A_{k,2i-(m-k)}^{(+1)})_* + p \sum_{i=m-k-2}^{m-2} p^i (A_{k,2i-(m-k-2)}^{(+1)})_*$$

$$= p\tau_{m-2}|_{I_k} + \sum_{\ell=0}^{m-k-1} p^\ell (A_{k,\ell}^{((-1)^{m-k-\ell})})_* + \sum_{i=m-k}^{m} p^i (A_{k,2i-(m-k)}^{(+1)})_*.$$

This completes the proof of the induction step and of the lemma. □

Proof of Proposition 5.6. Let $\hat{v}_p$ and $(\tau_m)_{m=0}^{\infty}$ be as in Proposition 4.5.

To prove (i), let r in $]0, 1[$ be given. By Proposition 4.3 there is an integer $\ell \geq 0$ such that $\hat{v}_p(\boldsymbol{B}(r)) \subseteq [x_\ell, x_0]$. Then the desired assertion follows from Proposition 4.5 and by the observation that by Lemma 5.7 for every x in $]0, x_0]$ we have

$$\deg(\tau_m(x)|_{[x_\ell, x_0]}) \leq 1 + p + \cdots + p^\ell.$$

To prove (ii), let r_0 in $]0, 1[$ and an integer $m_0 \geq 0$ be given. By Proposition 4.3 there is an integer $\ell \geq 0$ such that $\hat{v}_p(\boldsymbol{B}(r_0)) \subseteq [x_\ell, x_0]$ and r in $]0, 1[$ such that $\hat{v}_p^{-1}([x_{\ell+m_0}, x_0]) \subseteq \boldsymbol{B}(r)$. Then the desired inclusion follows from Proposition 4.5 by noting that by Lemma 5.7 for every x in $[x_\ell, x_0]$ and every m in $\{0, \ldots, m_0\}$, we have $\operatorname{supp}(\tau_m(x)) \subseteq [x_{\ell+m_0}, x_0]$. □

Appendix A: Lifting the Hasse invariant in characteristic 2 and 3

When p equals 2 or 3 it is not possible to lift the Hasse invariant A_{p-1} to a modular form of level one, holomorphic at infinity, over $\mathbb{Z}_{(p)}$. There are two approaches to solve this issue. On the one hand, there are liftings of A_1^4 and A_2^3 in the desired space (namely, the Eisenstein series E_4 and E_6). On the other hand, considering level structures, liftings can be constructed as algebraic modular forms over $\mathbb{Z}_{(p)}$ of the expected weight but higher level. In this appendix we recall both approaches, following [Katz 1973, Section 2.1], and give a quantitative comparison between them, embodied in Proposition A.1 below. Such comparison is needed in Section 4B.

We start by recalling level structures. Let R be a ring and let $n \geq 1$ be an integer which is assumed to be invertible in R. Let E be an elliptic curve over R in the sense of Section 4A. A *level n structure* on E over R is an isomorphism $\alpha_n \colon E[n] \to (\mathbb{Z}/n\mathbb{Z})^2$ of group schemes over R.

Given an integer $n \geq 1$ and an arbitrary ring R_0 where n is invertible, an algebraic modular form of level $n \geq 1$ over R_0 is a family of maps $F = (F_R)_{R \in R_0\text{-Alg}}$ such that for any $R \in R_0$-Alg, the R-valued map F_R is defined on the set of triples (E, ω, α_n), where E is an elliptic curve over $R \in R_0$-Alg, together with a differential form in $\Omega^1_{E/R}(E)'$ and a level n structure. The element $F_R(E, \omega, \alpha_n) \in R$ must define an assignment satisfying properties analogous to (i), (ii) and (iii) stated in Section 4A. See [Katz 1973, Section 1.2] for further details.

When R_0 contains $1/n$ and a primitive n-th root of unity, the q-expansions of an algebraic modular form F of level n over R_0 are defined as the elements of $\mathbb{Z}((q)) \otimes_{\mathbb{Z}} R_0$ obtained by evaluating F at the triples $(\mathrm{Tate}(q^n), \omega_{\mathrm{can}}, \alpha_n)_{R_0}$ consisting of the Tate curve $\mathrm{Tate}(q^n)$ (see Section 5A) with its canonical differential ω_{can}, regarded over $\mathbb{Z}((q)) \otimes_{\mathbb{Z}} R_0$, with α_n varying over all level n structures of $\mathrm{Tate}(q^n)$ over $\mathbb{Z}((q)) \otimes_{\mathbb{Z}} R_0$. If all of the q-expansions of F lie in $\mathbb{Z}[\![q]\!] \otimes_{\mathbb{Z}} R_0$ then F is called holomorphic at infinity. For algebraic modular forms F of level one there is only one q-expansion, which coincides with the previously defined $F(q)$.

According to [Katz 1973, page 98], for any level $3 \leq n \leq 11$ odd, there exists a lifting of A_1 to a modular form of level n, weight one, holomorphic at infinity, over $\mathbb{Z}[1/n]$. We define E_1 as any such lifting and set $n(\mathrm{E}_1) := n$. Similarly, when $m \geq 4$ and $3 \nmid m$, there exists a lifting of A_2 to a modular form of level m, weight two, holomorphic at infinity, over $\mathbb{Z}[1/m]$. We define E_2 as any such lifting and set $n(\mathrm{E}_2) := m$.

The following statement is a comparison between both approaches.

Proposition A.1. *Let $E \in Y_{\mathrm{sups}}(\mathbb{C}_p)$ and let ω be a differential form in $\Omega^1_{E/\mathcal{O}_p}(E)'$:*

(i) *For any level $n(\mathrm{E}_1)$ structure α on E we have*

$$\mathrm{ord}_2(\mathrm{E}_4(E, \omega)) < 3 \Leftrightarrow \mathrm{ord}_2(\mathrm{E}_1^4(E, \omega, \alpha)) < 3,$$

in which case $\mathrm{ord}_2(\mathrm{E}_4(E, \omega)) = \mathrm{ord}_2(\mathrm{E}_1^4(E, \omega, \alpha))$.

(ii) *For any level $n(\mathrm{E}_2)$ structure α on E we have*

$$\mathrm{ord}_3(\mathrm{E}_6(E, \omega)) < \tfrac{5}{2} \Leftrightarrow \mathrm{ord}_3(\mathrm{E}_2^3(E, \omega, \alpha)) < \tfrac{5}{2},$$

in which case $\mathrm{ord}_3(\mathrm{E}_4(E, \omega)) = \mathrm{ord}_3(\mathrm{E}_2^3(E, \omega, \alpha))$.

Proof. In order to prove (i), we start by recalling the q-expansion

$$\mathrm{E}_4(q) = 1 + 240 \sum_{n=1}^{\infty} \sigma_3(n) q^n,$$

obtained by setting $k = 4$ in (4-2). Since $\mathrm{ord}_2(240) = 4$, we have $\mathrm{E}_4(q) \equiv 1 \bmod 2^4$. Now, put $n_1 := n(\mathrm{E}_1)$, let ζ_{n_1} be a primitive n_1-th roof of unity and define $R_1 := \mathbb{Z}[1/n_1, \zeta_{n_1}]$. By the definition of E_1 we have

$$\mathrm{E}_1(\mathrm{Tate}(q^{n_1}), \omega_{\mathrm{can}}, \alpha_{n_1}) \equiv A_1(q^{n_1}) \equiv 1 \bmod 2R_1,$$

hence

$$\mathrm{E}_1^4(\mathrm{Tate}(q^{n_1}), \omega_{\mathrm{can}}, \alpha_{n_1}) \equiv 1 \equiv \mathrm{E}_4(q^{n_1}) \bmod 2^3 R_1,$$

for any level n_1 structure α_{n_1} on $\mathrm{Tate}(q^{n_1})$. We conclude that the form f obtained by reducing modulo $2^3\mathbb{Z}[1/n_1]$ the form $\mathrm{E}_4 - \mathrm{E}_1^4$ is an algebraic modular form of weight 4, level n_1 over $\mathbb{Z}/2^3\mathbb{Z}$, whose q-expansions over $(\mathbb{Z}/2^3\mathbb{Z})[\zeta_{n_1}]$ vanish identically. By [Katz 1973, Theorem 1.6.1] we deduce that $f = 0$. By compatibility with base change we conclude that for any $\mathbb{Z}[1/n_1]$-algebra R and any triple $(E, \omega, \alpha_{n_1})$ over R we have

$$\mathrm{E}_4(E, \omega) - \mathrm{E}_1^4(E, \omega, \alpha_{n_1}) \equiv f((E, \omega, \alpha_{n_1})_{R/2^3R}) \equiv 0 \bmod 2^3 R.$$

In particular, choosing $R = \mathcal{O}_p$, we get

$$\mathrm{ord}_2(\mathrm{E}_4(E, \omega) - \mathrm{E}_1^4(E, \omega, \alpha_{n_1})) \geq 3, \tag{A-1}$$

for every $E \in Y_{\mathrm{sups}}(\mathbb{C}_p)$, every basis ω of $\Omega^1_{E/\mathcal{O}_p}$ and every level n_1 structure α_{n_1} on E. Then, (i) is a direct consequence of (A-1) and the ultrametric inequality.

The proof of (ii) is unfortunately less straightforward. This is because the same argument used to prove (A-1) only yields the inequality

$$\mathrm{ord}_3(\mathrm{E}_6(E, \omega) - \mathrm{E}_2^3(E, \omega, \alpha_{n_2})) \geq 2,$$

valid for any level $n_2 := n(\mathrm{E}_2)$ structure α_{n_2} on E, but such inequality does not imply the desired result. On the other hand, the above argument allows us to infer

$$\mathrm{ord}_3(\mathrm{E}_4(E, \omega) - \mathrm{E}_2^2(E, \omega, \alpha_{n_2})) \geq 1. \tag{A-2}$$

In order to prove (ii) we introduce the series

$$G_2(\tau) = 1 + 24 \sum_{n=1}^{\infty} \left(\sigma_1(n) - 2\sigma_1\left(\frac{n}{2}\right) \right) e^{2\pi i n \tau}, \qquad \tau \in \mathbb{H}, \tag{A-3}$$

where $\sigma_1\left(\frac{n}{2}\right)$ is defined as zero when n is odd. It is known that G_2 is a classical holomorphic modular form of weight two for the group $\Gamma_0(2) = \left\{ g \in \mathrm{SL}_2(\mathbb{Z}) : g \equiv \left(\begin{smallmatrix} * & * \\ 0 & * \end{smallmatrix}\right) \bmod 2 \right\}$.‖ By [Katz 1973, Corollary 1.9.1], G_2 defines an algebraic modular over $\mathbb{Z}\left[\frac{1}{2}\right]$ of weight two and level two. This form satisfies the identity

$$4G_2^3 = \mathrm{E}_6 + 3\,\mathrm{E}_4\, G_2. \tag{A-4}$$

Indeed, the space of modular forms over $\mathbb{C}$ of weight six for $\Gamma_0(2)$ has dimension 2, see the dimension formulas in [Diamond and Shurman 2005, Chapter 3]. By comparing Fourier expansions, it is easy

‖Up to an explicit multiplicative factor, this is denoted by $G_{2,2}$ in [Diamond and Shurman 2005, Section 1.2].

to check that E_6 and $\mathrm{E}_4\, G_2$ are linearly independent over $\mathbb{C}$, hence they form a basis of such space. This implies that there exist $a, b \in \mathbb{C}$ with $G_2^3 = a\,\mathrm{E}_6 + b\,\mathrm{E}_4\, G_2$. Then, (A-4) follows at the level of classical modular forms by computing the values of a and b, which can be done by comparing Fourier expansions. Finally, the fact that (A-4) holds as an identity between algebraic modular forms over $\mathbb{Z}[\frac{1}{2}]$ is a consequence of [Katz 1973, Corollary 1.9.1].

We also recall the identity

$$\mathrm{E}_6^2 - \mathrm{E}_4^3 = 1728\Delta.$$

At the level of classical modular forms, see for example [Diamond and Shurman 2005, Sections 1.1 and 1.2]. Then, this identity holds at the level of algebraic modular forms by the same reasoning as before. Given $E \in Y_{\mathrm{sups}}(\mathbb{C}_p)$ and a differential form ω in $\Omega^1_{E/\mathcal{O}_p}(E)'$, we have $\Delta(E, \omega) \in \mathcal{O}_p^\times$ since E has good reduction. This implies

$$\mathrm{ord}_3(\mathrm{E}_6^2(E, \omega) - \mathrm{E}_4^3(E, \omega)) = 3. \tag{A-5}$$

By using (A-4) and (A-5), we will now prove (ii). Let α be a level n_2 structure on E. First, assume that $\mathrm{ord}_3(\mathrm{E}_2(E, \omega, \alpha)) < \frac{5}{6}$. From (A-4) we see that the reduction modulo 3 of G_2 equals A_2. Since the same holds for E_2, we conclude that

$$\mathrm{ord}_3(\mathrm{E}_2(E, \omega, \alpha) - G_2(E, \omega, \beta)) \geq 1, \tag{A-6}$$

for any level two structure β. In particular

$$\mathrm{ord}_3(G_2(E, \omega, \beta)) = \mathrm{ord}_3(\mathrm{E}_2(E, \omega, \alpha)) < \tfrac{5}{6}.$$

By (A-4) we have

$$\mathrm{E}_6(E, \omega) = G_2(E, \omega, \beta)(4G_2^2(E, \omega, \beta) - 3\,\mathrm{E}_4(E, \omega)).$$

But by (A-2) and (A-6) we also have

$$\mathrm{ord}_3(3\,\mathrm{E}_4(E, \omega)) = 1 + \mathrm{ord}_3(\mathrm{E}_4(E, \omega)) \geq 1 + \min\{1, \mathrm{ord}_3(G_2^2(E, \omega, \beta))\} > \mathrm{ord}_3(G_2^2(E, \omega, \beta)),$$

hence

$$\mathrm{ord}_3(\mathrm{E}_6(E, \omega)) = \mathrm{ord}_3(G_2^3(E, \omega, \beta)) = \mathrm{ord}_3(\mathrm{E}_2^3(E, \omega, \alpha)).$$

This proves one implication. Let us now prove the reciprocal. We start by assuming that $\mathrm{ord}_3(\mathrm{E}_6(E, \omega)) < \frac{5}{2}$. If $\mathrm{ord}_3(\mathrm{E}_4(E, \omega)) < 1$, then we can use (A-2), (A-5) and (A-6) to deduce that $\mathrm{ord}_3(\mathrm{E}_4^3(E, \omega)) = \mathrm{ord}_3(\mathrm{E}_6^2(E, \omega))$ and $\mathrm{ord}_3(G_2^2(E, \omega, \beta)) = \mathrm{ord}_3(\mathrm{E}_4(E, \omega))$. This implies

$$\mathrm{ord}_3(3G_2(E, \omega, \beta)\,\mathrm{E}_4(E, \omega)) = 1 + \mathrm{ord}_3(\mathrm{E}_6(E, \omega)) > \mathrm{ord}_3(\mathrm{E}_6(E, \omega)).$$

By (A-4) and (A-6) we conclude

$$\mathrm{ord}_3(\mathrm{E}_2^3(E, \omega, \alpha)) = \mathrm{ord}_3(G_2^3(E, \omega, \beta)) = \mathrm{ord}_3(\mathrm{E}_6(E, \omega)).$$

Now, if $\operatorname{ord}_3(\mathrm{E}_4(E,\omega)) \geq 1$ then (A-2) and (A-6) imply $\operatorname{ord}_3(G_2^2(E,\omega,\beta)) \geq 1$, giving

$$\operatorname{ord}_3(3G_2(E,\omega,\beta)\,\mathrm{E}_4(E,\omega)) \geq \tfrac{5}{2} > \operatorname{ord}_3(\mathrm{E}_6(E,\omega)).$$

As before, we conclude $\operatorname{ord}_3(\mathrm{E}_2^3(E,\omega,\alpha)) = \operatorname{ord}_3(\mathrm{E}_6(E,\omega))$. This proves the reciprocal implication and completes the proof of the proposition. □

Appendix B: Eichler–Shimura analytic relation

In this appendix we further study the canonical branch $\boldsymbol{t}$ of T_p that is defined on $Y_{\mathrm{ord}}(\mathbb{C}_p)$ in Section 3A. We start extending $\boldsymbol{t}$, as follows. Recall that v_p denotes Katz' valuation, defined in Section 4A. Extend v_p to $Y(\mathbb{C}_p)$ as $v_p \equiv 0$ outside $Y_{\mathrm{sups}}(\mathbb{C}_p)$, and put

$$N_p := \left\{ E \in Y(\mathbb{C}_p) : v_p(E) < \frac{p}{p+1} \right\}. \tag{B-1}$$

On $N_p \cap Y_{\mathrm{sups}}(\mathbb{C}_p)$, we use the definition of $\boldsymbol{t}$ in Lemma 4.6. To define $\boldsymbol{t}$ at a point E in $Y_{\mathrm{bad}}(\mathbb{C}_p)$, let z in $\boldsymbol{D}(0,1)^*$ and let $\varphi_z \colon \mathbb{C}_p^\times / z^{\mathbb{Z}} \to \mathrm{Tate}(z)(\mathbb{C}_p)$ be the isomorphism of analytic groups as in Section 5A. Then we define

$$H(E) := \varphi_z(\{\zeta z^n \in \mathbb{C}_p^\times : \zeta^p = 1, n \in \mathbb{Z}\}/z^{\mathbb{Z}}), \quad \text{and} \quad \boldsymbol{t}(E) := E/H(E).$$

Note that in the notation (5-2) of Section 5A, we have $H(E) = C_{p,z^p}$. The map $\boldsymbol{t} \colon N_p \to Y(\mathbb{C}_p)$ so defined is the *canonical branch of* T_p.

The goal of this appendix is to prove the following result.

Theorem B.1 (Eichler–Shimura analytic relation). *The canonical branch* $\boldsymbol{t}$ *of* T_p *is given by a finite sum of Laurent series, each of which converges on all of* N_p. *Furthermore, for every* E *in* $N_p \setminus Y_{\mathrm{bad}}(\mathbb{C}_p)$ *we have*

$$\operatorname{ord}_p(\boldsymbol{t}(j(E)) - j(E)^p) \geq 1 - v_p(E), \tag{B-2}$$

and for every E *in* $Y(\mathbb{C}_p)$ *we have*

$$T_p(E) = \begin{cases} \boldsymbol{t}^*(E) + [\boldsymbol{t}(E)] & \text{if } v_p(E) \leq 1/(p+1); \\ \boldsymbol{t}^*(E) & \text{if } v_p(E) > 1/(p+1). \end{cases} \tag{B-3}$$

In view of (B-2), the relation (B-3) can be seen as refinement and a lift to N_p of the classical Eichler–Shimura congruence relation; see for example [Shimura 1971, Section 7.4] or [Diamond and Shurman 2005, Section 8.7].

The proof of Theorem B.1 is at the end of this appendix. When restricted to $Y_{\mathrm{ord}}(\mathbb{C}_p)$, it is a direct consequence of Theorem 3.3 and Proposition 3.4 with $m = 1$. To prove (B-3) for E in $Y_{\mathrm{sups}}(\mathbb{C}_p)$, we use Lemma 4.6. To prove this relation on $Y_{\mathrm{bad}}(\mathbb{C}_p)$, we use the results on the uniformization of p-adic elliptic curves with multiplicative reduction, recalled in Section 5A. To prove (B-2) and that $\boldsymbol{t}$ is a finite sum of Laurent series for $p \geq 5$, we use Theorem 3.3 in Section 3A. For $p = 2$ and 3, we use Proposition B.2

below, whose proof is based on the explicit formulae in [Mestre 1986, Appendice]. This result also provides a proof of Theorem 3.3 when $p = 2$ and 3

Note that for $p = 2$ and 3, the set $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$ consists of a single point whose j-invariant is equal to 0 and to 1728; see for example [Silverman 2009, Chapter V, Section 4].

Proposition B.2. *Put* $\mathrm{j}_2 := 0$ *and* $\mathrm{j}_3 := 1728$, *and consider the polynomials*

$$\check{k}_2(z) := -93 \cdot 2^4 z + 627 \cdot 2^8 \quad \text{and} \quad \check{k}_3(z) := 328 \cdot 3^2 z^2 + 85708 \cdot 3^3 z + 1263704 \cdot 3^5.$$

Then for $p = 2$ *and* 3, *the canonical branch* $\boldsymbol{t}$ *of* T_p *admits a Laurent series expansion of the form*

$$\boldsymbol{t}(z) = (z - \mathrm{j}_p)^p + \mathrm{j}_p + \check{k}_p(z - \mathrm{j}_p) + \sum_{n=1}^{\infty} \frac{A_n^{(p)}}{(z - \mathrm{j}_p)^n},$$

where for every $n \geq 1$ *the coefficient* $A_n^{(p)}$ *is in* $\mathbb{Z}$ *and satisfies*

$$\mathrm{ord}_p(A_n^{(p)}) \geq \begin{cases} 4 + 8n & \text{if } p = 2; \\ \frac{3}{2} + \frac{9}{2} n & \text{if } p = 3, \end{cases}$$

with equality if $n = 1$.

To prove this proposition, we introduce some notation and recall the explicit formulae in [Mestre 1986, Appendice]. For $\mathbb{K} = \mathbb{C}$ or $\mathbb{C}_p$, we use j to identify $Y(\mathbb{K})$ with $\mathbb{K}$ and consider T_p as a correspondence acting on $\mathrm{Div}(\mathbb{K})$. Let $Y_0(p)$, α_p and β_p be as in Section 2B, so that $T_p = (j \circ \alpha_p)_* \circ (j \circ \beta_p)^*$. Denote by

$$w_p \colon Y_0(p)(\mathbb{K}) \to Y_0(p)(\mathbb{K})$$

the *Atkin–Lehner* or *Fricke involution*, defined by $w_p(E, C) := (E/C, E[p]/C)$ and note that $\beta_p = \alpha_p \circ w_p$. Identify $Y_0(p)(\mathbb{C})$ with the quotient $\Gamma_0(p) \backslash \mathbb{H}$ and denote by $\eta \colon \mathbb{H} \to \mathbb{C}$ *Dedekind's eta function*, defined by

$$\eta(\tau) := \exp\left(\frac{\pi i \tau}{12}\right) \prod_{n=1}^{\infty} (1 - \exp(2\pi i n \tau)).$$

Then for $p = 2$ or 3, the function $\hat{x}_p \colon \mathbb{H} \to \mathbb{C}$ defined by

$$\hat{x}_p(\tau) := \left(\frac{\eta(\tau)}{\eta(p\tau)}\right)^{24/(p-1)}$$

descends to a complex analytic isomorphism $x_p \colon Y_0(p)(\mathbb{C}) \to \mathbb{C}$. Moreover, defining

$$\hat{\alpha}_p(z) := \begin{cases} (z + 2^4)^3/z & \text{if } p = 2; \\ (z + 3^3)(z + 3)^3/z & \text{if } p = 3, \end{cases} \quad \text{and} \quad \hat{w}_p(z) := \begin{cases} 2^{12}/z & \text{if } p = 2; \\ 3^6/z & \text{if } p = 3, \end{cases}$$

we have $j \circ \alpha_p = \hat{\alpha}_p \circ x_p$ and $x_p \circ w_p = \hat{w}_p \circ x_p$; see [Mestre 1986, pages 238 and 239]. It follows that, if we put

$$\hat{\beta}_p(z) := \hat{\alpha}_p \circ \hat{w}_p(z) = \begin{cases} (z + 2^8)^3/z^2 & \text{if } p = 2; \\ (z + 3^3)(z + 3^5)^3/z^3 & \text{if } p = 3, \end{cases}$$

then $j \circ \beta_p = \hat{\beta}_p \circ x_p$ and therefore $T_p = (\hat{\alpha}_p)_* \circ \hat{\beta}_p^*$ as algebraic correspondences over $\mathbb{C}$. Since T_p, $\hat{\alpha}_p$ and $\hat{\beta}_p$ are all defined over $\mathbb{Q}$, we have that the equality $T_p = (\hat{\alpha}_p)_* \circ \hat{\beta}_p^*$ also holds as algebraic correspondences over $\mathrm{Div}(Y(\mathbb{C}_p))$.

The following elementary lemma is used the proof of Proposition B.2. Given r in $]0, 1[$, and a Laurent series $\sum_{n=0}^{\infty} \frac{A_n}{z^n}$ in $\mathbb{Z}[\![\frac{1}{z}]\!]$, put

$$\left\| \sum_{n=0}^{\infty} \frac{A_n}{z^n} \right\|_r := \sup\{|A_n|_p r^{-n} : n \geq 0\}.$$

Lemma B.3. *Let $\delta(z)$ in $\frac{1}{z}\mathbb{Z}[\![\frac{1}{z}]\!]$ be given and put $f(z) := z(1+\delta(z))$. Then there is $\Delta(z)$ in $\frac{1}{z}\mathbb{Z}[\![\frac{1}{z}]\!]$ such that $F(z) := z(1+\Delta(z))$ satisfies $F(f(z)) = z$. If in addition for some r in $]0, 1[$ we have $\|\delta\|_r \leq 1$, then $\|\Delta\|_r \leq 1$.*

Proof. We start defining recursively a sequence $(\Delta_n)_{n=0}^{\infty}$ in $\frac{1}{z}\mathbb{Z}\left[\frac{1}{z}\right]$ such that for every integer $n \geq 0$,

$$z^n \Delta_n(z) \in \mathbb{Z}[z], \quad \Delta_{n+1}(z) \equiv \Delta_n(z) \bmod \tfrac{1}{z^{n+1}}\mathbb{Z}\left[\tfrac{1}{z}\right],$$

and the Laurent polynomial $F_n(z) := z(1+\Delta_n(z))$ satisfies

$$F_n(f(z)) \equiv z \bmod \tfrac{1}{z^n}\mathbb{Z}\left[\tfrac{1}{z}\right].$$

For $n = 0$ put $\Delta_0(z) = 0$, so $F_0(f(z)) = f(z) \equiv z \bmod \mathbb{Z}\left[\frac{1}{z}\right]$. Let $n \geq 0$ be an integer so that Δ_n is already defined and let A in $\mathbb{Z}$ be the coefficient of $1/z^n$ in $F_n(f(z))$. Then for $\Delta_{n+1}(z) := \Delta_n(z) - A/z^{n+1}$, we have

$$(F_{n+1} - F_n)(f(z)) = -\frac{A}{z^n(1+\delta(z))^n} = -\frac{A}{z^n}\left(1 + \sum_{k=1}^{\infty}(-\delta(z))^k\right)^n \equiv -\frac{A}{z^n} \bmod \frac{1}{z^{n+1}}\mathbb{Z}\left[\frac{1}{z}\right],$$

and therefore

$$F_{n+1}(f(z)) - z = F_n(f(z)) - z + (F_{n+1} - F_n)(f(z)) \equiv 0 \bmod \tfrac{1}{z^{n+1}}\mathbb{Z}\left[\tfrac{1}{z}\right].$$

This completes the definition of the sequence $(\Delta_n)_{n=0}^{\infty}$. It follows that the unique series Δ in $\frac{1}{z}\mathbb{Z}[\![\frac{1}{z}]\!]$ satisfying for every $n \geq 0$ the congruence

$$\Delta(z) \equiv \Delta_n(z) \bmod \tfrac{1}{z^{n+1}}\mathbb{Z}[\![\tfrac{1}{z}]\!],$$

satisfies $F(f(z)) = z$.

To prove the last assertion, note that for every r in $]0, 1[$,

$$I_r := \left\{ z(1+g(z)) : g(z) \in \tfrac{1}{z}\mathbb{Z}[\![\tfrac{1}{z}]\!], \|g\|_r \leq 1 \right\}$$

is a collection of series in $\mathbb{Z}[\![\frac{1}{z}]\!]$ that is closed under composition. It follows from the above construction that, if for some r in $]0, 1[$ we have $\|\delta\|_r \leq 1$, then for every integer $n \geq 0$ the series F_n and $F_n \circ f$ are both in I_r. This implies that F is in I_r, as wanted. □

The proof of Proposition B.2 is given after the following lemma, which is also used in the proof of Theorem B.1.

Lemma B.4. *For an arbitrary prime number p, the right-hand side of (3-3) converges to $\boldsymbol{t}$ on $Y_{\mathrm{ord}}(\mathbb{C}_p) \cup Y_{\mathrm{bad}}(\mathbb{C}_p)$.*

Proof. Let $\Phi_p(X, Y)$ be the modular polynomial of level p, as defined in Section 2B, so that for every z in $Y_{\mathrm{ord}}(\mathbb{C}_p)$ we have $\Phi_p(z, \boldsymbol{t}(z)) = 0$. By Theorem 3.3, the finite sum of Laurent series on the right-hand side of (3-3) converges on $Y_{\mathrm{ord}}(\mathbb{C}_p) \cup Y_{\mathrm{bad}}(\mathbb{C}_p)$ to a function $\hat{\boldsymbol{t}}$ extending $\boldsymbol{t}$, and for z in $Y_{\mathrm{bad}}(\mathbb{C}_p)$ we have $|\hat{\boldsymbol{t}}(z)|_p = |z|_p^p$. It follows that for every z in $Y_{\mathrm{bad}}(\mathbb{C}_p)$ we have $\Phi_p(z, \hat{\boldsymbol{t}}(z)) = 0$, so $\hat{\boldsymbol{t}}(z)$ is in the support of $T_p(z)$. Combining (5-1) and (5-3), we conclude that $\hat{\boldsymbol{t}}(z) = \boldsymbol{t}(z)$. □

Proof of Proposition B.2. Note that if we put $r_2 := 2^{-8}$ and $r_3 := 3^{-9/2}$, then for $p = 2$ and 3 we have by Proposition 4.3,

$$N_p = \{z \in \mathbb{C}_p : |z - \mathrm{j}_p|_p > r_p\}.$$

For $p = 2$ and 3, put

$$\check{\alpha}_p := \hat{\alpha}_p - \mathrm{j}_p \quad \text{and} \quad \check{\beta}_p := \hat{\beta}_p - \mathrm{j}_p.$$

Note that for $p = 3$, we have

$$\check{\alpha}_3(z) = \frac{(z^2 + 2 \cdot 3^2 z - 3^3)^2}{z} \quad \text{and} \quad \check{\beta}_3(z) = \frac{(z^2 - 2 \cdot 3^5 z - 3^9)^2}{z^3}.$$

So, for $p = 2$ and 3 the rational map $\delta_p(z) := z^{-1}\check{\beta}_p(z) - 1$ is a Laurent polynomial in $\frac{1}{z}\mathbb{Z}\left[\frac{1}{z}\right]$ satisfying $\|\delta_p\|_{r_p} \le 1$. In particular, for every z in the set

$$\check{N}_p := \{z' \in \mathbb{C}_p : |z'|_p > r_p\},$$

we have $|\check{\beta}_p(z)|_p = |z|_p$, so $\check{\beta}_p$ maps $\check{N}_p$ into itself. By Lemma B.3 there is $\Delta_p(w)$ in $\frac{1}{w}\mathbb{Z}[\![\frac{1}{w}]\!]$ such that $\|\Delta_p\|_{r_p} \le 1$ and such that the map

$$\begin{aligned} F_p \colon \check{N}_p &\to \check{N}_p \\ w &\mapsto F_p(w) := w(1 + \Delta_p(w)) \end{aligned}$$

is an inverse of $\check{\beta}_p|_{\check{N}_p}$.

We show below that $\boldsymbol{t}$ coincides with the map

$$\begin{aligned} \check{\boldsymbol{t}} \colon N_p &\to \mathbb{C}_p \\ z &\mapsto \check{\boldsymbol{t}}(z) := (\check{\alpha}_p \circ F_p)(z - \mathrm{j}_p) + \mathrm{j}_p. \end{aligned}$$

Once this is established, the proposition follows from explicit computations using the estimates,

$$\|\Delta_p\|_{r_p} \le 1, \quad \left\| \frac{\check{\alpha}_2(w)}{w^2} \right\|_{2^{-4}} \le 1 \text{ for } p = 2, \quad \text{and} \quad \left\| \frac{\check{\alpha}_3(w)}{w^3} \right\|_{3^{-3/2}} \le 1 \text{ for } p = 3.$$

By definition, for each z in $\hat{N}_p$ the point $\check{\boldsymbol{t}}(z)$ is in the support of $T_p(z) = (\alpha_p)_* \circ \beta_p^*(z)$. Moreover, for every z in $Y_{\mathrm{bad}}(\mathbb{C}_p)$ we have $|\check{\boldsymbol{t}}(z)|_p = |z|_p^p$, so by (5-1) and (5-3) we have $\check{\boldsymbol{t}}(z) = \boldsymbol{t}(z)$. Combined with Lemma B.4, this implies that $\check{\boldsymbol{t}}$ and $\boldsymbol{t}$ agree on $Y_{\mathrm{ord}}(\mathbb{C}_p) \cup Y_{\mathrm{bad}}(\mathbb{C}_p)$. In view of Proposition 4.3 and Lemma 4.6, to prove that $\check{\boldsymbol{t}}$ and $\boldsymbol{t}$ agree on $N_p \cap Y_{\mathrm{sups}}(\mathbb{C}_p)$ it is sufficient to show that for every w in $\check{N}_p \cap \mathcal{M}_p$ we have $|(\check{\alpha}_p \circ F_p)(w)|_p \neq |w|_p^{1/p}$. Note that for every w in $\check{N}_p$ we have $|F_p(w)|_p = |w|_p$. A direct computation shows that for $p = 2$ we have

$$|(\check{\alpha}_2 \circ F_2)(w)|_2 \begin{cases} = |w|_2^2 & \text{if } 2^{-4} < |w|_2 < 1; \\ \leq 2^{-8} & \text{if } |w|_2 = 2^{-4}; \\ = 2^{-12}/|w|_2 & \text{if } r_2 < |w|_2 < 2^{-4}, \end{cases}$$

and that for $p = 3$ we have

$$|(\check{\alpha}_3 \circ F_3)(w)|_3 \begin{cases} = |w|_3^3 & \text{if } 3^{-3/2} < |w|_3 < 1; \\ \leq 3^{-9/2} & \text{if } |w|_3 = 3^{-3/2}; \\ = 3^{-6}/|w|_3 & \text{if } r_3 < |w|_3 < 3^{-3/2}. \end{cases}$$

In all the cases we have $|(\check{\alpha}_p \circ F_p)(w)|_p \neq |w|_p^{1/p}$. This completes the proof of $\boldsymbol{t} = \check{\boldsymbol{t}}$, and of the proposition. □

Proof of Theorem B.1. We first prove (B-2), and the assertions about the Laurent series expansion. For $p = 2$ and 3, these are given by Proposition B.2. Assume $p \geq 5$. For each e in $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$, let j_e be given by Proposition 4.3, and define $P_{\mathrm{sups}}(X) = \prod_{e \in Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)} (X - \mathrm{j}_e)$ as in the proof of this proposition. Since the reduction modulo p of the polynomial P_{sups} is separable, for every e in $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$ we have that j_e is in $\mathbb{Q}_p^{\mathrm{unr}}$. Put $\beta_e := \mathrm{j}_e$. Denote by $\hat{\boldsymbol{t}}$ the finite sum of Laurent series in the right-hand side of (3-3) for these choices of $(\beta_e)_{e \in Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)}$. It follows from Theorem 3.3 and Proposition 4.3 that $\hat{\boldsymbol{t}}$ converges on N_p, and by Lemma B.4 that for every z in $Y_{\mathrm{bad}}(\mathbb{C}_p) \cup Y_{\mathrm{ord}}(\mathbb{C}_p)$ we have $\hat{\boldsymbol{t}}(z) = \boldsymbol{t}(z)$. We proceed to prove that for every z in $\hat{N}_p := N_p \cap Y_{\mathrm{sups}}(\mathbb{C}_p)$ we also have $\hat{\boldsymbol{t}}(z) = \boldsymbol{t}(z)$.

Denote by $\Phi_p(X, Y)$ the modular polynomial of level p defined in Section 2B. Note that for every z in $Y_{\mathrm{bad}}(\mathbb{C}_p) \cup Y_{\mathrm{ord}}(\mathbb{C}_p)$ we have

$$\Phi_p(\hat{\boldsymbol{t}}(z), z) = \Phi_p(z, \hat{\boldsymbol{t}}(z)) = 0. \tag{B-4}$$

Since $\hat{\boldsymbol{t}}$ is analytic, (B-4) holds for every z in N_p. In view of Lemma 4.6, this implies that for every E in $\hat{N}_p$ we have either $v_p(\hat{\boldsymbol{t}}(E)) = \frac{1}{p} v_p(E)$, or

$$v_p(\hat{\boldsymbol{t}}(E)) \begin{cases} = p v_p(E) & \text{if } v_p(E) \in \,]0, 1/(p+1)]; \\ \geq p v_p(E) & \text{if } v_p(E) = 1/p + 1; \\ = 1 - v_p(E) & \text{if } v_p(E) \in \,]1/(p+1), p/(p+1)[. \end{cases} \tag{B-5}$$

We now prove that (B-5) holds for every E in $\hat{N}_p$. Fix e in $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$, and note that the function

$$\begin{aligned} \nu \colon \,]0, p/(p+1)[\cap \mathbb{Q} &\to \mathbb{Q} \\ r &\mapsto \nu(r) := \inf\{v_p(\hat{\boldsymbol{t}}(E)) : E \in \boldsymbol{D}(j(e)),\, v_p(E) = r\}, \end{aligned}$$

extends continuously to $]0, p/(p+1)[$. Thus, either (B-5) holds for every E in $N_p \cap \boldsymbol{D}(j(e))$, or for every E in this set we have $v_p(\hat{\boldsymbol{t}}(E)) = \frac{1}{p} v_p(E)$. So, to prove that (B-5) holds for every E in $N_p \cap \boldsymbol{D}(j(e))$ it is sufficient to prove that it holds for some E_0 in $N_p \cap \boldsymbol{D}(j(e))$. Choose E_0 in $N_p \cap \boldsymbol{D}(j(e))$ such that $z_0 := j(E_0)$ satisfies

$$0 < \operatorname{ord}_p(z_0 - \mathrm{j}_e) < \frac{1}{p+1}.$$

By Theorem 3.3 we have

$$\operatorname{ord}_p(\hat{\boldsymbol{t}}(z_0) - z_0^p - pk(z_0)) \geq 1 - \operatorname{ord}_p(z_0 - \mathrm{j}_e) > \frac{p}{p+1}.$$

Since $\operatorname{ord}_p(z_0 - \mathrm{j}_e) < \frac{1}{p}$, we also have

$$\operatorname{ord}_p(\hat{\boldsymbol{t}}(z_0) - \mathrm{j}_e^p) = p \operatorname{ord}_p(z_0 - \mathrm{j}_e) < \frac{p}{p+1}.$$

Combined with $\operatorname{ord}_p(\mathrm{j}_e^p - \mathrm{j}_{e^{(p)}}) \geq 1$ and $\operatorname{ord}_p(pk(z_0)) \geq 1$, this implies

$$\operatorname{ord}_p(\hat{\boldsymbol{t}}(z_0) - \mathrm{j}_{e^{(p)}}) = p \operatorname{ord}_p(z_0 - \mathrm{j}_e), \tag{B-6}$$

and therefore (B-5) with $E = E_0$. This completes the proof that (B-5) holds for every E in $\hat{N}_p$. In view of (B-4), Proposition 4.3, and Lemma 4.6, it follows that for every z in $\hat{N}_p$ we have $\hat{\boldsymbol{t}}(z) = \boldsymbol{t}(z)$. By Theorem 3.3 we also obtain (B-2).

It remains to prove (B-3) for an arbitrary prime number p. Note that for E in $Y_{\mathrm{ord}}(\mathbb{C}_p)$ this is given by Proposition 3.4 with $m = 1$, and that for E in $Y_{\mathrm{bad}}(\mathbb{C}_p)$ this follows from the combination of (5-1), and of (5-3) with $n = p$. It remains to prove (B-3) for E in $\hat{N}_p$. By the considerations above, and the proof of Proposition B.2, we have that (B-5) holds for every prime number p and for every E in $\hat{N}_p$. By Lemma 4.6 we deduce that:

(1) $\boldsymbol{t}$ maps
$$N'_p := \left\{ E \in Y(\mathbb{C}_p) : 0 < v_p(E) < \frac{1}{p+1} \right\}$$
onto $\hat{N}_p$, and for every E in $\hat{N}_p$ the divisor $(\boldsymbol{t}|_{N'_p})^*(E)$ has degree p.

(2) $\boldsymbol{t}$ maps
$$S_p := \left\{ E \in Y(\mathbb{C}_p) : v_p(E) = \frac{1}{p+1} \right\}$$
onto $B_p := Y_{\mathrm{sups}}(\mathbb{C}_p) \setminus \hat{N}_p$, and for every E in B_p the divisor $(\boldsymbol{t}|_{S_p})^*(E)$ has degree $p+1$.

(3) $\boldsymbol{t}$ maps $A_p := \hat{N}_p \setminus (N'_p \cup S_p)$ onto itself, and for every E in A_p we have $(\boldsymbol{t}|_{A_p})^*(E) = [\boldsymbol{t}(E)]$.

The proof of (B-3) is divided in the following cases:

(1) For E in B_p, we have $\boldsymbol{t}^*(E) = (\boldsymbol{t}|_{S_p})^*(E)$ and this divisor has degree $p+1$. Together with (B-4) this implies $T_p(E) = \boldsymbol{t}^*(E)$.

(2) For E in A_p, we have $\boldsymbol{t}^*(E) = (\boldsymbol{t}|_{N'_p})^*(E) + (\boldsymbol{t}|_{A_p})^*(E)$ and this divisor has degree $p+1$. As in the previous case we conclude that $T_p(E) = \boldsymbol{t}^*(E)$.

(3) For E in $N'_p \cup S_p$, we have $\boldsymbol{t}^*(E) = (\boldsymbol{t}|_{N'_p})^*(E)$ and this divisor is of degree p. Combined with (B-4) this implies that the divisor $T_p(E) - \boldsymbol{t}^*(E)$ has degree 1. On the other hand, by (B-5) the point $\boldsymbol{t}(E)$ is not in the support of $\boldsymbol{t}^*(E)$, so by (B-4) we have $T(E) - \boldsymbol{t}^*(E) = [\boldsymbol{t}(E)]$.

This completes the proof of (B-3), and of the theorem. □

Acknowledgments

Menares and Rivera-Letelier thank Leon Takhtajan for references. Menares thanks Emmanuel Ullmo for sharing his interest in the questions we study here. He also acknowledges Rodolphe Richard for explaining him the basic ideas leading to Propositions 5.1 and 5.4. We thank the anonymous referees for their valuable comments that helped us improve the exposition.

During the preparation of this work the Herrero was supported by the Chilean CONICYT grant 21130412 and the Royal Swedish Academy of Sciences. Menares acknowledges partial support from FONDECYT grant 1171329. Rivera-Letelier acknowledges partial support from NSF grant DMS-1700291. The authors would like to thank the Pontificia Universidad Católica de Valparaíso, the University of Rochester and Universitat de Barcelona for hospitality during the preparation of this work.

References

[Apostol 1976] T. M. Apostol, *Introduction to analytic number theory*, Springer, 1976. MR Zbl

[Baker and Rumely 2010] M. Baker and R. Rumely, *Potential theory and dynamics on the Berkovich projective line*, Math. Surv. Monogr. **159**, Amer. Math. Soc., Providence, RI, 2010. MR Zbl

[Berkovich 1990] V. G. Berkovich, *Spectral theory and analytic geometry over non-Archimedean fields*, Math. Surv. Monogr. **33**, Amer. Math. Soc., Providence, RI, 1990. MR Zbl

[Billingsley 1968] P. Billingsley, *Convergence of probability measures*, Wiley, New York, 1968. MR Zbl

[Bilu 1997] Y. Bilu, "Limit distribution of small points on algebraic tori", *Duke Math. J.* **89**:3 (1997), 465–476. MR Zbl

[Brink 2006] D. Brink, "New light on Hensel's lemma", *Expo. Math.* **24**:4 (2006), 291–306. MR Zbl

[Buzzard 2003] K. Buzzard, "Analytic continuation of overconvergent eigenforms", *J. Amer. Math. Soc.* **16**:1 (2003), 29–55. MR Zbl

[Clozel and Ullmo 2004] L. Clozel and E. Ullmo, "Équidistribution des points de Hecke", pp. 193–254 in *Contributions to automorphic forms, geometry, and number theory* (Baltimore, MD, 2002), edited by H. Hida et al., Johns Hopkins Univ. Press, Baltimore, MD, 2004. MR Zbl

[Clozel et al. 2001] L. Clozel, H. Oh, and E. Ullmo, "Hecke operators and equidistribution of Hecke points", *Invent. Math.* **144**:2 (2001), 327–351. MR Zbl

[Cohen 2007] H. Cohen, *Number theory, II: Analytic and modern tools*, Graduate Texts in Math. **240**, Springer, 2007. MR Zbl

[Coleman and McMurdy 2006] R. Coleman and K. McMurdy, "Fake CM and the stable model of $X_0(Np^3)$", *Doc. Math.* extra volume (2006), 261–300. MR Zbl

[Cox 2013] D. A. Cox, *Primes of the form x^2+ny^2: Fermat, class field theory, and complex multiplication*, 2nd ed., Wiley, Hoboken, NJ, 2013. MR Zbl

[Deuring 1941] M. Deuring, "Die Typen der Multiplikatorenringe elliptischer Funktionenkörper", *Abh. Math. Sem. Hansischen Univ.* **14** (1941), 197–272. MR Zbl

[Diamond and Im 1995] F. Diamond and J. Im, "Modular forms and modular curves", pp. 39–133 in *Seminar on Fermat's last theorem* (Toronto, 1993-1994), edited by V. K. Murty, CMS Conf. Proc. **17**, Amer. Math. Soc., Providence, RI, 1995. MR Zbl

[Diamond and Shurman 2005] F. Diamond and J. Shurman, *A first course in modular forms*, Graduate Texts in Math. **228**, Springer, 2005. MR Zbl

[Dinh et al. 2020] T.-C. Dinh, L. Kaufmann, and H. Wu, "Dynamics of holomorphic correspondences on Riemann surfaces", *Int. J. Math.* **31**:5 (2020), 2050036, 21. MR Zbl

[Duke 1988] W. Duke, "Hyperbolic distribution problems and half-integral weight Maass forms", *Invent. Math.* **92**:1 (1988), 73–90. MR Zbl

[Dwork 1969] B. Dwork, "p-adic cycles", *Inst. Hautes Études Sci. Publ. Math.* **37** (1969), 27–115. MR Zbl

[Eskin and Oh 2006] A. Eskin and H. Oh, "Ergodic theoretic proof of equidistribution of Hecke points", *Ergodic Theory Dynam. Systems* **26**:1 (2006), 163–167. MR Zbl

[Fresnel and van der Put 2004] J. Fresnel and M. van der Put, *Rigid analytic geometry and its applications*, Progr. Math. **218**, Birkhäuser, Boston, 2004. MR Zbl

[Goren and Kassaei 2017] E. Z. Goren and P. L. Kassaei, "p-adic dynamics of Hecke operators on modular curves", preprint, 2017. arXiv

[Gross 1986] B. H. Gross, "On canonical and quasicanonical liftings", *Invent. Math.* **84**:2 (1986), 321–326. MR

[Hazewinkel 1978] M. Hazewinkel, *Formal groups and applications*, Pure Appl. Math. **78**, Academic Press, New York, 1978. MR Zbl

[Herrero et al. 2019] S. Herrero, R. Menares, and J. Rivera-Letelier, "p-adic distribution of CM points and Hecke orbits. II: Linnik equidistribution on the supersingular locus", preprint, 2019.

[Iwaniec 1987] H. Iwaniec, "Fourier coefficients of modular forms of half-integral weight", *Invent. Math.* **87**:2 (1987), 385–401. MR Zbl

[de Jong and Noot 1991] J. de Jong and R. Noot, "Jacobians with complex multiplication", pp. 177–192 in *Arithmetic algebraic geometry* (Texel, Netherlands, 1989), edited by G. van der Geer et al., Progr. Math. **89**, Birkhäuser, Boston, 1991. MR Zbl

[Kaneko and Zagier 1998] M. Kaneko and D. Zagier, "Supersingular j-invariants, hypergeometric series, and Atkin's orthogonal polynomials", pp. 97–126 in *Computational perspectives on number theory* (Chicago, 1995), edited by D. A. Buell and J. T. Teitelbaum, AMS/IP Stud. Adv. Math. **7**, Amer. Math. Soc., Providence, RI, 1998. MR Zbl

[Katz 1973] N. M. Katz, "p-adic properties of modular schemes and modular forms", pp. 69–190 in *Modular functions of one variable, III* (Antwerp, Belgium, 1972), edited by W. Kuyk and J.-P. Serre, Lecture Notes in Math. **350**, Springer, 1973. MR Zbl

[Kohnen 2003] W. Kohnen, "Transcendence of zeros of Eisenstein series and other modular functions", *Comment. Math. Univ. St. Pauli* **52**:1 (2003), 55–57. MR Zbl

[Lang 1973] S. Lang, *Elliptic functions*, Addison-Wesley, Reading, MA, 1973. MR Zbl

[Lang 1994] S. Lang, *Algebraic number theory*, 2nd ed., Graduate Texts in Math. **110**, Springer, 1994. MR Zbl

[Linnik 1968] Y. V. Linnik, *Ergodic properties of algebraic fields*, Ergebnisse der Mathematik **45**, Springer, 1968. MR Zbl

[Linnik and Skubenko 1964] Y. V. Linnik and B. F. Skubenko, "Asymptotic distribution of integral matrices of third order", *Vestnik Leningrad. Univ. Ser. Mat. Meh. Astronom.* **19**:3 (1964), 25–36. In Russian. MR Zbl

[Mestre 1986] J.-F. Mestre, "La méthode des graphes: exemples et applications", pp. 217–242 in *Proc. Int. Conf. on Class Numbers and Fundamental Units of Algebraic Number Fields* (Katata, Japan, 1986), edited by Y. Yamamoto and H. Yokoi, Nagoya Univ., 1986. MR Zbl

[Michel and Venkatesh 2006] P. Michel and A. Venkatesh, "Equidistribution, L-functions and ergodic theory: on some problems of Yu. Linnik", pp. 421–457 in *Proc. Int. Congr. Math., II* (Madrid, 2006), edited by M. Sanz-Solé et al., Eur. Math. Soc., Zürich, 2006. MR Zbl

[Richard 2018] R. Richard, "Répartition galoisienne ultramétrique d'une classe d'isogénie de courbes elliptiques: le cas de la mauvaise réduction", *J. Théor. Nombres Bordeaux* **30**:1 (2018), 1–18. MR Zbl

[Rivera-Letelier 2003] J. Rivera-Letelier, "Dynamique des fonctions rationnelles sur des corps locaux", pp. 147–230 in *Geometric methods in dynamics, II*, edited by W. de Melo et al., Astérisque **287**, Soc. Math. France, Paris, 2003. MR Zbl

[Roquette 1970] P. Roquette, *Analytic theory of elliptic functions over local fields*, Hamburger Math. Einzelschriften (N.F.) **1**, Vandenhoeck & Ruprecht, Göttingen, 1970. MR Zbl

[Shimura 1971] G. Shimura, *Introduction to the arithmetic theory of automorphic functions*, Publ. Math. Soc. Japan **11**, Iwanami Shoten, Tokyo, 1971. MR Zbl

[Siegel 1935] C. L. Siegel, "Über die Classenzahl quadratischer Zahlkörper", *Acta Arith.* **1**:1 (1935), 83–86. Zbl

[Silverman 1994] J. H. Silverman, *Advanced topics in the arithmetic of elliptic curves*, Graduate Texts in Math. **151**, Springer, 1994. MR Zbl

[Silverman 2009] J. H. Silverman, *The arithmetic of elliptic curves*, 2nd ed., Graduate Texts in Math. **106**, Springer, 2009. MR Zbl

[Tate 1995] J. Tate, "A review of non-Archimedean elliptic functions", pp. 162–184 in *Elliptic curves, modular forms, & Fermat's last theorem* (Hong Kong, 1993), edited by J. Coates and S.-T. Yau, Ser. Number Theory **1**, Int. Press, Cambridge, MA, 1995. MR Zbl

[Zhang 2001] S. Zhang, "Heights of Heegner points on Shimura curves", *Ann. of Math.* (2) **153**:1 (2001), 27–147. MR Zbl

Communicated by Shou-Wu Zhang
Received 2018-11-03 Revised 2019-11-16 Accepted 2020-02-06

sebastian.herrero.m@gmail.com *Instituto de Matemáticas, Pontificia Universidad Católica de Valparaíso, Chile*

rmenares@mat.uc.cl *Facultad de Matemáticas, Pontificia Universidad Católica de Chile, Santiago, Chile*

riveraletelier@gmail.com *Mathematics Department, University of Rochester, NY, United States*

Algebra & Number Theory

msp.org/ant

See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2020 is US $415/year for the electronic version, and $620/year (+$60, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

ANT peer review and production are managed by EditFLOW® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

Algebra & Number Theory

Volume 14 No. 5 2020