NONARCHIMEDEAN EQUIDISTRIBUTION ON ELLIPTIC
CURVES WITH GLOBAL APPLICATIONS

CLAYTON PETSCHE

NONARCHIMEDEAN EQUIDISTRIBUTION ON ELLIPTIC CURVES WITH GLOBAL APPLICATIONS

CLAYTON PETSCHE

Let $\mathbb{K}$ be an algebraically closed, complete, nonarchimedean field, let $E/\mathbb{K}$ be an elliptic curve, and let E denote the Berkovich analytic space associated to $E/\mathbb{K}$. We study the μ-equidistribution of finite subsets of $E(\mathbb{K})$, where μ is a certain canonical unit Borel measure on E. Our main result is an inequality bounding the error term when testing against a certain class of continuous functions on E. We then give two applications to elliptic curves over global function fields: We prove a function field analogue of the Szpiro–Ullmo–Zhang equidistribution theorem for small points, and a function field analogue of a result of Baker, Ih, and Rumely on the finiteness of S-integral torsion points. Both applications are given in explicit quantitative form.

Introduction

The main local inequality. Let $\mathbb{K}$ be a field that is algebraically closed and complete with respect to a nontrivial, nonarchimedean absolute value $|\cdot|$, and let $E/\mathbb{K}$ be an elliptic curve. The main result of this paper is an inequality that measures the equidistribution of finite subsets of $E(\mathbb{K})$ with respect to a certain canonical unit Borel measure μ defined on the Berkovich analytic space E associated to $E/\mathbb{K}$.

We will give a thorough review of the space E in Section 1, but briefly it is a path-connected, compact, Hausdorff topological space that contains $E(\mathbb{K})$ (with the topology induced by the absolute value on $\mathbb{K}$) as a dense subspace. The set $E \setminus E(\mathbb{K})$ is endowed with a canonical path metric $\rho(x, y)$ giving it the structure of an (infinite) metrized graph. In particular, there exists a distinguished finite subgraph Σ of E called the *skeleton*, along with a strong deformation retraction $r_\Sigma : E \to \Sigma$, such that each connected component of $E \setminus \Sigma$ is homeomorphic to a simply connected open Berkovich disc $B°$. If the j-invariant j_E is integral ($|j_E| \leq 1$), the set Σ consists of a single point, and so in particular E is simply connected. If j_E is not integral ($|j_E| > 1$), the set Σ is isometric to a circle of circumference $\log|j_E|$, and in this case the fundamental group of E is infinite cyclic.

The canonical unit Borel measure μ on E is supported on Σ, and is defined as follows: If j_E is integral, μ is the Dirac measure supported at the point Σ, while if j_E is not integral, μ is the normalized uniform measure supported on the circle Σ.

Let Z be a nonempty finite subset of $E(\mathbb{K})$. In Section 2 we will define a nonnegative real number $D(Z)$ called the local discrepancy of Z, which is closely related to the Néron local height function on $E(\mathbb{K})$, and which is small precisely when the set Z is close to being μ-equidistributed. Let $\mathscr{C}(\mathsf{E}, \mathbb{R})$ denote the space of continuous real-valued functions on E. Our main local result is summarized in the following theorem.

Theorem 1. *There exists a dense subspace $\mathscr{S}(\mathsf{E}, \mathbb{R})$ of $\mathscr{C}(\mathsf{E}, \mathbb{R})$ such that*

$$\left| \frac{1}{|Z|} \sum_{P \in Z} F(P) - \int F \, d\mu \right| \leq C_1(F) \Big(D(Z) + \frac{C_2(F)}{|Z|} \Big)^{1/2}$$

for all $F \in \mathscr{S}(\mathsf{E}, \mathbb{R})$ and all nonempty finite subsets Z of $E(\mathbb{K})$, where $C_1(F)$ and $C_2(F)$ are constants depending only on F.

Roughly speaking, $\mathscr{S}(\mathsf{E}, \mathbb{R})$ is the space of continuous functions $F : \mathsf{E} \to \mathbb{R}$ whose derivative F' is supported on some finite subgraph of E and is square-integrable. The constant $C_1(F)$ is essentially the L^2-norm of F', and $C_2(F)$ is a quantity measuring the size of the support of F'. In Section 2 we will restate this result with precise definitions of $\mathscr{S}(\mathsf{E}, \mathbb{R})$, $C_1(F)$, and $C_2(F)$. As applications of Theorem 1 we prove several global results summarized in the following two sections.

The global equidistribution theorem. Let K be a global function field. We will state precisely what this means in Section 3, but an example is the field $K = k(C)$ of rational functions on an integral, proper, geometrically connected curve C defined over an arbitrary field k. At each place v of K, let K_v be the completion of K at v, and let $\mathbb{K}_v$ be the completion of the algebraic closure of K_v; the field $\mathbb{K}_v$ is both complete and algebraically closed [Bosch et al. 1984, Section 3.4].

Let E/K be an elliptic curve, and at each place v of K, let E_v be the Berkovich analytic space associated to $E/\mathbb{K}_v$; more generally, we affix a subscript v to all local objects associated to $E/\mathbb{K}_v$, including the canonical measure μ_v on E_v. Recall that the Néron–Tate canonical height function $\hat{h} : E(\overline{K}) \to \mathbb{R}$ is nonnegative and vanishes on the torsion subgroup $E(\overline{K})_{\mathrm{tor}}$ of $E(\overline{K})$. Given a finite set Z of points in $E(\overline{K})$, define the height of Z to be the average $\hat{h}(Z) = |Z|^{-1} \sum_{P \in Z} \hat{h}(P)$ of the heights of its points. Using Theorem 1 we will prove an inequality measuring the local μ_v-equidistribution of a large set Z of small global points in $E(\overline{K})$:

Theorem 2. *Let E/K be an elliptic curve over a global function field, let v be a place of K, and let $\epsilon : \overline{K} \hookrightarrow \mathbb{K}_v$ be a K-embedding. There exists a dense subspace*

$\mathscr{S}(\mathsf{E}_v, \mathbb{R})$ *of* $\mathscr{C}(\mathsf{E}_v, \mathbb{R})$ *such that*

$$\left| \frac{1}{|Z|} \sum_{P \in Z} F(\epsilon(P)) - \int F d\mu_v \right| \leq C_1(F) \left(4\hat{h}(Z) + \frac{h(j_E)}{12|Z|} + \frac{C_2(F)}{|Z|} \right)^{1/2}$$

for all $F \in \mathscr{S}(\mathsf{E}_v, \mathbb{R})$ *and all nonempty finite* $\mathrm{Aut}(\overline{K}/K)$-*stable subsets* Z *of* $E(\overline{K})$, *where* $h(j_E)$ *denotes the absolute Weil height of the* j-*invariant of* E, *and where* $C_1(F)$ *and* $C_2(F)$ *are the same constants associated to the function* $F \in \mathscr{S}(\mathsf{E}_v, \mathbb{R})$ *as in Theorem 1.*

Using the fact that $\mathscr{S}(\mathsf{E}_v, \mathbb{R})$ is dense in $\mathscr{C}(\mathsf{E}_v, \mathbb{R})$, we will deduce the following corollary, a function field analogue of the Szpiro–Ullmo–Zhang equidistribution theorem for elliptic curves.

Corollary 3. *Let* E/K *be an elliptic curve over a global function field, let* v *be a place of* K, *and let* $\epsilon : \overline{K} \hookrightarrow \mathbb{K}_v$ *be a* K-*embedding. Let* $\{Z_n\}_{n=1}^{\infty}$ *be a sequence of finite* $\mathrm{Aut}(\overline{K}/K)$-*stable subsets of* $E(\overline{K})$ *with* $\hat{h}(Z_n) \to 0$ *and* $|Z_n| \to +\infty$. *Then*

(1) $$\lim_{n \to +\infty} \frac{1}{|Z_n|} \sum_{P \in Z_n} F(\epsilon(P)) = \int F d\mu_v$$

for all continuous functions $F : \mathsf{E}_v \to \mathbb{R}$.

We will prove Theorem 2 and Corollary 3 in Section 3.

The finiteness of S-integral torsion points. Let K be a global function field or a number field, and let E/K be an elliptic curve given by a Weierstrass equation

(2) $$y^2 + a_1 xy + a_3 y = x^3 + a_2 x^2 + a_4 x + a_6.$$

Let S be a finite set of places of K (including the archimedean places if K is a number field) such that (2) is defined over the ring

$$\mathbb{O}_S = \{ a \in K \mid |a|_v \leq 1 \text{ for all places } v \notin S \}$$

of S-integers in K. Given two points $P, Q \in E(\overline{K})$, we say that P is S-integral *with respect to* Q if the Zariski closures of P and Q do not meet in the model $\mathscr{E}/\mathbb{O}_S$ of E associated to the Weierstrass equation (2).

For fixed $Q \in E(\overline{K})$, let $\mathscr{E}(\overline{\mathbb{O}}_S, Q)$ denote the set of points in $E(\overline{K})$ that are S-integral with respect to Q, and let $\mathscr{E}(\overline{\mathbb{O}}_S, Q)_{\mathrm{tor}} = E(\overline{K})_{\mathrm{tor}} \cap \mathscr{E}(\overline{\mathbb{O}}_S, Q)$ be the set of *torsion* points that are S-integral with respect to Q. In the number field case, Baker, Ih and Rumely [Baker et al. 2008] have shown that if $Q \in E(\overline{K})$ is a nontorsion point, then $\mathscr{E}(\overline{\mathbb{O}}_S, Q)_{\mathrm{tor}}$ is finite. More generally, Ih has conjectured analogous finiteness results for torsion points on abelian varieties and preperiodic points for dynamical systems on $\mathbb{P}^1$ over number fields. We will prove the following function field analogue of Baker, Ih and Rumely's theorem for elliptic curves.

Theorem 4. *Let E/K be an elliptic curve over a global function field. Let* (2) *be a Weierstrass equation for E/K, let S be a finite set of places of K such that* (2) *is defined over $\mathbb{O}_S$, and let $\mathscr{E}/\mathbb{O}_S$ denote the associated S-integral model for E. If Q is a point in $E(\overline{K})$ with $\hat{h}(Q) > 0$, then the set $\mathscr{E}(\overline{\mathbb{O}}_S, Q)_{\text{tor}}$ is finite.*

In the number field case, Baker, Ih and Rumely [Baker et al. 2008] treat the nonarchimedean places using a result of Chambert-Loir (which in our notation is equivalent to the equidistribution of torsion points after they have been retracted onto the skeleton Σ_v), in combination with a result of Cassels bounding the denominators of torsion points in characteristic zero. Our proof over function fields, which holds in arbitrary characteristic, requires the full equidistribution theorem on E_v, and in place of Cassels's inequality it uses the discreteness of torsion points in $E(\mathbb{K}_v)$. On the other hand, the lack of archimedean places makes the proof of Theorem 4 quite a bit less complicated than its number field analogue: The treatment of the archimedean places in [Baker et al. 2008] requires a strong quantitative equidistribution result on torsion points in $E(\mathbb{C})$, along with an effective diophantine inequality on linear forms in elliptic logarithms due to David and Hirata-Kohno.

Using Theorem 2, we can prove a quantitative version of Theorem 4. First observe that for fixed $Q \in E(\overline{K})$, the set $\mathscr{E}(\overline{\mathbb{O}}_S, Q)$ depends on the choice of Weierstrass equation and the set S. However, enlarging the set S only makes the set $\mathscr{E}(\overline{\mathbb{O}}_S, Q)_{\text{tor}}$ larger, and since any two Weierstrass equations for E/K give rise to isomorphic models over $\mathbb{O}_S$ for some (sufficiently large) finite set S of places of K, we see that the finiteness of $\mathscr{E}(\overline{\mathbb{O}}_S, Q)_{\text{tor}}$ for all S is independent of the Weierstrass equation (2). Moreover, by enlarging the set S if necessary, we may assume that the Weierstrass equation (2) has good reduction outside S. Under this additional hypothesis we can give an explicit bound on $\mathscr{E}(\overline{\mathbb{O}}_S, Q)_{\text{tor}}$.

Theorem 5. *Let E/K, S, and $\mathscr{E}/\mathbb{O}_S$ satisfy the same hypotheses as in Theorem 4, and assume in addition that S contains all places v of K such that $|\Delta|_v \neq 1$, where $\Delta \in K$ is the discriminant of* (2). *If Q is a point in $E(\overline{K})$ with $\hat{h}(Q) > 0$, then*

$$(3) \qquad |\mathscr{E}(\overline{\mathbb{O}}_S, Q)_{\text{tor}}| \leq \frac{1}{\hat{h}(Q)^2} \left(\frac{|S| h(j_E)}{12} + \sum_{v \in S} m_v(Q) \right)^2.$$

In this inequality, $h(j_E)$ denotes the absolute Weil height of the j-invariant of E, and for each $v \in S$, $m_v(Q)$ is a certain nonnegative quantity that is large when some embedding of Q into $E(\mathbb{K}_v)$ is v-adically close to a torsion point; we will define this precisely in Section 4. We point out that Theorem 5 holds even when S is empty; in that case E must be defined over the constant field K_0 of K, and our proof shows that $\mathscr{E}(\overline{\mathbb{O}}_S, Q)_{\text{tor}}$ is empty.

Other work. It is now well understood that Berkovich's theory of analytic spaces provides a suitable framework in which to prove nonarchimedean equidistribution theorems of the type of Szpiro, Ullmo and Zhang's; see for example [Baker and Rumely 2006; Chambert-Loir 2006; Favre and Rivera-Letelier 2006; Thuillier 2005]. Until very recently such results have been worked out in the number field case only. However, independently Faber [2008] and Gubler [2008] have now proved a general equidistribution result for dynamical systems on varieties over function fields. When $K = k(C)$, the results of Faber and Gubler contain our Corollary 3 as a special case; on the other hand, we treat a wider class of function fields than in [Faber 2008], and our results are quantitative.

Favre and Rivera-Letelier [2006] have proved a quantitative equidistribution result for small points with respect to dynamical systems on $\mathbb{P}^1$ over number fields. The local results therein are proved using potential theory on the Berkovich projective line P^1 and work for a large class of Borel measures on P^1, while our techniques are more elementary and are formulated specifically for the *canonical* measure μ on E. However, Thuillier [2005] has developed potential theory on an arbitrary nonarchimedean curve; it would be interesting to use this theory to obtain more general equidistribution results in the spirit of Theorem 1.

Baker and Petsche [2005] established quantitative equidistribution results for elliptic curves over number fields, although they did not prove an explicit inequality on the error term as in Theorem 2. Rather, their main inequality is a bound on the local discrepancy $D_v(Z)$ at each place v, which is enough to deduce the qualitative equidistribution theorem and several quantitative corollaries. As is often the case with nonarchimedean equidistribution theorems of this type, they use the countability of the residue field of $\mathbb{C}_v$ (over a number field) to show that the limiting measure is supported on the skeleton Σ_v of E_v. Our proof of Corollary 3 gets around this assumption by establishing the limit formula (via Theorem 2) for the dense class $\mathscr{S}(\mathsf{E}_v, \mathbb{R})$ of test functions. Note that the field $\mathbb{K}_v$ over a function field K may have uncountable residue field.

Our analytic treatment of E using the path metric $\rho(x, y)$ is heavily influenced by the work of Baker and Rumely, especially their monograph [2009] on the analytic theory of the Berkovich projective line P^1.

The rest of this paper is organized as follows. In Section 1 we give a detailed review of the topological and analytic structure of the Berkovich analytic space associated to an elliptic curve. We prove our main local result Theorem 1 in Section 2, and in Sections 3 and 4 we treat the global applications.

1. Nonarchimedean preliminaries

Overview. Let $\mathbb{K}$ be a field that is algebraically closed and complete with respect to a nontrivial, nonarchimedean absolute value $|\cdot|$. Let

$$\mathbb{O} = \{a \in \mathbb{K} \mid |a| \leq 1\} \quad \text{and} \quad \mathbb{M} = \{a \in \mathbb{K} \mid |a| < 1\}$$

respectively be the ring of integers in $\mathbb{K}$ and its maximal ideal. Let $k = \mathbb{O}/\mathbb{M}$ denote the residue field, which is algebraically closed [Bosch et al. 1984, Section 3.4].

Let $E/\mathbb{K}$ be an elliptic curve with j-invariant $j_E \in \mathbb{K}$. In this section we will give a fairly detailed review of the Berkovich analytic space E associated to $E/\mathbb{K}$, its canonical path metric $\rho(x, y)$, and its relation to the Néron local height function λ on $E(\mathbb{K})$. This is partly to make this paper as self-contained as possible; in particular, to understand our results it is not necessary to have any prior knowledge of the general theory of Berkovich analytic spaces. Moreover, we intend to make essential use of the metrized graph structure on E, and we will put special emphasis on the completely explicit nature of our main inequality and its corollaries. We will therefore need to have an explicit description of the space E and its canonical path metric $\rho(x, y)$. See [Berkovich 1990] for the theoretical foundation of Berkovich analytic spaces in general, and see [Baker and Rumely 2009; Baker 2008] for more detailed expositions on the Berkovich projective line and more general curves.

1.1. ***Elliptic curves and integral models.*** In this section we will fix a model $\mathscr{E}/\mathbb{O}$ associated to an integral Weierstrass equation for $E/\mathbb{K}$, and we will record a basic lemma. In later sections we will use this model to explicitly construct the Berkovich analytic space E and its associated path metric $\rho(x, y)$, working out the details in the cases of integral and nonintegral j-invariant separately. We will then explain the topology on E and its metrized graph structure in Section 1.7.

First suppose that j_E is integral ($|j_E| \leq 1$). Then there exists a Weierstrass equation

$$(4) \qquad y^2 + a_1 xy + a_3 y = x^3 + a_2 x^2 + a_4 x + a_6$$

for E with integral coefficients $a_1, a_3, a_2, a_4, a_6 \in \mathbb{O}$ and unit discriminant $\Delta \in \mathbb{O}^\times$. Letting $\mathscr{E}/\mathbb{O}$ denote the associated integral model for E, the special fiber $\overline{\mathscr{E}}$ is an elliptic curve over k, and the reduction map $\pi : E(\mathbb{K}) \to \overline{\mathscr{E}}(k)$ is a surjective group homomorphism. In this situation $E/\mathbb{K}$ is said to have good reduction.

Suppose now that j_E is not integral ($|j_E| > 1$). Then we let $\mathscr{E}/\mathbb{O}$ denote the integral model for E associated to the Weierstrass equation

$$(5) \qquad y^2 + xy = x^3 + a_4(q)x + a_6(q)$$

afforded by Tate's uniformization theory; see [Tate 1995] and [Silverman 1994, Section V.3]. Here $q \in \mathbb{K}^\times$ is a uniformizing parameter satisfying $|q| = |1/j_E| < 1$,

and $a_4(q), a_6(q)$ are certain elements of $\mathbb{K}$ defined by convergent integral power series in q. In particular we have a group isomorphism $\phi : \mathbb{K}^\times/q^{\mathbb{Z}} \xrightarrow{\sim} E(\mathbb{K})$ given by $\phi(u) = (X(u, q), Y(u, q))$ for $u \notin q^{\mathbb{Z}}$, where $X(u, q)$ and $Y(u, q)$ can be given explicitly in terms of u and q. Let $\pi : E(\mathbb{K}) \to \overline{\mathscr{E}}(k)$ denote the reduction map onto the special fiber (which is singular). The set $\overline{\mathscr{E}}_{\mathrm{ns}}(k)$ of nonsingular points on the special fiber is a group variety that is isomorphic to $k^\times$, and thus $E/\mathbb{K}$ is said to have (split) multiplicative reduction. Letting $E_0(\mathbb{K}) = \pi^{-1}(\overline{\mathscr{E}}_{\mathrm{ns}}(k))$ denote the set of points with nonsingular reduction, the map π restricts to a surjective group homomorphism $\pi_0 : E_0(\mathbb{K}) \to \overline{\mathscr{E}}_{\mathrm{ns}}(k) \simeq k^\times$. The set $\{u \in \mathbb{K}^\times \mid |q| < |u| \le 1\}$ is a fundamental domain for the quotient $\mathbb{K}^\times/q^{\mathbb{Z}}$. Given an element u in this domain, we have $\phi(u) \in E_0(\mathbb{K})$ if and only if $|u| = 1$. Moreover, when $|u| = 1$, the point $\phi(u) \in E_0(\mathbb{K})$ reduces to the identity element of $\overline{\mathscr{E}}_{\mathrm{ns}}(k)$ if and only if $|u - 1| < 1$.

Regardless of the reduction type of $E/\mathbb{K}$, we let $z = -x/y$ denote the standard local parameter at the origin with respect to the chosen Weierstrass equation. Let $B^\circ = \{P \in E(\mathbb{K}) \mid \pi(P) = \pi(O)\}$ denote the "kernel of reduction", that is, the set of points in $E(\mathbb{K})$ that reduce to the point $\pi(O) = (0 : 1 : 0) \in \overline{\mathscr{E}}(k) \subset \mathbb{P}^2(k)$. Note that B° is a subgroup of $E(\mathbb{K})$.

Lemma 6. *The map $z : B^\circ \to B^\circ(0, 1)$ is a bijection of B° with the open unit disc $B^\circ(0, 1)$ of $\mathbb{K}$, and $|z(P \pm Q)| = |z(P) \pm z(Q)|$ for all $P, Q \in B^\circ$.*

Before we get to the proof of this lemma, define

$$(6) \qquad d(P, Q) = \begin{cases} |z(P - Q)| & \text{if } P - Q \in B^\circ, \\ 1 & \text{if } P - Q \notin B^\circ \end{cases}$$

for any two points $P, Q \in E(\mathbb{K})$. This definition does not depend on our choice of Weierstrass equation. It is easy to check using Lemma 6 that $d(P, Q)$ defines a nonarchimedean metric on $E(\mathbb{K})$; we take the resulting metric topology as the definition of the topology on $E(\mathbb{K})$.

Proof of Lemma 6. Let $F(Z_1, Z_2) \in \mathbb{O}[[Z_1, Z_2]]$ be the formal group law associated to the integral Weierstrass equation, with inverse $i(Z) \in \mathbb{O}[[Z]]$, and let $w(Z) \in \mathbb{O}[[Z]]$ be the expansion of the function $z \mapsto w = -1/y$ on $E(\mathbb{K})$ at $z = 0$. These are formal power series with integral coefficients that converge when the variables Z, Z_1, Z_2 take values $z, z_1, z_1 \in B^\circ(0, 1)$; see [Silverman 1986, Section IV.1] for their definitions and basic properties.

Given $z \in B^\circ(0, 1)$, the point $\sigma(z) := (-z : 1 : -w(z)) \in \mathbb{P}^2(\mathbb{K})$ satisfies the Weierstrass equation and thus constitutes a point in $E(\mathbb{K})$. Since $w(Z) = Z^3 + u(Z)$, where all terms of $u(Z) \in \mathbb{O}[[Z]]$ have degree at least 4, we see that $|w(z)| < 1$, and thus $\sigma(z)$ reduces to $(0 : 1 : 0) \in \mathbb{P}^2(k)$; in other words $\sigma(z) \in B^\circ$. It follows that the map $\sigma : B^\circ(0, 1) \to B^\circ$ is the inverse of $z : B^\circ \to B^\circ(0, 1)$, and thus both maps are bijections.

In fact the maps $\sigma : \hat{E}(\mathbb{M}) \to B^\circ$ and $z : B^\circ \to \hat{E}(\mathbb{M})$ are group isomorphisms, where $\hat{E}(\mathbb{M})$ denotes the group that is equal to the set $B^\circ(0, 1)$ endowed with the group law $(z_1, z_2) \mapsto F(z_1, z_2)$, with inverse $z \mapsto i(z)$. Since $F(Z_1, Z_2) = Z_1 + Z_2 + G(Z_1, Z_2)$, where all terms of $G(Z_1, Z_2) \in \mathbb{O}[\![Z_1, Z_2]\!]$ have degree at least 2, we have $|z(P + Q)| = |F(z(P), z(Q))| = |z(P) + z(Q)|$. Moreover $i(Z) = -Z + j(Z)$, where all terms of $j(Z) \in \mathbb{O}[\![Z]\!]$ have degree at least 2, and thus $|z(-P)| = |i(z(P))| = |z(P)|$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

1.2. *The Berkovich unit disc.* In this section we will pause from the discussion of elliptic curves. We define the Berkovich unit disc $\mathsf{B}(0, 1)$ over $\mathbb{K}$ and the *open* Berkovich unit disc $\mathsf{B}^\circ(0, 1)$. We summarize their basic properties and describe their canonical path metric $\rho(x, y)$. Our treatment of $\mathsf{B}(0, 1)$ follows [Baker and Rumely 2009, Chapter 1].

Let $\mathbb{K}\langle T \rangle = \{f(T) = \sum_{\ell \geq 0} a_\ell T^\ell \mid a_\ell \in \mathbb{K}, \ |a_\ell| \to 0\}$ be the ring of formal power series converging in the unit disc $B(0, 1) = \{a \in \mathbb{K} \mid |a| \leq 1\}$ of $\mathbb{K}$. A multiplicative seminorm $[\,\cdot\,]$ on $\mathbb{K}\langle T \rangle$ is a nonnegative real-valued function on $\mathbb{K}\langle T \rangle$ that satisfies the axioms $[0] = 0$, $[1] = 1$, $[f(T) + g(T)] \leq [f(T)] + [g(T)]$, and $[f(T)g(T)] = [f(T)][g(T)]$ for all $f(T), g(T) \in \mathbb{K}\langle T \rangle$. A seminorm may fail to be a true norm in the functional-analytic sense since it may vanish on nonzero elements of $\mathbb{K}\langle T \rangle$. A multiplicative seminorm $[\,\cdot\,]$ is said to be bounded if there exists a constant $C > 0$ such that $[f(T)] \leq C \max_{\ell \geq 0} |a_\ell|$ for all $f(T) = \sum_{\ell \geq 0} a_\ell T^\ell \in \mathbb{K}\langle T \rangle$.

The Berkovich unit disc $\mathsf{B}(0, 1)$ is defined to be the set of bounded multiplicative seminorms $[\,\cdot\,]$ on the ring $\mathbb{K}\langle T \rangle$ that extend the absolute value $|\cdot|$ on $\mathbb{K}$. There exists a natural compact, Hausdorff topology on $\mathsf{B}(0, 1)$; it is defined as the weakest topology such that those subsets of the form

$$(7) \qquad \{x \in \mathsf{B}(0, 1) \mid \alpha < [f(T)]_x < \beta\} \quad \text{for } \alpha, \beta \in \mathbb{R} \text{ and } f(T) \in \mathbb{K}\langle T \rangle$$

are open, where here and throughout $[\,\cdot\,]_x$ denotes the seminorm corresponding to the point $x \in \mathsf{B}(0, 1)$. In other words, the collection of sets of the form (7) is a subbase for the topology on $\mathsf{B}(0, 1)$; see [Berkovich 1990, Section 1.2] and [Baker and Rumely 2009, Section 1.1].

To see some examples of elements of the Berkovich unit disc $\mathsf{B}(0, 1)$, note that each element a in the ordinary unit disc $B(0, 1)$ gives rise to the evaluation seminorm $[f(T)]_a = |f(a)|$, and this defines a dense embedding $B(0, 1) \hookrightarrow \mathsf{B}(0, 1)$. More generally, each closed disc $B(a, r) = \{z \in \mathbb{K} \mid |z - a| \leq r\}$, where $a \in B(0, 1)$ and $r \in \{0\} \cup |\mathbb{K}^\times|$ with $0 \leq r \leq 1$, defines a point $\zeta_{a,r}$ in $\mathsf{B}(0, 1)$ corresponding to the sup norm $[f(T)]_{\zeta_{a,r}} = \sup_{z \in B(a,r)} |f(z)|$. Thus we identify the "classical point" $a \in B(0, 1)$ with the "Berkovich point" $\zeta_{a,0} \in \mathsf{B}(0, 1)$ under this notation. It may happen (depending on the field $\mathbb{K}$) that the points $\zeta_{a,r}$ account for all elements of $\mathsf{B}(0, 1)$, but this is generally not the case. In fact, Berkovich [1990, Section 1.4.4]

has classified of all points of B(0, 1) into four *types*, of which the points $\zeta_{a,0}$ for $a \in B(0, 1)$ compose type 1, and the points $\zeta_{a,r}$ for $a \in B(0, 1)$ and $r \in |\mathbb{K}^\times|$ with $0 < r \leq 1$ compose type 2. Loosely speaking, the points of type 3 and 4 "fill in the holes" of the set $\{\zeta_{a,r}\}$ to create a compact, path-connected space B(0, 1).

The path-connectedness of B(0, 1) can be easily seen. First observe that B(0, 1) carries a natural partial order under which $x \leq y$ if and only if $[f(T)]_x \leq [f(T)]_y$ for all $f(T) \in \mathbb{K}\langle T \rangle$. Thus if $\zeta_{a,r}$ and $\zeta_{a',r'}$ are points of type 1 or 2, then $\zeta_{a,r} \leq \zeta_{a',r'}$ if and only if $B(a, r) \subseteq B(a', r')$. It follows from Berkovich's classification that the point $\zeta_{0,1}$ corresponding to the sup norm on the unit disc $B(0, 1)$ itself is the unique maximal point in B(0, 1); this point is commonly known as the *Gauss point* of B(0, 1). On the other hand, each type 1 point $a = \zeta_{a,0}$ in $B(0, 1)$ is minimal. Given two points $x, y \in$ B(0, 1) with $x \leq y$, denote by $[x, y]$ the set of all $z \in$ B(0, 1) satisfying $x \leq z \leq y$. It is a straightforward exercise to show that $[x, y]$ is homeomorphic to a closed subinterval of the real line. More generally, given any two $x, y \in$ B(0, 1) (perhaps with $x \not\leq y$), there exists a unique least upper bound $x \vee y$ of x and y; we then define $[x, y] = [x, x \vee y] \cup [y, x \vee y]$, which is again homeomorphic to a closed real interval; we will call such sets $[x, y]$ *line segments* in B(0, 1). We remark that $[x, y] = [y, x]$ for all $x, y \in$ B(0, 1); that is, the notation does not imply an order relationship between x and y.

In addition to the topological considerations above, there is a natural way to assign a notion of length to each line segment in B(0, 1). To do this we first consider the diameter function diam : B(0, 1) $\to$ [0, 1], $x \mapsto \inf_{a \in B(0,1)}[T - a]_x$; thus $\mathrm{diam}(\zeta_{a,r}) = r$ for points of type 1 and 2. For two points $x, y \in$ B(0, 1) \ $B(0, 1)$, define the length of the line segment $[x, y]$ by

$$\rho(x, y) = \log\left(\frac{\mathrm{diam}(x \vee y)^2}{\mathrm{diam}(x)\,\mathrm{diam}(y)}\right).$$

This defines a path metric on B(0, 1) \ $B(0, 1)$ under which the points of $B(0, 1)$ can be viewed as being at infinite distance from all other points. Moreover, any line segment $[x, y]$ in B(0, 1) \ $B(0, 1)$ is isometric to a real interval of length $\rho(x, y)$.

We note for future use the following fact: Let $x, y \in B(0, 1)$ be two distinct type 1 points. Then it is easy to see that $x \vee y = \zeta_{x,|x-y|}$, and therefore the distance from the Gauss point $\zeta_{0,1}$ to the point $x \vee y$ is

(8) $$\rho(\zeta_{0,1}, x \vee y) = -\log|x - y|.$$

For our purposes it will be more useful to work with the *open* Berkovich unit disc B°(0, 1), which is a certain proper subset of B(0, 1). To define B°(0, 1), let $B°(0, 1) = \{a \in \mathbb{K} \mid |a| < 1\}$ be the open unit disc in $B(0, 1)$. Given $a \in B(0, 1)$, it is clear that $a \in B°(0, 1)$ if and only if $0 \vee a \neq \zeta_{0,1}$ in B(0, 1). Motivated by this,

we define the open Berkovich unit disc by

$$\mathsf{B}^\circ(0, 1) = \{x \in \mathsf{B}(0, 1) \mid 0 \vee x \neq \zeta_{0,1}\}.$$

Thus $\mathsf{B}^\circ(0, 1) \cap B(0, 1) = B^\circ(0, 1)$. Like $\mathsf{B}(0, 1)$, the $\mathsf{B}^\circ(0, 1)$ is path connected, but unlike $\mathsf{B}(0, 1)$, it has no maximal element. On the other hand, the set $\mathsf{B}^\circ(0, 1) \cup \{\zeta_{0,1}\}$ obtained by adjoining the Gauss point is path connected, with $\zeta_{0,1}$ as its unique maximal element.

1.3. The space E in the good reduction case. Returning to our elliptic curve $E/\mathbb{K}$, suppose that $|j_E| \leq 1$. For each $\alpha \in \overline{\mathscr{E}}(k)$, let $B_\alpha^\circ = \pi^{-1}(\alpha) = \{P \in E(\mathbb{K}) \mid \pi(P) = \alpha\}$ be the set of points reducing to α, and fix a point $P_\alpha \in B_\alpha^\circ$. Lemma 6 implies that the map

$$(9) \qquad\qquad \nu_\alpha : B_\alpha^\circ \to B^\circ(0, 1), \quad P \mapsto z(P - P_\alpha)$$

is a homeomorphism of B_α° with the open unit disc $B^\circ(0, 1)$ of $\mathbb{K}$. We thus have a decomposition of $E(\mathbb{K})$ into a disjoint union

$$(10) \qquad\qquad E(\mathbb{K}) = \coprod_{\alpha \in \overline{\mathscr{E}}(k)} B_\alpha^\circ$$

of subsets that are homeomorphic to open discs.

For each $\alpha \in \overline{\mathscr{E}}(k)$, let B_α° denote the open Berkovich unit disc associated to the open unit disc B_α°, and let $\mathsf{B}_\alpha^\circ \cup \{\zeta_\alpha\}$ denote the union of B_α° with its associated Gauss point ζ_α, as described in Section 1.2. The Berkovich analytic space E is the union of the $\mathsf{B}_\alpha^\circ \cup \{\zeta_\alpha\}$ over all $\alpha \in \overline{\mathscr{E}}(k)$, where the points ζ_α are identified to form a single point $\zeta \in \mathsf{E}$. The set $\Sigma = \{\zeta\}$ is the skeleton of E, and we therefore have a decomposition

$$(11) \qquad\qquad \mathsf{E} \setminus \Sigma = \coprod_{\alpha \in \overline{\mathscr{E}}(k)} \mathsf{B}_\alpha^\circ$$

of $\mathsf{E} \setminus \Sigma$ into its connected components B_α°. The constant map $r_\Sigma : \mathsf{E} \to \Sigma$ defines a strong deformation retraction of E onto Σ; in particular, E is simply connected.

1.4. The space E in the multiplicative reduction case. Suppose now that $|j_E| > 1$. Let $u : E(\mathbb{K}) \xrightarrow{\sim} \mathbb{K}^\times / q^{\mathbb{Z}}$ be the inverse of the Tate isomorphism ϕ discussed in Section 1.1. Then

$$(12) \qquad\qquad r : E(\mathbb{K}) \to \Sigma, \quad P \mapsto -\log|u(P)|$$

is a group homomorphism, where $\Sigma = \mathbb{R}/(\log|j_E|)\mathbb{Z}$ denotes the circle group of circumference $\log|j_E| > 0$. As explained in Section 1.1, the kernel of the map r is precisely the set $E_0(\mathbb{K})$ of points in $E(\mathbb{K})$ with nonsingular reduction. More

generally, let $\Sigma_r \subseteq \Sigma$ denote the image of r, and for each $s \in \Sigma_r$, let $E_s(\mathbb{K}) = r^{-1}(s)$ denote its r-preimage.

For each $s \in \Sigma_r$, select $P_{s,0} \in E_s(\mathbb{K})$. Then for each pair $(s, \alpha) \in \Sigma_r \times \overline{\mathscr{E}}_{\mathrm{ns}}(k)$ define a subset

$$B_{s,\alpha}^\circ = \{P \in E_s(\mathbb{K}) \mid \pi_0(P - P_{s,0}) = \alpha\}$$

of $E_s(\mathbb{K})$. We have decomposed $E(\mathbb{K})$ into a disjoint union

(13)
$$E(\mathbb{K}) = \coprod_{(s,\alpha)\in\Sigma_r\times\overline{\mathscr{E}}_{\mathrm{ns}}(k)} B_{s,\alpha}^\circ.$$

Note that $P_{s,0} \in B_{s,0}^\circ$. More generally, for each pair (s, α) with $\alpha \neq 0$, select a point $P_{s,\alpha} \in B_{s,\alpha}^\circ$. It follows from Lemma 6 that the map

(14)
$$v_{s,\alpha} : B_{s,\alpha}^\circ \to B^\circ(0, 1), \quad P \mapsto z(P - P_{s,\alpha})$$

is a homeomorphism of $B_{s,\alpha}^\circ$ with the open unit disc $B^\circ(0, 1)$ of $\mathbb{K}$.

For each pair $(s, \alpha) \in \Sigma_r \times \overline{\mathscr{E}}_{\mathrm{ns}}(k)$, let $\mathsf{B}_{s,\alpha}^\circ$ denote the open Berkovich unit disc associated to the open unit disc $B_{s,\alpha}^\circ$, and let $\mathsf{B}_{s,\alpha}^\circ \cup \{\zeta_{s,\alpha}\}$ denote the union of $\mathsf{B}_{s,\alpha}^\circ$ with its associated Gauss point $\zeta_{s,\alpha}$, as described in Section 1.2. The Berkovich analytic space E is the union of the circle Σ with the sets $\mathsf{B}_{s,\alpha}^\circ \cup \{\zeta_{s,\alpha}\}$ over all $(s, \alpha) \in \Sigma_r \times \overline{\mathscr{E}}_{\mathrm{ns}}(k)$, where for each $s \in \Sigma_r$ the points $\zeta_{s,\alpha}$ are identified with each other and with the point $s \in \Sigma_r$ to form a single point s of E. Thus E can be visualized as a collection of open Berkovich discs $\mathsf{B}_{s,\alpha}^\circ$ glued together along the skeleton Σ. We have a decomposition

(15)
$$\mathsf{E} \setminus \Sigma = \coprod_{(s,\alpha)\in\Sigma_r\times\overline{\mathscr{E}}_{\mathrm{ns}}(k)} \mathsf{B}_{s,\alpha}^\circ$$

of $\mathsf{E} \setminus \Sigma$ into its connected components $\mathsf{B}_{s,\alpha}^\circ$. The homomorphism $r : E(\mathbb{K}) \to \Sigma$ defined in (12) extends to a strong deformation retraction $r_\Sigma : \mathsf{E} \to \Sigma$; in particular, the fundamental group of E is isomorphic to $\pi_1(\Sigma) \simeq \mathbb{Z}$.

Note that the parametrization of the sets $\mathsf{B}_{s,\alpha}^\circ$ by $\Sigma_r \times \overline{\mathscr{E}}_{\mathrm{ns}}(k)$ is noncanonical, due to the arbitrary choice of each point $P_{s,\alpha}$ in $B_{s,\alpha}^\circ$.

1.5. *The path metric*. The path metric $\rho(x, y)$ on $\mathsf{E} \setminus E(\mathbb{K})$ is the unique metric that restricts to the canonical path metric on each connected component B° of $\mathsf{E} \setminus E(\mathbb{K})$, as described in Section 1.2, and that in the multiplicative reduction case restricts to the usual path metric on the circle $\Sigma \simeq \mathbb{R}/(\log|j_E|)\mathbb{Z}$. Under this metric, each point of $E(\mathbb{K})$ can be viewed as being at "infinite distance" from all other points in E.

We caution that the metric topology on $\mathsf{E} \setminus E(\mathbb{K})$ is not the same as the subspace topology induced by the Hausdorff topology on E. However, the metric $\rho(x, y)$ and the topology on E are related to one another, as described in Section 1.7.

1.6. *The Néron function.* Let $\lambda : E(\mathbb{K}) \backslash \{O\} \to \mathbb{R}$ be the Néron function as defined and normalized in [Silverman 1994, Section VI.1]. There exists a natural extension of λ to a continuous function on $\mathsf{E} \backslash \{O\}$; this extension is closely related to the analytic structure of the space E, its skeleton Σ, and the path metric $\rho(x, y)$.

To describe this extension, we begin by defining two continuous functions

$$i : \mathsf{E} \times \mathsf{E} \to \mathbb{R} \cup \{+\infty\} \quad \text{and} \quad j : \mathsf{E} \times \mathsf{E} \to \mathbb{R}.$$

Letting $\operatorname{diag}(E(\mathbb{K}))$ denote the diagonal of $E(\mathbb{K}) \times E(\mathbb{K})$ in $\mathsf{E} \times \mathsf{E}$, we first declare that $i(x, y) = +\infty$ if $(x, y) \in \operatorname{diag}(E(\mathbb{K}))$. For $(x, y) \in \mathsf{E} \times \mathsf{E} \backslash \operatorname{diag}(E(\mathbb{K}))$, define $i(x, y) = 0$ if x and y are not both elements of the same connected component B° of $\mathsf{E} \backslash \Sigma$, as in one of the decompositions (11) or (15). On the other hand, if both x and y are elements of the same connected component B°, let $x \vee y \in \mathsf{E} \backslash E(\mathbb{K})$ be their least upper bound with respect to the partial order on B°, and define

$$(16) \qquad\qquad i(x, y) = \rho(r_{\Sigma}(x \vee y), x \vee y).$$

In other words, $i(x, y)$ is the distance between $x \vee y$ and its nearest point on the skeleton Σ. Note that $i(x, y)$ is nonnegative, finite outside of $\operatorname{diag}(E(\mathbb{K}))$, and positive if and only if x and y lie in the same connected component of $\mathsf{E} \backslash \Sigma$. In particular, if $P, Q \in E(\mathbb{K})$ then

$$(17) \qquad i(P, Q) > 0 \ \text{ if and only if } \ \pi(P) = \pi(Q) \ \text{ and } \ r_{\Sigma}(P) = r_{\Sigma}(Q).$$

The number $i(P, Q)$ is closely related to the intersection multiplicity of the two points $P, Q \in E(\mathbb{K})$ in the integral model $\mathscr{E}/\mathbb{O}$; for example, if $E/\mathbb{K}$ has good reduction, (17) shows that $i(P, Q) = 0$ if and only if P and Q do not meet in $\mathscr{E}$.

Finally we note that on $E(\mathbb{K})$ the function $i(P, Q)$ is related to the metric (6) by $i(P, Q) = -\log d(P, Q)$; this is easy to see using the definitions (6) and (16) and the rule (8).

Let $j(x, y)$ be identically zero if $E/\mathbb{K}$ has good reduction. If $E/\mathbb{K}$ has multiplicative reduction, let $\ell = \log|j_E| = \log|1/q| > 0$, let $r_{\Sigma} : \mathsf{E} \to \Sigma \simeq \mathbb{R}/\ell\mathbb{Z}$ denote the retraction map, and define

$$(18) \qquad\qquad j(x, y) = \tfrac{1}{2}\ell\Phi((r_{\Sigma}(x) - r_{\Sigma}(y))/\ell),$$

where $\Phi(t) = (t - [t])^2 - (t - [t]) + 1/6$ is the second periodic Bernoulli polynomial. Regardless of the reduction type of $E/\mathbb{K}$, we see that $j(x, y)$ is bounded and symmetric, it factors through the retraction $r_{\Sigma} : \mathsf{E} \to \Sigma$ onto the skeleton, and $j(x, y) = \frac{1}{12}\log^{+}|j_E|$ if $r_{\Sigma}(x) = r_{\Sigma}(y)$.

The function $j(x, y)$ is natural when viewed in the context of the Laplacian on Σ as a metrized graph. In the language of [Baker and Rumely 2007, Section 1.8], $j(x, y)$ is the unique normalized Arakelov–Green's function on the circle Σ with respect to the uniform probability measure.

Proposition 7. *Let P and Q be distinct points in $E(\mathbb{K})$. Then*

$$(19) \qquad \lambda(P - Q) = i(P, Q) + j(P, Q).$$

In view of this proposition we define a function $\lambda : \mathsf{E} \times \mathsf{E} \to \mathbb{R} \cup \{+\infty\}$ by $\lambda(x, y) = i(x, y) + j(x, y)$. Thus $\lambda(x, y) = +\infty$ if and only if $(x, y) \in \mathrm{diag}(E(\mathbb{K}))$, and Proposition 7 states that $\lambda(P, Q) = \lambda(P - Q)$ for distinct points $P, Q \in E(\mathbb{K})$. The decomposition (19) is analogous to the one Chinburg and Rumely [1993] introduced in the context of the reduction graph associated to an arithmetic surface over a discrete valuation ring; from the point of view of Berkovich analytic spaces, this is closely related to the skeleton Σ of E.

Proof of Proposition 7. First suppose that $E/\mathbb{K}$ has good reduction. Then

$$(20) \qquad \lambda(P - Q) = \tfrac{1}{2} \log^+ |x(P - Q)| = \log^+ |z(P - Q)^{-1}|,$$

where $z = -x/y$ is the local parameter at the origin associated to the integral Weierstrass equation (4) in the variables x and y. The first identity in (20) is proved in [Silverman 1994, Theorem VI.4.1].

To see the second identity in (20), let us abbreviate $x = x(P-Q)$, $y = y(P-Q)$, and $z = z(P - Q)$. First suppose that $|x| > 1$. Since the Weierstrass equation (4) has coefficients in $\mathbb{O}$, it follows from the ultrametric inequality that $|y|^2 = |x|^3$, whereby $|z^{-1}| = |y|/|x| = |x|^{1/2} > 1$. This proves the second identity in (20) in this case.

Now consider the case $|x| \leq 1$. It suffices to show that $|z^{-1}| \leq 1$, which will imply that both $\log^+ |x|$ and $\log^+ |z^{-1}|$ vanish, completing the proof of the second identity in (20). Suppose on the contrary that $|z^{-1}| > 1$. Then $|z| < 1$, and therefore $|w(z)| < 1$, where $w(Z) \in \mathbb{O}[\![Z]\!]$ is the power series expansion of the function $z \mapsto w = -1/y$ on $E(\mathbb{K})$ at $z = 0$, as discussed in the proof of Lemma 6. Thus $|y|^{-1} = |w| < 1$, whereby $|y| > 1$. But since $|x| \leq 1$, it follows from the ultrametric inequality that the left side of the Weierstrass equation (4) has absolute value $|y|^2 > 1$, while the right side of (4) has absolute value ≤ 1. The contradiction proves that $|z^{-1}| \leq 1$, as desired, and completes the proof of the second identity in (20).

We will now use (20) to prove (19). By definition $j(P, Q) = 0$, and $i(P, Q) = 0$ if P and Q are not both elements of the same disc B_α° as defined in Section 1.3. On the other hand, suppose P and Q are elements of B_α°, let $v_\alpha : B_\alpha^\circ \to B^\circ(0, 1)$ denote the homeomorphism (9), and let $v_\alpha(P) \vee v_\alpha(Q)$ denote their least upper bound in $B^\circ(0, 1)$. Then

$$(21) \qquad \begin{aligned} i(P, Q) &= \rho(r_\Sigma(v_\alpha(P) \vee v_\alpha(Q)), v_\alpha(P) \vee v_\alpha(Q)) \\ &= -\log|v_\alpha(P) - v_\alpha(Q)| \qquad \text{(by (8))} \\ &= -\log|z(P - P_\alpha) - z(Q - P_\alpha)| \\ &= -\log|z(P - Q)| > 0 \qquad \text{(by Lemma 6).} \end{aligned}$$

It follows that $\lambda(P - Q) = \log^+|z(P - Q)^{-1}| = i(P, Q) = i(P, Q) + j(P, Q)$, completing the proof of (19) in the good reduction case.

We now turn to the case that $E/\mathbb{K}$ has multiplicative reduction. Denote by $\phi : \mathbb{K}^\times/q^{\mathbb{Z}} \xrightarrow{\sim} E(\mathbb{K})$ the Tate isomorphism. By periodicity we may select an element $u \in \mathbb{K}^\times \setminus \{1\}$ such that $\phi(u) = P - Q$ and $|q| < |u| \leq 1$. We are going to calculate the sum $i(P, Q) + j(P, Q)$ in three separate cases.

Case 1: $|q| < |u| < 1$. In this case, $r_\Sigma(P) \neq r_\Sigma(Q)$ by (12), whereby $i(P, Q) = 0$ by (17). It follows from the definition (12) of the retraction map r_Σ and from the definition (18) of $j(P, Q)$ that

$$(22) \qquad i(P, Q) + j(P, Q) = j(P, Q) = \tfrac{1}{2} \log|1/q| \Phi\left(\frac{\log|1/u|}{\log|1/q|}\right).$$

Case 2: $|u| = 1$ and $|u - 1| = 1$. In this case $r_\Sigma(P) = r_\Sigma(Q)$ by (12), whereby $j(P, Q) = \tfrac{1}{12} \log|1/q|$. As explained in Section 1.1, the condition $|u - 1| = 1$ means that $P - Q$ does not reduce to the identity element of $\overline{\mathscr{E}}_{\mathrm{ns}}(k)$, whereby $\pi(P) \neq \pi(Q)$. Thus $i(P, Q) = 0$ by (17), and so we have

$$(23) \qquad\qquad i(P, Q) + j(P, Q) = j(P, Q) = \tfrac{1}{12} \log|1/q|.$$

Case 3: $|u| = 1$ and $|u - 1| < 1$. We again have $r_\Sigma(P) = r_\Sigma(Q)$ by (12), and thus $j(P, Q) = \tfrac{1}{12} \log|1/q|$. Writing the point $P - Q$ in the affine coordinates $(x, y) \in \mathbb{K}^2$ associated to the Weierstrass equation (5), and letting $z = -x/y$, we have

$$(24) \qquad\qquad i(P, Q) = -\log|z| = -\log|u - 1|.$$

The first identity in (24) follows from the exact same calculation (21) used in the good reduction case. To see the second identity in (24), we note that the coordinates $x = X(u, q)$ and $y = Y(u, q)$ are defined via certain convergent series; see [Silverman 1994, Theorem V.3.1(c)]. Our assumptions that $|u| = 1$ and $|u - 1| < 1$, along with the ultrametric inequality, imply that $|x| = |X(u, q)| = |u - 1|^{-2}$ and $|y| = |Y(u, q)| = |u - 1|^{-3}$, whereby $|z| = |x|/|y| = |u - 1|$, and the second identity in (24) follows. We have shown that

$$(25) \qquad\qquad i(P, Q) + j(P, Q) = -\log|u - 1| + \tfrac{1}{12} \log|1/q|.$$

To complete the proof of (19) in the multiplicative reduction case, we observe that the right sides of the identities (22), (23), and (25) coincide with the formulas for $\lambda(P - Q)$ given in [Silverman 1994, Theorem VI.4.2(b)]. $\qquad\square$

1.7. *Connected metrized subgraphs of* $\mathsf{E} \setminus E(\mathbb{K})$ *containing* Σ.
Given a finite subset S of $\mathsf{E} \setminus (\Sigma \cup E(\mathbb{K}))$, there exists a unique smallest path-connected subset Γ of E containing $\Sigma \cup S$. Each set Γ so arising inherits from the path metric $\rho(x, y)$

the structure of a connected metrized graph, and the induced topology on each such subgraph Γ is closely related to the topology on E. These metrized subgraphs Γ of E are the basic analytic objects of study in this paper, and so in this section we will give a detailed description of their properties.

First, for each point b in $\mathsf{E} \setminus (\Sigma \cup E(\mathbb{K}))$, let $[r_\Sigma(b), b]$ denote the line segment in E from $r_\Sigma(b) \in \Sigma$ to b, and recall from Section 1.2 that $[r_\Sigma(b), b]$ is isometric (with respect to the path metric $\rho(x, y)$ on $\mathsf{E} \setminus E(\mathbb{K})$) to a closed real interval of length $\rho(r_\Sigma(b), b) > 0$. Given a (possibly empty) finite subset S of $\mathsf{E} \setminus (\Sigma \cup E(\mathbb{K}))$, define a subset Γ of E by

$$(26) \qquad \Gamma = \Sigma \cup \bigcup_{b \in S} [r_\Sigma(b), b]$$

The path metric $\rho(x, y)$ endows the set Γ with the structure of a connected metrized graph, and in particular Γ carries a compact Hausdorff topology. Define a map $r_\Gamma : \mathsf{E} \to \Gamma$ as follows. Set $r_\Gamma(x) = x$ if $x \in \Gamma$, but if $x \notin \Gamma$, let $\mathsf{B}^\circ \subset \mathsf{E} \setminus \Sigma$ denote the open Berkovich disc containing x, as in either (11) or (15). Define $r_\Gamma(x)$ to be the unique smallest element in $[r_\Sigma(x), x]$ (with respect to the partial order $\leq$ on B°) that is contained in Γ. We note for future use that if x, y are two points in the same connected component B° of $\mathsf{E} \setminus \Sigma$, then

$$(27) \qquad r_\Gamma(x) \vee r_\Gamma(y) = r_\Gamma(x \vee y),$$

where $\vee$ denotes the least upper bound with respect to the partial order $\leq$ on B°.

Let U be a subset of Γ that is open with respect to the metrized graph topology. Then the set $r_\Gamma^{-1}(U)$ is an open subset of E, and the collection of such sets $r_\Gamma^{-1}(U)$, over all connected metrized subgraphs Γ of $\mathsf{E} \setminus E(\mathbb{K})$ containing Σ and all open subsets U of Γ, form a base of open sets for the topology on E. Under this topology E is a path-connected, compact, Hausdorff space, the map $r_\Gamma : \mathsf{E} \to \Gamma$ is a strong deformation retraction, and $E(\mathbb{K})$ is a dense subspace of E (see [Berkovich 1990, Chapter 4] and [Baker 2008, Section 5]).

We do not claim that the sets $[r_\Sigma(b), b]$ ($b \in S$) in the union (26) are pairwise disjoint; any one of them may be contained in another, or any two of them may coincide on some initial subsegment. On the other hand, we may write any such connected subgraph Γ of $\mathsf{E} \setminus E(\mathbb{K})$ containing Σ as a disjoint union

$$(28) \qquad \Gamma = \Sigma \sqcup (a_1, b_1] \sqcup \cdots \sqcup (a_M, b_M],$$

where each $(a_m, b_m]$ is a subset of E that is isometric to a half-open real interval of length $\rho(a_m, b_m) > 0$. Recall that the length of the skeleton is $\ell(\Sigma) = \log^+ |j_E|$, and thus the length of Γ is $\ell(\Gamma) = \log^+ |j_E| + \sum_{m=1}^{M} \rho(a_m, b_m)$. Thus $\ell(\Gamma) = 0$ if and only if $E/\mathbb{K}$ has good reduction and $\Gamma = \Sigma$.

Assuming for the rest of this section that $\ell(\Gamma) > 0$, denote by m_Γ the uniform measure on Γ with respect to the metric $\rho(x, y)$. Thus if $F : \Gamma \to \mathbb{R}$ is m_Γ-integrable we have,

$$\int_\Gamma F(x) dm_\Gamma(x) = \int_\Sigma F(x) dm_\Sigma(x) + \sum_{m=1}^{M} \int_{(a_m, b_m]} F(x) dx,$$

where the first term on the right side is present only in the multiplicative reduction case, and dx is the uniform (Lebesgue) measure on $(a_m, b_m]$ for each $1 \le m \le M$. Denote by $L^2(\Gamma)$ the space of square m_Γ-integrable, real-valued functions on Γ, with norm $\|F\|_{L^2(\Gamma)} = \{\int_\Gamma |F|^2 dm_\Gamma\}^{1/2}$.

1.8. *The discreteness of torsion.* We conclude this preliminary section with a lemma. We will only need it for a global application in Section 4, but as it is purely local we include it here.

Lemma 8. *The torsion subgroup $E(\mathbb{K})_{\mathrm{tor}}$ is discrete in $E(\mathbb{K})$.*

This is well known, at least in the characteristic-zero case where it follows at once from properties of the formal logarithm. The following proof holds in arbitrary characteristic and was pointed out to us by Felipe Voloch.

Proof of Lemma 8. By translation it suffices to show that the origin $O \in E(\mathbb{K})$ has a torsion-free open neighborhood U. Let B° be the kernel of reduction associated to an integral Weierstrass equation for E, let $z : B^\circ \to \hat{E}(\mathbb{M})$ be the isomorphism of Lemma 6, where $\hat{E}(\mathbb{M})$ denotes the formal group over $\mathbb{M}$ with formal group law $F(Z_1, Z_2) \in \mathbb{O}[[Z_1, Z_2]]$ associated to the Weierstrass equation.

If the characteristic of the residue field k is zero, then $\hat{E}(\mathbb{M})$ is torsion free [Silverman 1986, IV.3.2(b)], and so we may take $U = B^\circ$ itself. Thus we may assume $\mathrm{char}(k) = p \ne 0$; then $\hat{E}(\mathbb{M})$ has only p-torsion [Silverman 1986, IV.3.2(b)]. The multiplication-by-p map $[p] : \hat{E}(\mathbb{M}) \to \hat{E}(\mathbb{M})$ is given by $[p](z) = pz + \cdots$ if $\mathrm{char}(\mathbb{K}) = 0$, and by $[p](z) = az^p + \cdots$ for some $a \in \mathbb{O}$ if $\mathrm{char}(\mathbb{K}) = p$. In either case it is easy to see that $[p](B(0, R)) \subseteq B(0, R)$ for all sufficiently small $R > 0$, where $B(0, R) = \{z \in \mathbb{M} \mid |z| \le R\}$. Since the set $E[p]$ of p-torsion on $E(\mathbb{K})$ is finite, if we take R small enough we may assume that $[p](B(0, R)) \subseteq B(0, R)$ and that $B(0, R)$ contains no p-torsion other than 0.

Suppose that $B(0, R)$ contains a point z of exact order p^r with $r \ge 2$. Then $[p^{r-1}](z)$ is nonzero, has order p, and is in $B(0, R)$ since $[p](B(0, R)) \subseteq B(0, R)$. This contradicts the fact that $B(0, R)$ contains no p-torsion other than 0. Thus $B(0, R)$ contains no nonzero p-power torsion, so it is torsion-free. Let U be the pullback of $B(0, R)$ under the isomorphism $z : B^\circ \to \hat{E}(\mathbb{M})$. $\square$

2. The main local inequality

Overview. We continue with the notation of the last section. In this section we will define the canonical measure μ on E, and we will prove the main local result of this paper, which is a bound on the error term when testing the μ-equidistribution of a set Z against a certain class of test functions on E.

2.1. *The space $\mathscr{S}(\mathsf{E}, \mathbb{R})$ of test functions.* Let $\mathscr{C}(\mathsf{E}, \mathbb{R})$ denote the Banach space of continuous real-valued functions on E, equipped with the supremum norm. We will now define a dense subspace $\mathscr{S}(\mathsf{E}, \mathbb{R})$ of $\mathscr{C}(\mathsf{E}, \mathbb{R})$ that will serve as our space of test functions.

First, let Γ be a connected metrized subgraph of $\mathsf{E} \setminus E(\mathbb{K})$ containing Σ, and let $F : \Gamma \to \mathbb{R}$ be a function. Given $x \in (r_\Sigma(b), b)$, where Γ is written as in (26), define the derivative $F'(x)$ of F at x via the usual difference quotient limit, assuming it exists, by identifying $(r_\Sigma(b), b)$ with an open real interval. If $E/\mathbb{K}$ has multiplicative reduction and $x \in \Sigma$, define $F'(x)$ similarly by identifying $\Sigma \simeq \mathbb{R}/(\log|j_E|)\mathbb{Z}$. Thus the derivative $F'(x)$, assuming it exists, has an unambiguous meaning for m_Γ-almost all $x \in \Gamma$. Define $\mathscr{S}_\Gamma(\mathsf{E}, \mathbb{R})$ to be the space of functions $F : \mathsf{E} \to \mathbb{R}$ satisfying the following conditions:

- F factors through the retraction map $r_\Gamma : \mathsf{E} \to \Gamma$; thus $F = F \circ r_\Gamma$.

- F is continuous on Γ with respect to its metrized graph topology.

- The derivative F' exists m_Γ-almost everywhere on Γ, and $F' \in L^2(\Gamma)$.

We call $\mathscr{S}_\Gamma(\mathsf{E}, \mathbb{R})$ the space of test functions associated to Γ.

Now define $\mathscr{S}(\mathsf{E}, \mathbb{R}) = \bigcup_\Gamma \mathscr{S}_\Gamma(\mathsf{E}, \mathbb{R})$, the union over all connected metrized subgraphs Γ of $\mathsf{E} \setminus E(\mathbb{K})$ containing Σ. It follows from the definition of the topology on E that $\mathscr{S}(\mathsf{E}, \mathbb{R}) \subset \mathscr{C}(\mathsf{E}, \mathbb{R})$, and it is straightforward to show using the Stone–Weierstrass theorem that $\mathscr{S}(\mathsf{E}, \mathbb{R})$ is dense in $\mathscr{C}(\mathsf{E}, \mathbb{R})$.

2.2. *The canonical measure and the local discrepancy.* The canonical measure μ on E is the unique positive unit Borel measure that is supported on the skeleton Σ, and that when restricted to Σ is given by

$$\mu = \begin{cases} \text{the Dirac measure at the point } \Sigma & \text{if } |j_E| \leq 1, \\ \text{the normalized uniform measure } (\log|j_E|)^{-1} m_\Sigma & \text{if } |j_E| > 1. \end{cases}$$

Let $Z = \{P_1, \ldots, P_N\}$ be a set of N distinct points in $E(\mathbb{K})$. The *local discrepancy* of the set Z is defined by

$$(29) \qquad D(Z) = \frac{1}{N^2} \sum_{\substack{1 \leq m,n \leq N \\ m \neq n}} \lambda(P_m - P_n) + \frac{1}{12N} \log^+|j_E|.$$

By the decomposition (19) and the fact that $j(x, x) = \frac{1}{12} \log^+ |j_E|$ for all $x \in \mathsf{E}$, we also have

$$(30) \qquad D(Z) = \frac{1}{N^2} \sum_{\substack{1 \le m,n \le N \\ m \ne n}} i(P_m, P_n) + \frac{1}{N^2} \sum_{1 \le m,n \le N} j(P_m, P_n).$$

We shall see that $D(Z)$ is nonnegative, and that it gives a quantitative measure of the μ-equidistribution of the set Z. This notion of local discrepancy was introduced in [Baker and Petsche 2005] to study equidistribution on elliptic curves over number fields.

2.3. *Fourier analysis on* $\mathbb{R}/\ell\mathbb{Z}$. For the proof of our main equidistribution result, and again for one of our global applications in Section 4, we will use Fourier analysis on the circle group $\mathbb{R}/\ell\mathbb{Z}$, where $\ell > 0$. In this section we will recall the basic theory and prove a lemma.

Given a Lebesgue integrable function $F : \mathbb{R}/\ell\mathbb{Z} \to \mathbb{R}$, its k-th Fourier coefficient is defined by $\hat{F}(k) = \ell^{-1} \int_0^\ell F(x) e^{-2\pi i k x/\ell} dx$, and its Fourier series is given by $F(x) \sim \sum_{k \in \mathbb{Z}} \hat{F}(k) e^{2\pi i k x/\ell}$. If $\sum_{k \in \mathbb{Z}} |\hat{F}(k)| < +\infty$, then the right side converges absolutely and the symbol "$\sim$" can be interpreted as an equality. Assuming F is square integrable on $\mathbb{R}/\ell\mathbb{Z}$, Parseval's formula $\ell^{-1} \int_0^\ell |F(x)|^2 dx = \sum_{k \in \mathbb{Z}} |\hat{F}(k)|^2$ holds. The convolution $F * G : \mathbb{R}/\ell\mathbb{Z} \to \mathbb{R}$ of two functions $F, G : \mathbb{R}/\ell\mathbb{Z} \to \mathbb{R}$ is defined by $F * G(x) = \ell^{-1} \int_0^\ell F(y) G(x - y) dy$ and has Fourier coefficients $\widehat{F * G}(k) = \hat{F}(k) \hat{G}(k)$.

Let $\Psi(x) = x - [x] - 1/2$ and $\Phi(x) = (x - [x])^2 - (x - [x]) + 1/6$ be the first and second $\mathbb{Z}$-periodic Bernoulli polynomials respectively, and let $\Psi_\ell(x) = \Psi(x/\ell)$ and $\Phi_\ell(x) = \Phi(x/\ell)$ denote their ℓ-periodic analogues. Then $\hat{\Psi}_\ell(0) = \hat{\Phi}_\ell(0) = 0$, while $\hat{\Psi}_\ell(k) = -1/(2\pi i k)$ and $\hat{\Phi}_\ell(k) = 1/(2\pi^2 k^2)$ for $k \ne 0$; thus Φ has an absolutely convergent Fourier series but Ψ does not.

Lemma 9. *Suppose that $F : \mathbb{R}/\ell\mathbb{Z} \to \mathbb{R}$ is continuous and differentiable almost everywhere, and that F' is Lebesgue integrable. Given N (possibly indistinct) points $p_1, \ldots, p_N \in \mathbb{R}/\ell\mathbb{Z}$, define $G : \mathbb{R}/\ell\mathbb{Z} \to \mathbb{R}$, $x \mapsto -N^{-1} \sum_{n=1}^N \Psi_\ell(p_n - x)$. Then*

$$(31) \qquad \frac{1}{N} \sum_{n=1}^N F(p_n) - \frac{1}{\ell} \int_0^\ell F(x) dx = \int_0^\ell F'(x) G(x) dx$$

and

$$(32) \qquad \int_0^\ell |G(x)|^2 dx = \frac{\ell}{2N^2} \sum_{1 \le i,j \le N} \Phi_\ell(p_i - p_j).$$

Proof. We have $F(x) - \hat{F}(0) = -\ell F' * \Psi_\ell(x)$ for all $x \in \mathbb{R}/\ell\mathbb{Z}$; this follows from integration by parts, or alternatively by noting that $\hat{F}'(k) = (2\pi ik/\ell)\hat{F}(k)$ and comparing the Fourier coefficients of both sides. Thus

$$\frac{1}{N}\sum_{n=1}^{N}(F(p_n) - \hat{F}(0)) = -\frac{\ell}{N}\sum_{n=1}^{N}F' * \Psi_\ell(p_n)$$

$$= \int_0^\ell F'(x)\Big(\frac{-1}{N}\sum_{n=1}^{N}\Psi_\ell(p_n - x)\Big)dx = \int_0^\ell F'(x)G(x)dx,$$

which implies (31). Note that

$$\hat{G}(k) = -\frac{1}{2\pi ik}\Big(\frac{1}{N}\sum_{n=1}^{N}e^{-2\pi ikp_n/\ell}\Big)$$

for $k \neq 0$ and $\hat{G}(0) = 0$, and thus by Parseval's formula, we get (32) by

$$\frac{1}{\ell}\int_0^\ell |G(x)|^2 dx = \sum_{k\neq 0}\frac{1}{4\pi^2 k^2}\Big|\frac{1}{N}\sum_{n=1}^{N}e^{-2\pi ikp_n/\ell}\Big|^2$$

$$= \frac{1}{2N^2}\sum_{1\leq i,j\leq N}\sum_{k\neq 0}\frac{1}{2\pi^2 k^2}e^{2\pi ik(p_i - p_j)/\ell}$$

$$= \frac{1}{2N^2}\sum_{1\leq i,j\leq N}\Phi_\ell(p_i - p_j). \qquad \square$$

2.4. *The main local inequality.* We can now state the main local results of this paper. First, given a connected metrized subgraph Γ of $\mathsf{E} \setminus E(\mathbb{K})$ containing Σ, written as in (26), define

(33) $$\ell_0(\Gamma) = \max_{b\in S}\rho(r_\Sigma(b), b).$$

In other words, $\ell_0(\Gamma)$ is the greatest distance of any point $x \in \Gamma$ to the skeleton Σ; thus $\ell_0(\Gamma) = 0$ if and only if $\Gamma = \Sigma$.

Theorem 10. *Let $F \in \mathscr{S}_\Gamma(\mathsf{E}, \mathbb{R})$ be a test function associated to a connected metrized subgraph Γ of $\mathsf{E} \setminus E(\mathbb{K})$ containing Σ. Let Z be a nonempty finite subset of $E(\mathbb{K})$. Then*

(34) $$\Big|\frac{1}{|Z|}\sum_{P\in Z}F(P) - \int F d\mu\Big| \leq \|F'\|_{L^2(\Gamma)}\Big(D(Z) + \frac{\ell_0(\Gamma)}{|Z|}\Big)^{1/2}.$$

Proof. First, if Γ consists of a single point, F is constant since $F = F \circ r_\Gamma$ factors through Γ. In this case the left side of (34) vanishes, and as the right side is nonnegative, the theorem is trivial. So we may now assume that $\ell(\Gamma) > 0$, or

equivalently that either $E/\mathbb{K}$ has multiplicative reduction or that $\Gamma \setminus \Sigma$ is nonempty. Let $Z = \{P_1, \ldots, P_N\}$. We are going to define a piecewise continuous function $G : \Gamma \to \mathbb{R}$, depending on the graph Γ and the set Z but not otherwise on the function F, that satisfies the identity

$$(35) \qquad \frac{1}{N} \sum_{n=1}^{N} F(P_n) - \int F d\mu = \int_{\Gamma} F'(x) G(x) dm_{\Gamma}(x)$$

and the inequality

$$(36) \qquad \|G\|_{L^2(\Gamma)}^2 \leq D(Z) + \ell_0(\Gamma)/N.$$

The inequality (34) will then follow immediately from Cauchy's inequality.

We first define G on the skeleton Σ. This is only necessary in the multiplicative reduction case, since in the good reduction case Σ has m_{Γ}-measure zero. Supposing that $E/\mathbb{K}$ has multiplicative reduction, we may identify $\Sigma \simeq \mathbb{R}/\ell\mathbb{Z}$ as described in Section 1.4, where $\ell = \log|j_E| > 0$. Thus $\mu = (1/\ell)m_{\Sigma}$ is the normalized uniform measure on Σ. For each $1 \leq n \leq N$, let $p_n = r_{\Sigma}(P_n) \in \Sigma$, and define $G : \Sigma \to \mathbb{R}$ as in Lemma 9. Then

$$(37) \qquad \frac{1}{N} \sum_{n=1}^{N} F(r_{\Sigma}(P_n)) - \int F d\mu = \int_{\Sigma} F'(x) G(x) dm_{\Gamma}(x)$$

by (31), and since $j(P, Q)$ factors through the retraction map $r_{\Sigma} : \mathsf{E} \to \Sigma$, we have

$$(38) \qquad \begin{aligned} \int_{\Sigma} |G(x)|^2 dm_{\Gamma}(x) &= \frac{1}{N^2} \sum_{1 \leq m,n \leq N} j(r_{\Sigma}(P_m), r_{\Sigma}(P_n)) \\ &= \frac{1}{N^2} \sum_{1 \leq m,n \leq N} j(P_m, P_n) \end{aligned}$$

by (32). If $\Gamma = \Sigma$ then the proof of the theorem is complete, since (37) implies (35), and (36) follows from (38) and the fact that the right side of (38) is at most $D(Z)$ by (30). Note that (38) and the decomposition (30), along with the nonnegativity of $i(x, y)$, shows that local discrepancy $D(Z)$ is nonnegative.

We will now define G on $\Gamma \setminus \Sigma$, which we may assume is nonempty since the $\Gamma = \Sigma$ case has been settled. Define $G : \Gamma \setminus \Sigma \to \mathbb{R}$ by

$$G(x) = \frac{1}{N} \sum_{n=1}^{N} \chi_{[r_{\Sigma}(P_n), r_{\Gamma}(P_n)]}(x).$$

Thus $NG(x)$ counts the number of points P_n in Z such that x lies in the subsegment of Γ from $r_\Sigma(P_n)$ to $r_\Gamma(P_n)$. Since $r_\Gamma(P_n) = r_\Sigma(P_n)$ if $r_\Gamma(P_n) \in \Sigma$, we have

$$(39) \qquad \frac{1}{N} \sum_{n=1}^{N} F(r_\Gamma(P_n)) - \frac{1}{N} \sum_{n=1}^{N} F(r_\Sigma(P_n)) = \int_{\Gamma \setminus \Sigma} F'(x) G(x) dm_\Gamma(x)$$

by elementary calculus.

For each n, let us abbreviate $\sigma_n = r_\Sigma(P_n) \in \Sigma$ and $\gamma_n = r_\Gamma(P_n) \in \Gamma$. Given two (possibly equal) points P_m and P_n in Z, if they are elements of two distinct connected components of $\mathsf{E} \setminus \Sigma$, then the intersection $[\sigma_m, \gamma_m] \cap [\sigma_n, \gamma_n]$ either is the empty set or contains only the single point $\sigma = \sigma_m = \sigma_n$. In either case the intersection has m_Γ-measure zero, and $i(P_m, P_n) = 0$. On the other hand, suppose that P_m and P_n are elements of the same connected component B° of $\mathsf{E} \setminus \Sigma$. Then $\sigma_m = \sigma_n \, (= \sigma)$ and

$$[\sigma_m, \gamma_m] \cap [\sigma_n, \gamma_n] = [\sigma, \gamma_m \vee \gamma_n] = [\sigma, r_\Gamma(P_m \vee P_n)] \qquad \text{(by (27))}$$
$$\subseteq [\sigma, P_m \vee P_n].$$

Recall that $\rho(\sigma, P_m \vee P_n) = i(P_m, P_n)$ by definition for $m \neq n$. When $m = n$, we use the inequality $\rho(\sigma, r_\Gamma(P_m \vee P_n)) \leq \ell_0(\Gamma)$ by the definition (33) of $\ell_0(\Gamma)$. Summing over all pairs, we have

$$\int_{\Gamma \setminus \Sigma} |G(x)|^2 dm_\Gamma(x) = \frac{1}{N^2} \sum_{1 \leq m, n \leq N} \int_{\Gamma \setminus \Sigma} \chi_{[\sigma_m, \gamma_m]}(x) \chi_{[\sigma_n, \gamma_n]}(x) dm_\Gamma(x)$$

$$(40) \qquad\qquad = \frac{1}{N^2} \sum_{1 \leq m, n \leq N} \int_{\Gamma \setminus \Sigma} \chi_{[\sigma_m, \gamma_m] \cap [\sigma_n, \gamma_n]}(x) dm_\Gamma(x)$$

$$\leq \frac{1}{N^2} \sum_{m \neq n} i(P_m, P_n) + \frac{\ell_0(\Gamma)}{N}.$$

The identities (37) and (39) imply (35), the identities (30) and (38) along with the inequality (40) imply (36), and Cauchy's inequality completes the proof of (34). $\square$

3. The global equidistribution theorem

3.1. *Global function fields.* Let K be a field having a set M_K of inequivalent, nontrivial, discrete valuations on it such that

- $\{v \in M_K \mid v(a) \neq 0\}$ is finite for each $a \in K^\times$;
- $\sum_{v \in M_K} v(a) = 0$ for each $a \in K^\times$ (product formula);
- $[L : K] = \sum_{w \mid v} [L_w : K_v]$ for each finite extension L/K and each $v \in M_K$ (local degree formula).

The sum in the local degree formula is over all valuations w on L whose restriction to K coincides with $v \in M_K$ (in this case we write $w | v$), and K_v (respectively L_w) denotes the completion of K at v (respectively L at w). We refer to the elements of M_K as *places* of K, and instead of the additive valuation v, we will usually use the corresponding nonarchimedean absolute value $|\cdot|_v = e^{-v(\cdot)}$. We will call a field K satisfying these axioms a *global function field*. The most commonly encountered example is given by the field $K = k(C)$ of rational functions on an integral, proper, geometrically connected curve C defined over an arbitrary field k; in this case the places of K correspond to the scheme-theoretic closed points of C. There are other examples, however; see [Bombieri and Gubler 2006, Section 1.4.6] and [Lang 1983, Section 2.3].

The local degree formula always holds for separable finite extensions L/K [Bombieri and Gubler 2006, Corollary 1.3.2]. The referee pointed out that even when L/K is inseparable and the local degree formula fails, a suitable product formula can be recovered for L, provided one renormalizes the valuations $w \in M_L$ appropriately. This is worked out in [Bombieri and Gubler 2006, Section 1.4] and [Serre 1989, Section 2.1].

We will now summarize the basic properties of global function fields; for detailed proofs see [Lang 1983]. First, it follows from the local degree formula that the product formula holds in the form $\prod_{w \in M_L} |a|_w^{[L_w:K_w]} = 1$ for each finite extension L/K, where M_L denotes the set of absolute values on L restricting to an absolute value in M_K.

Given a finite extension L/K, define the *constant field* of L by

$$L_0 = \{a \in L \mid |a|_w \leq 1 \text{ for all } w \in M_L\}.$$

It follows from the ultrametric inequality and the product formula that L_0 is a subfield of L, and that $|a|_w = 1$ for all $a \in L_0^\times$ and all $w \in M_L$. Let $\overline{K}$ be the algebraic closure of K, and let $\overline{K}_0 = \bigcup_{L/K} L_0$ denote the constant field of $\overline{K}$.

For each $v \in M_K$, let $\mathbb{K}_v$ denote the completion of the algebraic closure $\overline{K}_v$ of K_v; the absolute value $|\cdot|_v$ extends uniquely to the field $\mathbb{K}_v$, which is complete and algebraically closed [Bosch et al. 1984, Section 3.4]. Given a finite extension L/K and a place $v \in M_K$, each K-embedding $\epsilon : L \hookrightarrow \mathbb{K}_v$ determines a valuation $w \in M_L$ over v by $|\cdot|_w = |\epsilon(\cdot)|_v$, and conversely every place $w \in M_L$ arises in this way. If $\epsilon, \delta : L \hookrightarrow \mathbb{K}_v$ are two such K-embeddings, they determine the same element $w \in M_L$ if and only if $\epsilon = \sigma \circ \delta$ for some σ in the group $\mathrm{Aut}(\mathbb{K}_v/K_v)$ of continuous K_v-automorphisms of $\mathbb{K}_v$.

Finally, suppose that L/K is a finite normal extension, and suppose that $\epsilon, \delta : L \hookrightarrow \mathbb{K}_v$ are two K-embeddings. Then $\epsilon(L) = \delta(L)$, and therefore $\epsilon = \delta \circ \alpha$ for some $\alpha \in \mathrm{Aut}(L/K)$.

3.2. *The absolute Weil height.* The absolute Weil height $h : \mathbb{P}^1(\overline{K}) \to \mathbb{R}$ is defined as follows. Given a point $x = (x_0 : x_1) \in \mathbb{P}^1(\overline{K})$, select a finite extension L/K containing the coordinates of x, and set

$$(41) \qquad h(x) = \sum_{w \in M_L} \frac{[L_w : K_w]}{[L : K]} \log \max\{|x_0|_w, |x_1|_w\}.$$

By the product formula and the local degree formula, the value of $h(x)$ does not depend on the choice of coordinates or the choice of L containing them. We also denote by $h : \overline{K} \to \mathbb{R}$ the affine height given by $h(x) = h(1 : x)$ for $x \in \overline{K}$. It is easy to check that $h(x) \geq 0$ for all $x \in \overline{K}$, with $h(x) = 0$ if and only if $x \in \overline{K}_0$.

3.3. *Elliptic curves and the Néron–Tate height.* Let E/K be an elliptic curve over K. The Néron–Tate height function $\hat{h} : E(\overline{K}) \to \mathbb{R}$ is defined by

$$\hat{h}(P) = \tfrac{1}{2} \lim_{n \to +\infty} \frac{h(x(nP))}{n^2},$$

where $x : E(\overline{K}) \to \mathbb{P}^1(\overline{K})$ is the x-coordinate function associated to a Weierstrass equation for E/K. For the proofs of the following basic facts about $\hat{h}$ are found in [Silverman 1986, Section VIII.9]; strictly speaking this text treats the number field case only, although the arguments work in the function field setting. For a general treatment of Néron–Tate heights on abelian varieties over arbitrary global fields, see [Bombieri and Gubler 2006, Section 9.2].

The Néron–Tate height function satisfies $\hat{h}(P) = (1/2)h(x(P)) + O(1)$ and $\hat{h}(nP) = n^2\hat{h}(P)$ for all $P \in E(\overline{K})$, and in fact $\hat{h}$ can be characterized as the unique function satisfying these two properties. The definition of $\hat{h}$ is independent of the x-coordinate function used to define it.

Using these properties of $\hat{h}$, we see that $\hat{h}(P) \geq 0$ for all $P \in E(\overline{K})$, and

$$(42) \qquad P \in E(\overline{K})_{\mathrm{tor}} \quad \text{implies} \quad \hat{h}(P) = 0.$$

However, unlike in the number field case, the converse of (42) may fail. For example, if E is defined over the constant field K_0 of K, then every point in $E(\overline{K}_0)$ has canonical height zero, and $E(\overline{K}_0)$ may contain nontorsion points (if K_0 is infinite). See Lang [Lang 1983, Section 6.5] and Baker [Baker 2009, Appendix B] for results that imply the converse to (42) under certain conditions on E/K.

Finally, we note for future use a fundamental property of the Néron–Tate height, the parallelogram law

$$(43) \qquad \hat{h}(P + Q) + \hat{h}(P - Q) = 2\hat{h}(P) + 2\hat{h}(Q) \quad \text{for } P, Q \in E(\overline{K}).$$

3.4. *Local height functions.* For each place $v \in M_K$, we may consider E as an elliptic curve over $\mathbb{K}_v$, and we generally affix a subscript v to each local object

associated to $E/\mathbb{K}_v$ introduced in Sections 1 and 2. Thus E_v denotes the Berkovich analytic space associated to $E/\mathbb{K}_v$ and $\lambda_v : E(\mathbb{K}_v) \setminus \{O\} \to \mathbb{R}$ is the Néron local height function discussed in Section 1.6; likewise for i_v, j_v, Σ_v, μ_v, etc.

Let L/K be a finite extension. For each $w \in M_L$, let $\lambda_w : E(L_w) \setminus \{O\} \to \mathbb{R}$ denote the associated Néron local height function. By the uniqueness of such functions, if the place w corresponds to the K-embedding $\epsilon : L \hookrightarrow \mathbb{K}_v$, then $\lambda_w(\cdot) = \lambda_v(\epsilon(\cdot))$.

Given a global point $P \in E(L) \setminus \{O\}$, we have a decomposition

$$(44) \qquad \hat{h}(P) = \sum_{w \in M_L} \frac{[L_w : K_w]}{[L : K]} \lambda_w(P)$$

of the global height into local terms. For the proof of (44) see [Silverman 1994, Section VI.2] (the proof is written for number fields but works in this setting) or [Call and Silverman 1993].

3.5. *The global equidistribution theorem.* In this section we will prove the quantitative global equidistribution result Theorem 2 along with its qualitative corollary Corollary 3. Recall that the height $\hat{h}(Z)$ of a nonempty finite subset Z of $E(\overline{K})$ is defined to be the average $\hat{h}(Z) = |Z|^{-1} \sum_{P \in Z} \hat{h}(P)$ of the height of its points. The following is a more precise statement of Theorem 2.

Theorem 11. *Let E/K be an elliptic curve over a global function field, let $v \in M_K$ be a place of K, and let $\epsilon : \overline{K} \hookrightarrow \mathbb{K}_v$ be a K-embedding. Let Z be a nonempty finite $\mathrm{Aut}(\overline{K}/K)$-stable subset of $E(\overline{K})$, and let $F \in \mathscr{S}(\mathsf{E}_v, \mathbb{R})$ be a test function on E_v with associated connected metrized subgraph Γ containing Σ_v. Then*

$$\left| \frac{1}{|Z|} \sum_{P \in Z} F(\epsilon(P)) - \int F \, d\mu_v \right| \le \|F'\|_{L^2(\Gamma)} \left(4\hat{h}(Z) + \frac{h(j_E)}{12|Z|} + \frac{\ell_0(\Gamma)}{|Z|} \right)^{1/2}.$$

Proof. Let L/K be a finite normal extension such that $Z = \{P_1, \ldots, P_N\} \subseteq E(L)$. Given a place $w \in M_L$ corresponding to some K-embedding $\delta : L \hookrightarrow \mathbb{K}_u$, where $u \in M_K$ is the place below w, define $D_w(Z) = D_u(\delta(Z))$, where D_u denotes the local discrepancy defined on subsets of $E(\mathbb{K}_u)$, as in Section 2.2.

We are going to bound the sum $D = \sum_{w \in M_L} ([L_w : K_w]/[L : K]) D_w(Z)$ from above globally and from below locally. First the global bound:

$$D = \sum_{w \in M_L} \frac{[L_w : K_w]}{[L : K]} \left(\frac{1}{N^2} \sum_{m \ne n} \lambda_w(P_m - P_n) + \frac{\log^+|j_E|_w}{12N} \right)$$

$$(45) \qquad = \frac{1}{N^2} \sum_{1 \le m,n \le N} \hat{h}(P_m - P_n) + \frac{h(j_E)}{12N}$$

$$\le \frac{1}{N^2} \sum_{1 \le m,n \le N} (2\hat{h}(P_m) + 2\hat{h}(P_n)) + \frac{h(j_E)}{12N} = 4\hat{h}(Z) + \frac{h(j_E)}{12N}$$

by the decompositions (41) and (44), the parallelogram law, and the nonnegativity of the height $\hat{h}$. At places $w \in M_L$ with $w \nmid v$, we use the trivial lower bound $D_w(Z) \geq 0$ stemming from the fact that the local discrepancy (29) is nonnegative.

Finally, we treat the places $w \in M_L$ such that $w | v$. We first observe that $D_w(Z) = D_v(\epsilon(Z))$ for all $w | v$, where $\epsilon : L \hookrightarrow \mathbb{K}_v$ is the fixed K-embedding in the statement of the theorem. (Proof: Any $w | v$ is induced by *some* K-embedding $\delta : L \hookrightarrow \mathbb{K}_v$, and since L/K is normal, $\delta = \epsilon \circ \alpha$ for some $\alpha \in \mathrm{Aut}(L/K)$. By hypothesis Z is $\mathrm{Aut}(L/K)$-invariant.) We deduce from the local degree formula that $\sum_{w|v}([L_w : K_v]/[L : K])D_w(Z) = D_v(\epsilon(Z))$, and assembling (45) with the local considerations we have $D_v(\epsilon(Z)) \leq 4\hat{h}(Z) + h(j_E)/12N$. The inequality stated in Theorem 11 now follows from (34).   □

Proof of Corollary 3. Let $F : \mathsf{E}_v \to \mathbb{R}$ be a continuous function, and let $\theta > 0$. Since $\mathscr{S}(\mathsf{E}_v, \mathbb{R})$ is dense in $\mathscr{C}(\mathsf{E}_v, \mathbb{R})$ (with respect to the supremum norm), there exists a function $F^* \in \mathscr{S}(\mathsf{E}_v, \mathbb{R})$ such that $|F(x) - F^*(x)| \leq \theta$ for all $x \in \mathsf{E}_v$. Since $\hat{h}(Z_n) \to 0$ and $|Z_n| \to +\infty$, it follows from Theorem 11 that the desired limit formula (1) holds with F replaced by F^*. Using the triangle inequality, we have

$$\left| \frac{1}{|Z_n|} \sum_{P \in Z_n} F(\epsilon(P)) - \int F d\mu_v \right| \leq \left| \frac{1}{|Z_n|} \sum_{P \in Z_n} (F(\epsilon(P)) - F^*(\epsilon(P))) \right|$$

$$+ \left| \frac{1}{|Z_n|} \sum_{P \in Z_n} F^*(\epsilon(P)) - \int F^* d\mu_v \right|$$

$$+ \left| \int (F^* - F) d\mu_v \right|,$$

and since $\int 1 d\mu_v = 1$ and $|F(x) - F^*(x)| \leq \theta$ we deduce that

$$(46) \qquad \limsup_{n \to +\infty} \left| \frac{1}{|Z_n|} \sum_{P \in Z_n} F(\epsilon(P)) - \int F d\mu_v \right| \leq 2\theta.$$

Since $\theta > 0$ is arbitrary, the left side of (46) is zero, completing the proof of (1). □

4. The finiteness of S-integral torsion points

Introduction. In this section we will prove Theorem 5. For convenience we will restate the theorem below. Let K be a global function field and let E/K be an elliptic curve. Given a place v of K and a point $R \in E(\mathbb{K}_v)$, define

$$(47) \qquad M_v(R) = \inf_{P \in E(\mathbb{K}_v)_{\mathrm{tor}}} d_v(P, R),$$

where d_v is the v-adic metric on $E(\mathbb{K}_v)$ defined in Section 1.1. Thus $M_v(R)$ measures the distance between R and its nearest torsion point in $E(\mathbb{K}_v)$. Plainly

$M_v(R) = 0$ if R is a torsion point, but if R is nontorsion, then the discreteness of $E(\mathbb{K}_v)_{\text{tor}}$ (by Lemma 8) and the definition of the metric d_v shows that $0 < M_v(R) \leq 1$. If $Q \in E(\overline{K})$ is a global point, define

$$(48) \qquad M_v(Q) = \inf_\epsilon M_v(\epsilon(Q)) \quad \text{and} \quad m_v(Q) = -\log M_v(Q),$$

the first being the infimum over all K-embeddings $\epsilon : \overline{K} \hookrightarrow \mathbb{K}_v$; clearly we have $0 < M_v(Q) \leq 1$ if Q is nontorsion, since Q has only finitely many embeddings into $E(\mathbb{K}_v)$. Thus $m_v(Q)$ is nonnegative, and finite provided Q is nontorsion.

Theorem 12. *Let E/K be an elliptic curve over a global function field. Let (2) be a Weierstrass equation for E/K, let S be a finite set of places of K such that (2) is defined over $\mathbb{O}_S$, and such that $|\Delta|_v = 1$ for all $v \notin S$, where $\Delta \in K$ is the discriminant of (2). Let $\mathscr{E}/\mathbb{O}_S$ denote the associated S-integral Weierstrass model for E. If Q is a point in $E(\overline{K})$ with $\hat{h}(Q) > 0$, then*

$$(49) \qquad |\mathscr{E}(\overline{\mathbb{O}}_S, Q)_{\text{tor}}| \leq \frac{1}{\hat{h}(Q)^2} \left(\frac{|S|h(j_E)}{12} + \sum_{v \in S} m_v(Q) \right)^2.$$

Our proof follows the strategy used by Baker, Ih and Rumely [Baker et al. 2008] in the number field case, with some slight differences. The idea is to use the equidistribution theorem at each place $v \in S$, testing a large set of global torsion points against truncated Néron local height functions.

4.1. *Truncated Néron local height functions.* Fix a place v of K. In order to prove Theorem 12 we will use certain test functions in the space $\mathscr{S}(\mathsf{E}_v, \mathbb{R})$ that are formed by truncating linear combinations of Néron local height functions. The work in this section is entirely local, and thus we suppress the dependence on the place v in all of our notation.

Let $\mathscr{R}$ be a nonempty finite subset of $E(\mathbb{K})$, and define $F : \mathsf{E} \setminus \mathscr{R} \to \mathbb{R}$ by

$$F(x) = \frac{1}{|\mathscr{R}|} \sum_{R \in \mathscr{R}} \lambda(x - R).$$

Note that F is continuous on $\mathsf{E} \setminus \mathscr{R}$ but has logarithmic singularities at the points of $\mathscr{R}$. However, fix a constant $m > 0$ and define the truncated function $G : \mathsf{E} \to \mathbb{R}$ by

$$G(x) = \frac{1}{|\mathscr{R}|} \sum_{R \in \mathscr{R}} (j(x, R) + i^*(x, R)),$$

where $i^*(x, R) = \min\{i(x, R), m\}$. For each $R \in \mathscr{R}$, let κ_R be the unique point in the infinitely long line segment $[r_\Sigma(R), R]$ in E that satisfies $\rho(r_\Sigma(R), \kappa_R) = m$. Let

$$\Gamma = \Sigma \cup \left(\bigcup_{R \in \mathscr{R}} [r_\Sigma(R), \kappa_R] \right)$$

be the smallest connected metrized subgraph of $E \setminus E(\mathbb{K})$ that contains Σ and all of the κ_R.

Lemma 13. *The function* $G : E \to \mathbb{R}$ *is in the space* $\mathscr{S}_\Gamma(E, \mathbb{R})$ *of test functions associated to* Γ, $\ell_0(\Gamma) = m$, *and*

$$(50) \qquad \|G'\|^2_{L^2(\Gamma)} \le \tfrac{1}{12} \log^+ |j_E| + m.$$

Proof. To show that $G \in \mathscr{S}_\Gamma(E, \mathbb{R})$, we must show that $G = G \circ r_\Gamma$, where $r_\Gamma : E \to \Gamma$ is the retraction map. Fix a point $x \in E$; we are going to show that for each $R \in \mathscr{R}$,

$$(51) \qquad i^*(x, R) = i^*(r_\Gamma(x), R).$$

Since the function $j(x, y)$ factors through the retraction map $r_\Sigma : E \to \Sigma$, and therefore the map $r_\Gamma : E \to \Gamma$ as well, this will complete the proof that $G = G \circ r_\Gamma$.

First, suppose that x and R are elements of the same connected component B° of $E \setminus \Sigma$ as in the decomposition (11) or (15). Then set $x_R = x \vee R$, the least upper bound of x and R with respect to the partial order on B°. If $x_R \in [r_\Sigma(R), \kappa_R]$, then $i(x, R) = \rho(r_\Sigma(x_R), x_R) \le \rho(r_\Sigma(x_R), \kappa_R) = m$, and since $x_R = x \vee R = r_\Gamma(x) \vee R$, we have $i(x, R) = i(r_\Gamma(x), R)$. Since both of these quantities are at most m, we deduce (51) in this case. On the other hand, suppose that $x_R \notin [r_\Sigma(R), \kappa_R]$. Then by a similar argument as above, both $i(x, R)$ and $i(r_\Gamma(x), R)$ are greater than m, whereby both sides of (51) are equal to m. Finally, if x and R are *not* both elements of the same connected component B° of $E \setminus \Sigma$, then both sides of (51) vanish; the proof that $G = G \circ r_\Gamma$ is complete. The other axioms satisfied by functions in $\mathscr{S}_\Gamma(E, \mathbb{R})$ are trivially verified for G. The equality $\ell_0(\Gamma) = m$ follows at once from the fact that $\rho(r_\Sigma(R), \kappa_R) = m$ for all $R \in \mathscr{R}$.

Finally, we are going to show that

$$(52) \qquad \|G'\|^2_{L^2(\Sigma)} \le \tfrac{1}{12} \log^+ |j_E| \quad \text{and} \quad \|G'\|^2_{L^2(\Gamma \setminus \Sigma)} \le m,$$

which together imply (50). First, both sides of the first inequality in (52) vanish in the good reduction case, so we may assume that $E/\mathbb{K}$ has multiplicative reduction; thus $\ell := \log|j_E| > 0$. Since each $i(x, R)$ ($R \in \mathscr{R}$) vanishes on the skeleton, for $x \in \Sigma \simeq \mathbb{R}/\ell\mathbb{Z}$ we have

$$G(x) = \frac{1}{|\mathscr{R}|} \sum_{R \in \mathscr{R}} j(x, R) = \frac{1}{|\mathscr{R}|} \sum_{R \in \mathscr{R}} \frac{\ell}{2} \Phi\left(\frac{x - r_\Sigma(R)}{\ell}\right),$$

and since $\Phi'(x) = 2\Psi(x)$, we deduce that $G'(x) = |\mathscr{R}|^{-1} \sum_{R \in \mathscr{R}} \Psi_\ell(x - r_\Sigma(R))$; here we are using the notation of Section 2.3. Therefore $\hat{G}'(0) = 0$ and for $k \ne 0$,

$$\hat{G}'(k) = \frac{-1}{2\pi i k}\left(\frac{1}{|\mathscr{R}|} \sum_{R \in \mathscr{R}} e^{-2\pi i k r_\Sigma(R)/\ell}\right);$$

in particular $|\hat{G}'(k)| \leq 1/2\pi |k|$ for $k \neq 0$. It follows from Parseval's formula that

$$\int_0^\ell |G'(x)|^2 dx = \ell \sum_{k \in \mathbb{Z}} |\hat{G}'(k)|^2 \leq \ell \sum_{k \neq 0} \frac{1}{4\pi^2 k^2} = \tfrac{1}{12}\ell,$$

which proves the first inequality in (52) in this case.

To prove the second, fix $R \in \mathfrak{R}$, and identify the line segment $[r_\Sigma(R), \kappa_R]$, which by the definition of κ_R has length m, with the real interval $[0, m]$. Defining $g_R : [r_\Sigma(R), \kappa_R] \to \mathbb{R}$ by $g_R(x) = i(x, R)$, we have $g_R(x) = \rho(r_\Sigma(R), x) = x$, and thus $g_R'(x) \equiv 1$ on $[r_\Sigma(R), \kappa_R]$. Therefore $\int_{[r_\Sigma(R),\kappa_R]} |g_R'(x)|^2 dm_\Gamma(x) = m$. Extending the definition of g_R by declaring that $g_R(x) = 0$ for all $x \notin [r_\Sigma(R), \kappa_R]$, we have $G'(x) = |\mathfrak{R}|^{-1} \sum_{R \in \mathfrak{R}} g_R'(x)$ for $x \in \Gamma \setminus \Sigma$, and thus

$$\int_{\Gamma \setminus \Sigma} |G'(x)|^2 dm_\Gamma(x) = \int_{\Gamma \setminus \Sigma} \left| \frac{1}{|\mathfrak{R}|} \sum_{R \in \mathfrak{R}} g_R'(x) \right|^2 dm_\Gamma(x)$$

$$= \frac{1}{|\mathfrak{R}|^2} \sum_{Q, R \in \mathfrak{R}} \int_{\Gamma \setminus \Sigma} g_Q'(x) g_R'(x) dm_\Gamma(x)$$

$$\leq \frac{1}{|\mathfrak{R}|^2} \sum_{Q, R \in \mathfrak{R}} m = m,$$

which is the second inequality in (52). $\qquad\square$

4.2. *The proofs of Theorems 4 and 12.*

Proof of Theorem 4. This follows immediately from Theorem 12 since enlarging the set S only makes the set $\mathcal{E}(\overline{\mathbb{O}}_S, Q)_{\mathrm{tor}}$ larger. $\qquad\square$

Proof of Theorem 12. Note that the set $\mathcal{E}(\overline{\mathbb{O}}_S, Q)_{\mathrm{tor}}$ is $\mathrm{Aut}(\overline{K}/K)$-invariant. Given a finite $\mathrm{Aut}(\overline{K}/K)$-invariant subset Z of $\mathcal{E}(\overline{\mathbb{O}}_S, Q)_{\mathrm{tor}}$, we are going to show that $|Z|$ is at most the right side of (49), which will prove the theorem. We may assume that Z is nonempty, for if $\mathcal{E}(\overline{\mathbb{O}}_S, Q)_{\mathrm{tor}}$ is empty the theorem is trivial.

Let $\mathfrak{Q}$ be the set of $\mathrm{Aut}(\overline{K}/K)$-conjugates of Q, and let L/K be a finite normal extension such that $\mathfrak{Q} \cup Z \subseteq E(L)$. For each $w \in M_L$, define

$$\Lambda_w(Z, \mathfrak{Q}) = \frac{1}{|\mathfrak{Q}||Z|} \sum_{P \in Z} \sum_{R \in \mathfrak{Q}} \lambda_w(P - R),$$

and for each $v \in M_K$, let $\Lambda_v(Z, \mathfrak{Q}) = \sum_{w | v}([L_w : K_v]/[L : K])\Lambda_w(Z, \mathfrak{Q})$. Thus

(53) $$\hat{h}(Q) = \frac{1}{|\mathfrak{Q}||Z|} \sum_{P \in Z} \sum_{R \in \mathfrak{Q}} \hat{h}(P - R) = \sum_{v \in M_K} \Lambda_v(Z, \mathfrak{Q})$$

by (44); here we have used the fact that $\hat{h}(P - R) = \hat{h}(R) = \hat{h}(Q)$ for all $P \in Z$ and $R \in \mathfrak{Q}$, since the points P in Z are torsion points and $\hat{h}$ is $\mathrm{Aut}(\overline{K}/K)$-invariant.

We will use the S-integrality hypothesis to show that

$$\Lambda_v(Z, \mathcal{Q}) = 0 \quad \text{for all } v \in M_K \setminus S, \tag{54}$$

and we will use equidistribution (Theorem 11) to show that

$$\Lambda_v(Z, \mathcal{Q}) \leq |Z|^{-1/2}\left(\tfrac{1}{12}h(j_E) + m_v(Q)\right) \quad \text{for all } v \in S. \tag{55}$$

Assuming these two claims hold for now, we deduce from (53) that

$$\hat{h}(Q) \leq \sum_{v \in S} |Z|^{-1/2}\left(\tfrac{1}{12}h(j_E) + m_v(Q)\right) = |Z|^{-1/2}\left(\tfrac{1}{12}|S|h(j_E) + \sum_{v \in S} m_v(Q)\right),$$

and the desired inequality (49) follows immediately, completing the proof of the theorem. Moreover, our assumption that Z is nonempty ensures that the set S is nonempty; for if S were empty, then (53) and (54) would contradict the assumption that $\hat{h}(Q) > 0$. In other words, this shows that $\mathcal{E}(\overline{\mathbb{O}}_S, Q)_{\mathrm{tor}}$ is empty if S is empty.

We will now show (54). Assuming $v \in M_K \setminus S$, the Weierstrass equation (2) is defined over $\mathbb{O}_S$ with $|\Delta|_v = 1$ and $|j_E|_v \leq 1$. Thus the model $\mathcal{E}/\mathbb{O}_S$ has good reduction at v, meaning the fiber $\overline{\mathcal{E}}_v/k_v$ is an elliptic curve over the residue field $k_v = \mathbb{O}_v/\mathbb{M}_v$, and the associated reduction map $\pi_v : E(\mathbb{K}_v) \to \overline{\mathcal{E}}_v(k_v)$ is a group homomorphism. Since the Zariski closure in $\mathcal{E}$ of each point P in Z does not meet the Zariski closure of Q, this means that $\pi_v(\epsilon(P)) \neq \pi_v(\epsilon(R))$ for every embedding $\epsilon : L \hookrightarrow \mathbb{K}_v$ and every $R \in \mathcal{Q}$. Thus if $w \mid v$ is a place of L corresponding to the embedding ϵ, then $\lambda_w(P - R) = \lambda_v(\epsilon(P) - \epsilon(R)) = i_v(\epsilon(P), \epsilon(R)) = 0$ by (17) and Proposition 7. We conclude that $\Lambda_w(Z, \mathcal{Q}) = 0$ for all $w \mid v$, and thus $\Lambda_v(Z, \mathcal{Q}) = 0$ for all $v \in M_K \setminus S$, which is (54).

It now remains only to prove (55). Assume that $v \in S$, and fix once and for all a K-embedding $\epsilon : \overline{K} \hookrightarrow \mathbb{K}_v$. We claim that for each place $w \mid v$ of L,

$$\Lambda_v(Z, \mathcal{Q}) = \Lambda_w(Z, \mathcal{Q}) = \frac{1}{|\mathcal{Q}||Z|} \sum_{P \in Z} \sum_{R \in \mathcal{Q}} \lambda_v(\epsilon(P) - \epsilon(R)). \tag{56}$$

To see this fact, note that each place $w \mid v$ of L arises from *some* K-embedding $\delta : \overline{K} \hookrightarrow \mathbb{K}_v$ where $\lambda_w(\cdot) = \lambda_v(\delta(\cdot))$. Since L/K is a normal extension, $\delta = \epsilon \circ \alpha$ for some $\alpha \in \mathrm{Aut}(L/K)$, and therefore since both Z and $\mathcal{Q}$ are $\mathrm{Aut}(L/K)$-invariant, the second identity in (56) follows. The first identity in (56) follows from the local degree formula and the fact that the right side is independent of $w \mid v$.

Define a $F_v : \mathsf{E}_v \setminus \epsilon(\mathcal{Q}) \to \mathbb{R}$, $x \mapsto |\mathcal{Q}|^{-1} \sum_{R \in \mathcal{Q}} \lambda_v(x - \epsilon(R))$. Note that F_v is continuous on $\mathsf{E}_v \setminus \epsilon(\mathcal{Q})$ but has logarithmic singularities at the points of $\epsilon(\mathcal{Q})$. To get around this we define the truncated function $G_v : \mathsf{E}_v \to \mathbb{R}$ by

$$G_v(x) = \frac{1}{|\mathcal{Q}|} \sum_{R \in \mathcal{Q}} (j_v(x, R) + i_v^*(x, R)), \tag{57}$$

where $i_v^*(x, R) = \min\{i_v(x, R), m_v(Q)\}$, and where $m_v(Q)$ is defined in (48). Therefore, using that $i_v(P, Q) = -\log d_v(P, Q)$ from Section 1.6, we still have

$$(58) \qquad\qquad G_v(P) = F_v(P) \quad \text{for all } P \in E(\mathbb{K}_v)_{\text{tor}}.$$

Note that $0 \le m_v(Q) < +\infty$ since Q is nontorsion. By its definition, $m_v(Q)$ is the minimal constant at which to truncate $i_v(x, R)$ such that (58) holds. Also F_v and G_v agree on the skeleton Σ_v of E_v. In particular $\int G_v \, d\mu_v = \int F_v \, d\mu_v = 0$, since the canonical measure μ_v is supported on Σ_v, and $\int \lambda_v(x, y) \, d\mu_v(x) = \int j_v(x, y) \, d\mu_v(x) = \int_{\Sigma_v} j_v(x, y) \, dm_{\Sigma_v}(x) = 0$ for all $y \in \mathsf{E}_v$; here we have used that $i_v(x, y) = 0$ for $x \in \Sigma_v$.

For each $R \in \mathcal{Q}$, let $\kappa_R \in \mathsf{E}_v$ be the unique point in the (infinitely long) line segment $[r_{\Sigma_v}(\epsilon(R)), \epsilon(R)]$ that satisfies $\rho(r_{\Sigma_v}(\epsilon(R)), \kappa_R) = m_v(Q)$. Let

$$\Gamma_v = \Sigma_v \cup \left(\bigcup_{R \in \mathcal{Q}} [r_{\Sigma_v}(\epsilon(R)), \kappa_R] \right)$$

be the smallest connected subgraph of E_v containing Σ_v and all of the points κ_R. By Lemma 13, G_v is a test function in the space $\mathscr{S}_{\Gamma_v}(\mathsf{E}_v, \mathbb{R})$ associated to the connected metrized subgraph Γ_v of E_v,

$$(59) \qquad \|G_v'\|^2_{L^2(\Gamma_v)} \le \tfrac{1}{12} \log^+ |j_E|_v + m_v(Q) \le \tfrac{1}{12} h(j_E) + m_v(Q),$$

and $\ell_0(\Gamma_v) = m_v(Q)$. Applying (56), Theorem 11, and the fact that $\int G_v \, d\mu_v = 0$ and $\hat{h}(Z) = 0$, we get (55) by

$$\begin{aligned}
\Lambda_v(Z, \mathcal{Q}) = |Z|^{-1} \sum_{P \in Z} F_v(\epsilon(P)) &= |Z|^{-1} \sum_{P \in Z} G_v(\epsilon(P)) \\
&\le \|G_v'\|_{L^2(\Gamma_v)} \left(\tfrac{1}{12} |Z|^{-1} h(j_E) + |Z|^{-1} \ell_0(\Gamma_v) \right)^{1/2} \\
&\le |Z|^{-1/2} \left(\tfrac{1}{12} h(j_E) + m_v(Q) \right). \qquad \square
\end{aligned}$$

References

[Baker 2008] M. Baker, "An introduction to Berkovich analytic spaces and non-archimedean potential theory on curves", pp. 123–174 in *p-adic geometry* (Tucson, AZ, 2007), University Lecture Series **45**, American Mathematical Society, Providence, RI, 2008. Zbl 1153.14018

[Baker 2009] M. Baker, "A finiteness theorem for canonical heights attached to rational maps over function fields", *J. Reine Angew. Math.* **626** (2009), 205–233. MR 2492995 Zbl 05505186

[Baker and Petsche 2005] M. Baker and C. Petsche, "Global discrepancy and small points on elliptic curves", *Int. Math. Res. Not.* 61 (2005), 3791–3834. MR 2007k:11103 Zbl 1120.11025

[Baker and Rumely 2006] M. H. Baker and R. Rumely, "Equidistribution of small points, rational dynamics, and potential theory", *Ann. Inst. Fourier* (*Grenoble*) **56**:3 (2006), 625–688. MR 2007m: 11082 Zbl 05176555

[Baker and Rumely 2007] M. Baker and R. Rumely, "Harmonic analysis on metrized graphs", *Canad. J. Math.* **59**:2 (2007), 225–275. MR 2008h:43019 Zbl 1123.43006

[Baker and Rumely 2009] M. H. Baker and R. Rumely, "Potential theory on the Berkovich projective line", manuscript, 2009. To appear in Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI.

[Baker et al. 2008] M. Baker, S.-i. Ih, and R. Rumely, "A finiteness property of torsion points", *Algebra Number Theory* **2**:2 (2008), 217–248. MR 2377370 Zbl 05529325

[Berkovich 1990] V. G. Berkovich, *Spectral theory and analytic geometry over nonarchimedean fields*, Mathematical Surveys and Monographs **33**, American Mathematical Society, Providence, RI, 1990. MR 91k:32038 Zbl 0715.14013

[Bombieri and Gubler 2006] E. Bombieri and W. Gubler, *Heights in Diophantine geometry*, New Mathematical Monographs **4**, Cambridge Univ. Press, 2006. MR 2007a:11092 Zbl 1115.11034

[Bosch et al. 1984] S. Bosch, U. Güntzer, and R. Remmert, *Nonarchimedean analysis: A systematic approach to rigid analytic geometry*, Grundlehren der Mathematischen Wissenschaften **261**, Springer, Berlin, 1984. MR 86b:32031 Zbl 0539.14017

[Call and Silverman 1993] G. S. Call and J. H. Silverman, "Canonical heights on varieties with morphisms", *Compositio Math.* **89**:2 (1993), 163–205. MR 95d:11077 Zbl 0826.14015

[Chambert-Loir 2006] A. Chambert-Loir, "Mesures et équidistribution sur les espaces de Berkovich", *J. Reine Angew. Math.* **595** (2006), 215–235. MR 2008b:14040 Zbl 1112.14022

[Chinburg and Rumely 1993] T. Chinburg and R. Rumely, "The capacity pairing", *J. Reine Angew. Math.* **434** (1993), 1–44. MR 94b:14019 Zbl 0756.14013

[Faber 2008] X. W. C. Faber, "Equidistribution of dynamically small points over function fields of curves", preprint, 2008. arXiv 0801.4811

[Favre and Rivera-Letelier 2006] C. Favre and J. Rivera-Letelier, "Équidistribution quantitative des points de petite hauteur sur la droite projective", *Math. Ann.* **335** (2006), 311–361. MR 2007g:11074 Zbl 05035986

[Gubler 2008] W. Gubler, "Equidistribution over function fields", *Manuscripta Math.* **127**:4 (2008), 485–510. MR 2457191 Zbl 05520090

[Lang 1983] S. Lang, *Fundamentals of Diophantine geometry*, Springer, New York, 1983. MR 85j: 11005 Zbl 0528.14013

[Serre 1989] J.-P. Serre, *Lectures on the Mordell–Weil theorem*, Aspects of Mathematics **15**, Friedr. Vieweg & Sohn, Braunschweig, 1989. MR 90e:11086 Zbl 0676.14005

[Silverman 1986] J. H. Silverman, *The arithmetic of elliptic curves*, Graduate Texts in Mathematics **106**, Springer, New York, 1986. MR 87g:11070 Zbl 0585.14026

[Silverman 1994] J. H. Silverman, *Advanced topics in the arithmetic of elliptic curves*, Graduate Texts in Mathematics **151**, Springer, New York, 1994. MR 96b:11074 Zbl 0911.14015

[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 97d:11096 Zbl 1071.11508

[Thuillier 2005] A. Thuillier, *Théorie du potentiel dur les courbes en géométrie analytique non archimédienne: Applications à la théorie d'Arakelov*, thesis, University of Rennes, 2005.

CLAYTON PETSCHE
DEPARTMENT OF MATHEMATICS AND STATISTICS
HUNTER COLLEGE
695 PARK AVENUE
NEW YORK, NY 10065
UNITED STATES
cpetsche@hunter.cuny.edu