Canonical Heights on Abelian Varieties
Lecture Notes for the Arizona Winter School
March 2–6, 2024

JOSEPH H. SILVERMAN
joseph_silverman@brown.edu

Contents

Date: February 13, 2024.

2020 *Mathematics Subject Classification.* Primary: 11G10; Secondary: 11G35, 11G50, 14G40, 14K15.

Key words and phrases. abelian variety, canonical height.

1. About These Notes/Note to Students

These notes are for the Arizona Winter School on *Abelian Varieties*, March 2–6, 2024. We start with sections contains background material on algebraic geometry (§2), abelian varieties (§3), and height functions (§4). This is followed by sections with expanded versions of the material covered in the lectures:

Lecture 1: Construction and properties of canonical heights (§5)
Lecture 2: Applications (§6); Local canonical heights (§7)
Lecture 3: Lower bounds for canonical heights (§8)
Lecture 4: Canonical heights in families; specialization theorems (§9)

Two final sections briefly discuss further topics (§10) and provide some references to learn more about abelian varieties (§11). We have also included a List of Notation and an Index to assist in reading these notes.

Each section includes a number of exercises that are designed to help the reader gain some feel for the subject matter. There are also brief paragraphs in small type marked "Supplementary Material" that describe advanced concepts and generalizations. This material is not used in these notes and may be skipped on first reading. Appendix A contains a number of proposed projects for our Winter School Working Group. The specific questions described in these projects are meant to serve as guidelines, and we may well find ourselves pursuing other problems during the workshop.

Acknowledgements. I would like to thank the organizers of the Arizona Winter School for inviting me to deliver these lectures. The author used ChatGPT to help generate the code for the TiKZ diagrams.

2. Background Material: Algebraic Geometry

Figure 1 gives some standard algebraic geometry notation that we will use throughout these notes.

K	a field, typically a number field, a local field, or an algebraically closed field.
$\bar{K}$	an algebraic closure of K.
X/K	a smooth projective algebraic variety, defined over K.
$K(X)$, $\bar{K}(X)$	the function field of X over K, respectively over $\bar{K}$.
$\operatorname{Div}(X)$	the group of geometric divisors of X, i.e., divisors defined over $\bar{K}$.
$\operatorname{div}(f)$	the divisor of a function $f \in \bar{K}(X)$.
$\sim$	linear equivalence, $D \sim D'$ if $D - D' = \operatorname{div}(f)$ for some $f \in \bar{K}(X)$.
$\mathcal{L}(D)$	$= H^0\bigl(X, \mathcal{O}_X(D)\bigr) = \bigl\{f \in \bar{K}(X) : \operatorname{div}(f) + D \ge 0\bigr\}$.
$\ell(D)$	$= \dim_{\bar{K}} \mathcal{L}(D)$.
$\equiv$	algebraic equivalence of divisors.
$\operatorname{Pic}(X)$	$= \operatorname{Div}(X)/\sim$, the Picard group of X.
$\operatorname{Div}^0(X)$	$= \{D \in \operatorname{Div}(X) : D \equiv 0\}$.
$\operatorname{Pic}^0(X)$	$= \operatorname{Div}^0(X)/\sim$.
$\operatorname{NS}(X)$	$= \operatorname{Div}(X)/\equiv$, the Néron-Severi group of X.
$\rho(X)$	$= \operatorname{rank} \operatorname{NS}(X)$.
$\operatorname{End}(X)$	the ring of endomorphisms $X \to X$.
$\operatorname{Aut}(X)$	the group of automorphisms $X \to X$, i.e., $\operatorname{Aut}(X) = \operatorname{End}(X)^*$.

FIGURE 1. Notation that will be used throughout these notes

Definition 2.1. Let $D \in \operatorname{Div}(X)$ be a divisor such that $\ell(D) \ge 1$. Let $\{f_1, \ldots, f_{\ell(D)}\}$ be a $\bar{K}$-basis for $\mathcal{L}(D)$. The divisor D is *very ample* if the rational map

$$[f_0, \ldots, f_n] : X \longrightarrow \mathbb{P}^{\ell(D)-1}$$

is an embedding, i.e., an isomorphism onto its image. The divisor D is *ample* if some multiple mD with $m \ge 1$ is very ample. If D is very ample, we select a basis for $\mathcal{L}(D)$ and write

$$f_D : X \hookrightarrow \mathbb{P}^{\ell(D)-1}$$

for the associated embedding. If X and D are defined over K, we may assume that f_D is also defined over K.

Exercises for Section 2.

Exercise 2.A. Let $F(x_0,\dots,x_N) \in \bar{K}[x_0,\dots,x_N]$ be a non-zero homogeneous polynomial of degree $d \ge 1$, and let

$$D_F = \{F = 0\} \in \operatorname{Div}(\mathbb{P}^N).$$

(a) Write down a basis for $\mathcal{L}(D_F)$ and compute $\ell(D_F)$.
(b) Prove that D_F is very ample directly from the definition.

Exercise 2.B. Prove that

$$\operatorname{Pic}^0(\mathbb{P}^N) = 0 \quad \text{and} \quad \operatorname{Pic}(\mathbb{P}^N) = \operatorname{NS}(\mathbb{P}^N) \cong \mathbb{Z}.$$

3. Background Material: Abelian Varieties

Definition 3.1. An *abelian variety* is a projective group variety, i.e., a projective variety A with a base point $O \in A$ and morphisms

$$\mu : A \times A \to A \quad \text{and} \quad \iota : A \to A$$

that give A the structure of a group:

$$\mu(P,O) = \mu(O,P) = P.$$
$$\mu\bigl(P,\iota(P)\bigr) = \mu\bigl(\iota(P),P\bigr) = O.$$
$$\mu\bigl(\mu(P,Q),R\bigr) = \mu\bigl(P,\mu(Q,R)\bigr).$$

We say that the abelian variety is defined over K if A, μ, and ι are defined over K and $O \in A(K)$.

Theorem 3.2. *An abelian variety is a smooth projective variety, and its group law is abelian,*

$$\mu(P,Q) = \mu(Q,P).$$

We thus typically write the group operations as

$$\mu(P,Q) = P + Q \quad \text{and} \quad \iota(P) = -P.$$

Remark 3.3. Every group variety is smooth, since a variety has at least one non-singular point, and then translation can be used to show that every point is non-singular. However, the commutativity of the group law on an abelian variety is a consequence of the assumed projectivity. Note that there are many non-commutative group varieties, for example GL_n and SL_n, but they are only quasi-projective, not projective.

Definition 3.4. Let A and B be abelian varieties. An *isogeny from A to B* is a finite map $\varphi : A \to B$ satisfying $\varphi(O_A) = O_B$. We say that A and B are *isogenous* if there exists an isogeny from A to B. This is an equivalence relation, because one can show that if there is an isogeny from A to B, then there is also an isogeny from B to A; see Exercise 3.B.

Definition 3.5. Let A be an abelian variety, and let $Q \in A$. The *translation-by-Q map* is the map

$$T_Q : A \longrightarrow A, \quad T_Q(P) = P + Q.$$

Theorem 3.6. *Let A and B be abelian varieties.*

(a) *Every isogeny $\varphi : A \to B$ is a group homomorphism.*
(b) *Every finite morphism $A \to B$ is a composition of an isogeny and a translation.*
(c) *The collection of isogenies from A to B, which we denote by* $\operatorname{Hom}(A, B)$, *is a group via the group law*

$$(\varphi + \psi)(P) = \varphi(P) + \psi(P).$$

(d) *The collection of self-isogenies of A, which we denote by* $\operatorname{End}(A)$, *is a ring via*

$$(\varphi + \psi)(P) = \varphi(P) + \psi(P) \quad \textit{and} \quad (\varphi\psi)(P) = \varphi\bigl(\psi(P)\bigr).$$

Theorem 3.7 (Theorem of the cube)**.** *Let A and B be abelian varieties, let $\varphi_1, \varphi_2, \varphi_3 \in \operatorname{Hom}(A, B)$ be isogenies, and let $D \in \operatorname{Div}(B)$. Then*

$$(\varphi_1 + \varphi_2 + \varphi_3)^*D - (\varphi_1 + \varphi_2)^*D - (\varphi_1 + \varphi_3)^*D - (\varphi_2 + \varphi_3)^*D + \varphi_1^*D + \varphi_2^*D + \varphi_3^*D \sim 0.$$

Corollary 3.8. *Let A be an abelian variety, let $m \in \mathbb{Z}$, let $D \in \operatorname{Div}(A)$, and let $D^- = [-1]^*D$. Then*

$$[m]^*D \sim \frac{m^2 + m}{2} D + \frac{m^2 - m}{2} D^-.$$

In particular, if $D^- \sim D$, then $[m]^*D \sim m^2 D$.

Theorem 3.9 (Theorem of the square)**.** *Let A be an abelian variety, let $P, Q \in A$, and let $D \in \operatorname{Div}(B)$. Then*

$$T_{P+Q}^*D - T_P^*D - T_Q^*D + D \sim 0.$$

Theorem 3.10. *Let A be an abelian variety. The Picard group* $\operatorname{Pic}^0(A)$ *has a natural structure as an abelian variety $\hat{A}$, which is called the* dual of A. *The association $A \to \hat{A}$ is functorial in the sense that each homomorphism $\alpha : A \to B$ induces a homomorphism $\hat{\alpha} : \hat{B} \to \hat{A}$. At the level of divsor classes, the map $\hat{\alpha}$ is given by*

$$\hat{\alpha}\bigl([D]_B\bigr) = \bigl[\varphi^*(D)\bigr]_A,$$

where we use $[\,\cdot\,]_A$ and $[\,\cdot\,]_B$ to denote divisor classes on A and B.

Definition 3.11. Let A be an abelian variety, and let $D \in \operatorname{Div}(A)$. We define a map

$$\varphi_D : A \longrightarrow \hat{A}, \quad \varphi_D(P) = [T_{-P}^* D - D],$$

where in general $[D]$ denotes the divisor class of D in $\operatorname{Pic}(A)$.

Theorem 3.12. *Let A be an abelian variety, and let $D \in \operatorname{Div}(A)$.*

(a) *The map $\varphi_D : A \to \hat{A}$ depends only on the algebraic equivalence class of D in $\operatorname{NS}(A)$.*
(b) *The map φ_D is a homomorphism.*
(c) *The map φ_D is an isogeny if and only if D is an ample divisor.*

Definition 3.13. Let A be an abelian variety. A *polarization of A* is a divisor $D \in \operatorname{Div}(A)$ such that $\varphi_D : A \to \hat{A}$ is an isogeny. The *degree* of the polarization is the degree of φ_D. It is a *principal polarization* if it has degree 1. We say that A is *principally polarized* if it has a principal polarization.

Theorem 3.14. *Let A be an abelian variety.*

(a) *There exists a principally polarized abelian variety that is isogenous to A.*
(b) (Zarhin trick [93]) *The product $A^4 \times \hat{A}^4$ is principally polarized.*

Definition 3.15. An abelian variety A is *simple* if it is not possible to write it as a product $A \cong B \times C$ of positive dimensional abelian varieties B and C.

Theorem 3.16. *Let A be a simple abelian variety, and let* $\operatorname{End}(A)_{\mathbb{Q}} = \operatorname{End}(A) \otimes \mathbb{Q}$. *Then* $\operatorname{End}(A)_{\mathbb{Q}}$ *is a finite-dimensional semisimple* $\mathbb{Q}$-*algebra. Further,*

$$\dim_{\mathbb{Q}} \operatorname{End}(A)_{\mathbb{Q}} \le \begin{cases} 2\dim(A) & \text{if } \operatorname{char}(K) = 0, \\ 4\dim(A) & \text{if } \operatorname{char}(K) > 0. \end{cases}$$

The endomorphism ring $\operatorname{End}(A)$ *is an order in* $\operatorname{End}(A)_{\mathbb{Q}}$.

We remark that much more is known about the structure of $\operatorname{End}(A) \otimes \mathbb{Q}$; see for example [65, §§19–21].

Theorem 3.17 (Poincaré complete reducibility theorem)**.** *Let A be an abelian variety.*

(a) *There are simple pairwise non-isogenous abelian varieties $A_1, \ldots, A_r$ and exponents $e_1, \ldots, e_r \ge 1$ so that A is isogenous to the product*

$$A \xrightarrow{\text{isogeny}} A_1^{e_1} \times \cdots \times A_r^{e_r}.$$

Up to relabeling, the abelian varieties $A_1, \ldots, A_r$ and postive integers $e_1, \ldots, e_r$ are uniquely determined by A.

(b) *Continuing with the notation from* (a)*, the endomorphism algebra* $\operatorname{End}(A)_{\mathbb{Q}}$ *of* A *is a product of matrix algebras,*

$$\operatorname{End}(A)_{\mathbb{Q}} \cong \prod_{i=1}^{r} \operatorname{Mat}_{r_i \times r_i}\bigl(\operatorname{End}(A_i)_{\mathbb{Q}}\bigr).$$

An important tool in studying the endomorphism ring of a abelian variety A is the Rosatti involution, which is an involution of the endomorphism algebra $\operatorname{End}(A)_{\mathbb{Q}}$.

Definition 3.18. We fix an ample divisor $H \in \operatorname{Div}(A)$, and we let $\varphi_H : A \to \hat{A}$ be the associated isogeny; see Definition 3.11 and Theorem 3.12. For $\alpha \in \operatorname{End}(A)$, we let $\hat{\alpha} : \hat{A} \to A$ be the dual map as described in Theorem 3.10 with $A = B$. The *Rosatti involution* on $\operatorname{End}(A)_{\mathbb{Q}}$ is the map[1]

$$\operatorname{End}(A)_{\mathbb{Q}} \longrightarrow \operatorname{End}(A)_{\mathbb{Q}}, \quad \alpha \longmapsto \alpha' := \varphi_H^{-1} \circ \hat{\alpha} \circ \varphi_H.$$

We also use H to define a map from divisor classes to endomorphisms:

$$\Phi_D : \operatorname{NS}(A)_{\mathbb{Q}} \hookrightarrow \operatorname{End}(A)_{\mathbb{Q}}, \quad \Phi_D = \varphi_H^{-1} \circ \varphi_D.$$

We state some useful properties of the Rosatti involution.

Theorem 3.19. *The Rosatti involution and the map* Φ_D *interact in various ways, including the following*:

(a) $\Phi_D\bigl(\operatorname{NS}(A)_{\mathbb{Q}}\bigr) = \bigl\{\alpha \in \operatorname{End}(A)_{\mathbb{Q}} : \alpha' = \alpha\bigr\}$.
(b) $\Phi_{\alpha^* D} = \alpha' \circ \Phi_D \circ \alpha$ *for all* $\alpha \in \operatorname{End}(A)_{\mathbb{Q}}$.
(c) *If* $D \in \operatorname{NS}(A)_{\mathbb{R}}$ *is a nef divisor,*[2] *then there is an* $\alpha \in \operatorname{End}(A)_{\mathbb{R}}$ *satisfying*

$$\Phi_D = \alpha' \circ \alpha \quad \textit{and} \quad \alpha' = \alpha.$$

Proof. (a) See [65, page 208].
(b) See [43, Lemma 24].
(c) See [43, Proposition 26]. □

3.1. **Abelian Varieties over Number Fields.** The following fundamental theorem was proven originally by Mordell for elliptic curves over $\mathbb{Q}$ and subsequently generalized by Weil to abelian varieties over number fields.

[1] We note that φ_H will have an inverse in $\operatorname{Hom}(\hat{A}, A)$ if and only if H gives a principal polarization, but that φ_H always has an inverse in $\operatorname{Hom}(\hat{A}, A) \otimes \mathbb{Q}$.

[2] A nef (numerically effective) divisor is a divisor D whose intersection with every curve $C \subset A$ satisfies $D \cdot C \geq 0$. Equivalently, D is nef if it lies in the closure of the ample cone in $\operatorname{NS}(A)_{\mathbb{R}}$.

Theorem 3.20 (Mordell–Weil Theorem)**.** *Let $K/\mathbb{Q}$ be a number field, and let A/K be an abelian variety. Then the group of K-rational points $A(K)$ is a finitely generated abelian group.*

Proof. There are many places to read the proof of the Mordell–Weil theorem, including [39, Part C], [47, Chapter 6], [47, Appendix II], [88, Chapters VIII and X] (the last reference only covers elliptic curves). □

Remark 3.21. It follows from Theorem 3.20 that $A(K)$ is the product of a finite group and a free abelian group of finite rank,

$$A(K) \cong A(K)_{\text{tors}} \times \mathbb{Z}^{\operatorname{rank} A(K)}.$$

It is conjectured that

$$\#A(K)_{\text{tors}} \le C_1\bigl([K:\mathbb{Q}], \dim(A)\bigr),$$

but this has only been proven for $\dim(A) = 1$ [42, 61, 63], and there are still open problems concerning the optimal value of $C_1(n,d)$, although it is known, for example, that $C_1(1,1) = 16$.

The rank is even more mysterious. For many years it was conjectured that for every R there was an elliptic curve $E/\mathbb{Q}$ with $\operatorname{rank} E(\mathbb{Q}) \ge R$. However, there is an heuristic argument [71] suggesting that the converse is true, and indeed the authors of [71] suggest for example that

$$\bigl\{E/\mathbb{Q} : \dim(E) = 1 \text{ and } \operatorname{rank} E(\mathbb{Q}) \ge 22\bigr\}$$

is a finite set. So one might ask if there is a general bound of the form

$$\operatorname{rank} A(K) \le C_2\bigl([K:\mathbb{Q}], \dim(A)\bigr)?$$

The author expresses no opinion as to the likely validity of such a statement!

3.2. **Abelian Varieties over Function Fields.** Let k be a field, and let $K = k(C)$ be the function field of a smooth projective curve C/k. Then an algebraic variety X/K may actually arise from a variety X_0/k defined over the constant field k, where we simply take X_0 and view it as being defined over K.[3] It's also possible for a variety over K to be partially defined over k, for example, if it's a product $X = Y \times Z$, where Y comes from a variety Y_0/k. For abelian varieties A, the next definition gives a more refined way to define the part of A that comes from an abelian variety defined over k.

[3] In fancier terms, the variety X_0/k is a scheme over $\operatorname{Spec}(k)$, and we obtain X/K be extending the base field via the fiber product $X = X_0 \times_{\operatorname{Spec}(k)} \operatorname{Spec}(K)$.

Definition 3.22. Let k be a field, let $K = k(C)$ be the function field of a smooth projective curve C/k, and let A/K be an abelian variety. A *K/k-trace of A/K* is a pair $(B/k, \varphi)$ having the following properties:

- B/k is an abelian variety.
- $\varphi : B \times_k K \longrightarrow A$ is a homomorphism defined over K.
- If B'/k is an abelian variety and $\varphi' : B' \times_k K \to A$ is a K-homomorphism, then there exists a unique k-homomorphism $\psi : B' \to B$ such that $\varphi' = \varphi \circ \psi$.

We say that two K/k-traces $(B/k, \varphi)$ and $(B'/k, \varphi')$ are isomorphic if there is a k-isomorphism $\psi : B' \to B$ such that $\varphi' = \varphi \circ \psi$.

Intuition 3.23. One may view the K/k-trace of A as being the "constant part of A."

Theorem 3.24. *Let k be a field of characteristic 0, let $K = k(C)$ be the function field of a smooth projective curve C/k, and let A/K be an abelian variety. Then there exists a K/k-trace for A/K, and it is unique up to k-isomorphism.*

Proof. See [49] for the classical construction, and [17] for a modern formulation. □

The Mordell–Weil theorem (Theorem 3.20) for number fields extends to the function field setting, where it is often still given the same name, although in full generality the result is due to Lang and Néron.

Theorem 3.25 (Mordell–Weil (Lang–Néron) Theorem)**.** *Let k be a field of characteristic 0, let $K = k(C)$ be the function field of a smooth projective curve C/k, let A/K be an abelian variety, and let $(B/k, \varphi)$ be a K/k-trace of A/K. Then the group $A(K)/\varphi\bigl(B(k)\bigr)$ is a finitely generated group.*

Remark 3.26. If k is a number field, then the original Mordell–Weil theorem (Theorem 3.20) says that $B(k)$ is finitely generated. Further, we have defined traces and stated Theorem 3.25 when $K = k(C)$ is the function field of a curve, but this material can be extended to higher dimensional bases. Then one version of the Lang–Néron theorem is that if K is a field that is a finitely generated $\mathbb{Q}$-algebra and A/K is an abelian variety, then $A(K)$ is a finitely generated group.

3.3. **Twists of Abelian Varieties.** The geometric isomorphism class of an abelian variety fails to encompass the many possibilities for its arithmetic. For this section, we fix the following notation:

K a number field or a function field.

K^{sep} a separable closure of K.

Definition 3.27. Let A/K be a variety. The set of K^{sep}/K*-twists of* A is[4]

$$\operatorname{Twist}(A/K) = \frac{\{B/K : B \text{ is } K^{\text{sep}}\text{-isomorphic to } A\}}{K\text{-isomorphism}}.$$

Proposition 3.28. *There is a natural identification*

$$\operatorname{Twist}(A/K) \cong H^1\bigl(\operatorname{Gal}(K^{\text{sep}}/K), \operatorname{Aut}(A)\bigr)$$

defined as follows: For each $B \in \operatorname{Twist}(A/K)$, we choose a K^{sep}-isomorphism $\varphi_B : B \to A$, and then

$$B \longmapsto \left[\begin{array}{l}\textit{the cohomology class determined}\\ \textit{by the cocycle } \sigma \longmapsto \varphi_B^{-1} \circ \sigma(\varphi_B).\end{array}\right.$$

Proof. We leave the proof to the reader. □

Example 3.29 (Cyclic Twists)**.** Let $m \ge 2$ with $m \ne 0$ in K, and suppose that there is a subgroup of $\operatorname{Aut}(A/K)$ that is isomorphic as a $\operatorname{Gal}(K^{\text{sep}}/K)$-module to the group $\boldsymbol{\mu}_m \subset (K^{\text{sep}})^*$ of mth roots of unity. Then we can use the maps

$$\begin{aligned} K^*/K^{*m} &\cong H^1\bigl(\operatorname{Gal}(K^{\text{sep}}/K), \boldsymbol{\mu}_m\bigr) \\ &\longrightarrow H^1\bigl(\operatorname{Gal}(K^{\text{sep}}/K), \operatorname{Aut}(A/K)\bigr) \\ &\cong \operatorname{Twist}(A/K) \quad \text{from Exercise 3.C} \end{aligned}$$

to obtain a unique twist of A corresponding to each element of K^*/K^{*m}.

Explicitly, let $D \in K^*$, let A_D/K be the associated twist, and let $\delta \in \bar{K}^*$ satisfy $\delta^m = D$. Then there is a K^{sep}-isomorphism

$$\xi_D : A_D \longrightarrow A$$

characterized by

$$\xi_D^{-1} \circ \sigma(\xi_D) = [\delta^\sigma/\delta] \in \boldsymbol{\mu}_m \subseteq \operatorname{Aut}(A/K).$$

Example 3.30. We assume that $\operatorname{char}(K) \ne 2$. Then every abelian variety A has *quadratic twists*, since we can embed $\boldsymbol{\mu}_2$ into $\operatorname{Aut}(A)$ via

$$\boldsymbol{\mu}_2 \hookrightarrow \operatorname{Aut}(A), \quad \zeta \longmapsto \bigl(P \mapsto [\zeta](P)\bigr).$$

For an elliptic curve, i.e., for $\dim(A) = 1$, if we start with a Weierstrass equation

$$E : y^2 = x^3 + ax + b$$

[4]An abelian variety is group variety, and an isomorphism of an abelian varieties is a homomorphism. One can also define twists of general varieties that don't have marked points.

and let $D \in K^*$, then the associated quadratic twist is[5]

$$E^{(D)} : Dy^2 = x^3 + ax + b.$$

Example 3.31. Let $m \ge 2$ with $m \ne 0$ in K, and let $f(x) \in K[x]$ be a separable polynomial with $\deg(f) \ge 3$. Then the automorphism group of the curve

$$C_m : y^m = f(x)$$

contains a copy of $\boldsymbol{\mu}_m$ via

$$\boldsymbol{\mu}_m \hookrightarrow \operatorname{Aut}(C_m), \quad \zeta \longmapsto \bigl[(x,y) \mapsto (x, \zeta y)\bigr].$$

These automorphisms of C_m induce automorphisms of the Jacobian J_m of C_m, so we have an inclusion $\boldsymbol{\mu}_m \subseteq \operatorname{Aut}(J_m)$, and as described in each $D \in K^*/K^{*m}$ leads to a twist $J_m^{(D)}$ of J_m. Explicitliy, the abelian variety $J_m^{(D)}$ is the Jacobian of the twisted curve

$$C_m^{(D)} : Dy^m = f(x).$$

Exercises for Section 3.

Exercise 3.A. Use the theorem of the cube (Theorem 3.7) to prove Corollary 3.8.

Exercise 3.B. Let $\varphi : A \to B$ be an isogeny of abelian varieties.

(a) Prove that there is an isogeny $\hat{\varphi} : B \to A$.
(b) Prove that the relation of being isogenous is an equivalence relation.

Exercise 3.C. Prove Proposition 3.28. (If you're not familiar with non-abelian cohomology, you may assume that $\operatorname{Aut}(A)$ is abelian, or in any case restrict to an abelian subgroup of $\operatorname{Aut}(A)$.) In particular:

(a) Prove that $\sigma \longmapsto \varphi_B^{-1} \circ \sigma(\varphi_B)$ is a cocycle.
(b) Prove that choosing a different K^{sep}-isomorphism $\psi_B : B \to A$ changes the cocycle by a coboundary.
(c) From (a) and (b) we get a well-defined map

$$\operatorname{Twist}(A/K) \to H^1\bigl(\operatorname{Gal}(K^{\text{sep}}/K), \operatorname{Aut}(A)\bigr).$$

Prove that this map is injective.
(d) Prove that the map in (c) is surjective. [This is more difficult.]

Exercise 3.D. Let A/K be an abelian variety.

(a) Let A'/K be a quadratic twist of A/K, as described in Example 3.30. Prove that there is a unique quadratic extension L/K such that A' is L-isomorphic to A, i.e., there is an isomorphism $A' \to A$ that is defined over L.

[5]By a simple change of variable, we can also write the quadratic twist of E as $E^{(D)} : y^2 = x^3 + D^2ax + D^3b$.

K	a global field, i.e., a number field or the function field of a curve.
M_K	a complete set of normalized absolute values on K; see Definition 4.1.
M_K^∞	the archimedean absolute values in M_K.
M_K°	the non-archimedean absolute values in M_K.
K_v	the completion of K at the absolute value $v \in M_K$.
H_K, h_K	height on $\mathbb{P}^N(K)$ relative to K; see Definitions 4.4 and 4.11.
H, h	absolute height on $\mathbb{P}^N(\bar{K})$; see Definition 4.5.

FIGURE 2. Notation for global fields that we will use

(b) For L as in (a), prove that

$$\operatorname{rank} A(L) = \operatorname{rank} A(K) + \operatorname{rank} A'(K).$$

Hint: Construct homomorphisms

$$A(K) \oplus A'(K) \longrightarrow A(L) \quad \text{and} \quad A(L) \longrightarrow A(K) \oplus A'(K)$$

with finite kernels.

Exercise 3.E. Let A/K be an abelian variety such that $\boldsymbol{\mu}_m \subseteq \operatorname{Aut}(A/K)$ as $\operatorname{Gal}(K^{\text{sep}}/K)$-modules, and for $D \in K^*/K^{*m}$, let $A^{(D)}/K$ be the associated twist of A as described in Example 3.29. Suppsoe that m and D satisfy

$$\bigl[K(\sqrt[m]{D}) : K\bigr] = m.$$

Prove that

$$\operatorname{rank} A\bigl(K(\sqrt[m]{D})\bigr) = \sum_{i=0}^{m-1} \operatorname{rank} A^{(D^i)}(K).$$

Hint: Try the case $m = 2$ first. Construct a homomorphism

$$A\bigl(K(\sqrt{D}\,)\bigr) \longrightarrow A(K) \times A^{(D)}(K)$$

with finite kernel and cokernel.

Exercise 3.F. Let A be an abelian variety such that $\operatorname{Aut}(A)$ contains a copy of the mth roots of unity $\boldsymbol{\mu}_m$. Let $P \in A$ and $1 \ne \zeta \in \boldsymbol{\mu}_m$. Prove that

$$[\zeta]P = P \quad \Longrightarrow \quad [m]P = 0.$$

4. Background Material: Height Functions

For this section we set the following notation:

Definition 4.1 (Absolute Values on Number Fields)**.** The *standard set of absolute values on* $\mathbb{Q}$ is the set $M_{\mathbb{Q}}$ consisting of one archimedean absolute value

$$|\alpha|_\infty := \max\{\alpha,\ -\alpha\},$$

and for each prime p, one non-archimedean absolute value

$$|\alpha|_p := p^{-r}, \quad \text{where } \alpha = p^r a/b \in \mathbb{Q} \text{ with } a,b \in \mathbb{Z} \text{ and } \gcd(ab,p) = 1.$$

Let $K/\mathbb{Q}$ be a number field. We let M_K be the set of distinct absolute values on K that extend the absolute values in $M_{\mathbb{Q}}$.

Definition 4.2. The *valuation* associated to an absolute value $v \in M_K$ is the function

$$v : K^* \longrightarrow \mathbb{Z}, \quad v(\alpha) = -\log|\alpha|_v.$$

We sometimes extend the definition by setting $v(0) = \infty$.

We state two standard results from algebraic number theory.

Proposition 4.3. *Let $L/K\mathbb{Q}$ be number fields.*

(a) **[Product Formula]** *Let $\alpha \in K$. Then*

$$\prod_{v \in M_K} |\alpha|_v^{[K_v:\mathbb{Q}_v]} = \begin{cases} 1 & \text{if } \alpha \ne 0, \\ 0 & \text{if } \alpha = 0. \end{cases}$$

(b) **[Extension Formula]** *Let $v \in M_K$. Then*

$$\frac{1}{[L:K]} \sum_{\substack{w \in M_L \\ w \mid v}} [L_w : \mathbb{Q}_w] = [K_v : \mathbb{Q}_v].$$

Proof. See for example [48, II§1 and V§1]. □

Definition 4.4. The *logarithmic Weil height* on $\mathbb{P}^N(K)$ relative to the field K is the function

$$h_K : \mathbb{P}^n(K) \longrightarrow [0,\infty),$$
$$h_K\bigl([\alpha_0,\ldots,\alpha_n]\bigr) = \sum_{v \in M_K} [K_v : \mathbb{Q}_v] \cdot \log\max\bigl\{|\alpha_0|_v,\ |\alpha_1|_v,\ \ldots,\ |\alpha_n|_v\bigr\}.$$

Definition 4.5. The *absolute logarithmic Weil height* on $\mathbb{P}^N(\bar{K})$ is the function

$$h : \mathbb{P}^n(\bar{K}) \longrightarrow [0,\infty), \quad h(P) = \frac{1}{[L:\mathbb{Q}]} h_L(P),$$

where L is any number field over which P is defined, i.e., any number field for which we can write $P = [\alpha_0,\ldots,\alpha_n]$ with $\alpha_0,\ldots,\alpha_n \in L$.

Proposition 4.6. (a) *The Weil height h_K is well-defined and non-negative on $\mathbb{P}^N(K)$.*
(b) *The Weil height h is well-defined and non-negative on $\mathbb{P}^N(\bar{K})$.*

Proof. (a) Let $[\alpha_0,\ldots,\alpha_N]\in\mathbb{P}^N(K)$ and let $\gamma\in K^*$. Then the sum in Definition 4.4 for $[\alpha_0,\ldots,\alpha_N]$ and for $[\gamma\alpha_0,\ldots,\gamma\alpha_N]$ differ by

$$\sum_{v\in M_K}[K_v:\mathbb{Q}_v]\cdot\log|\gamma|_v=0,$$

where the equality comes from taking the logarithm of the product formula (Proposition 4.3(a)).

Now that we know that $h(P)$ doesn't depend on the choice of homogeneous coordinates, we can start with $P=[\alpha_0,\ldots,\alpha_N]$, choose a non-zero coordinate α_j, and divide all of the coordinates by α_j. We may thus assume that at least one of the coordinates of P is equal to 1. Since $|1|_v=1$ for every $v\in M_K$, it follows that

$$\log\max\bigl\{|\alpha_0|_v,\,|\alpha_1|_v,\,\ldots,\,|\alpha_N|_v\bigr\}\ge 0,$$

so $h_K(P)$ is a sum of non-negative terms. Hence $h_K(P)\ge 0$.
(b) Let $P=[\alpha_0,\ldots,\alpha_N]\in\mathbb{P}^N(K)$, and suppose that we view P as lying in $\mathbb{P}^N(L)$. Then

$$\begin{aligned}
h_L(P)&=\sum_{w\in M_L}[L_w:\mathbb{Q}_w]\cdot\log\max\bigl\{|\alpha_0|_w,\,|\alpha_1|_w,\,\ldots,\,|\alpha_n|_w\bigr\}\\
&\qquad\qquad\text{from Definition 4.4,}\\
&=\sum_{v\in M_K}\sum_{\substack{w\in M_L\\ w|v}}[L_w:\mathbb{Q}_w]\cdot\log\max\bigl\{|\alpha_0|_w,\,|\alpha_1|_w,\,\ldots,\,|\alpha_n|_w\bigr\}\\
&=\sum_{v\in M_K}\sum_{\substack{w\in M_L\\ w|v}}[L_w:\mathbb{Q}_w]\cdot\log\max\bigl\{|\alpha_0|_v,\,|\alpha_1|_v,\,\ldots,\,|\alpha_n|_v\bigr\}\\
&\qquad\qquad\text{since }\alpha_i\in K\text{ and }w\mid v,\\
&=\sum_{v\in M_K}\log\max\bigl\{|\alpha_0|_v,\,|\alpha_1|_v,\,\ldots,\,|\alpha_n|_v\bigr\}\sum_{\substack{w\in M_L\\ w|v}}[L_w:\mathbb{Q}_w]\\
&=\sum_{v\in M_K}\log\max\bigl\{|\alpha_0|_v,\,|\alpha_1|_v,\,\ldots,\,|\alpha_n|_v\bigr\}\cdot[K_v:\mathbb{Q}_v]\cdot[L:K]\\
&\qquad\qquad\text{by the extension formula (Proposition 4.3(b)),}\\
&=[L:K]\cdot h_K(P)\quad\text{from Definition 4.4.}
\end{aligned}$$

Hence using $[L:K]=[L:\mathbb{Q}]/[K:\mathbb{Q}]$, we find that

$$\frac{1}{[L:\mathbb{Q}]}h_L(P)=\frac{1}{[K:\mathbb{Q}]}h_K(P).$$

This concludes the proof that the definition of the absolute height does not depend on the choice of field.

Finally, the non-negativity of $h(P)$ follows from the non-negativity of $h_K(P)$ proven in (a). $\square$

Definition 4.7. Let X be a smooth projective variety defined over K. For each divisor $D \in \operatorname{Div}(X)$, we choose very ample divisors $D_1, \dots, D_r$ and integers $n_1, \dots, n_r$ so that

$$D \sim n_1 D_1 + \cdots + n_r D_r, \tag{4.1}$$

and for each $1 \le i \le r$, we choose projective embeddings

$$f_{D_i} : X \hookrightarrow \mathbb{P}^{\ell(D_i)-1} \tag{4.2}$$

as described in Definition 2.1. We then define the *absolute logarithmic Weil height function on X relative to the divisor D* to be the function

$$h_{X,D} : X(\bar{K}) \longrightarrow \mathbb{R}, \quad h_{X,D}(P) = \sum_{i=1}^{r} n_i h\bigl(f_{D_i}(P)\bigr).$$

Theorem 4.8 (Weil height machine)**.** *Let X be a smooth projective variety defined over K. In the following, the implied $O(1)$ constants depend on the indicated quantities, as well as on the choice of particular Weil height functions.*[6]

(a) (Linear Equivalence) *Let $D, D' \in \operatorname{Div}(X)$ be a divisor satisfying*

$$D' \sim D,$$

i.e., D and D' are linearly equivalent. Then

$$h_{X,D'} = h_{X,D} + O_{X,D,D'}(1).$$

where as indicated, the implied bounds depend on X, D, and D'. In particular, a Weil height $h_{X,D}$ is determined by X and D up the a function that is bounded.

(b) (Functoriality) *Let $\varphi : X \to Y$ be a morphism between two smooth projective varieties, and let $D \in \operatorname{Div}(Y)$. Then*

$$h_{X,\varphi^* D} = h_{Y,D} \circ \varphi + O_{X,Y,\varphi}.$$

(c) (Additivity) *Let $D_1, D_2 \in \operatorname{Div}(X)$. Then*

$$h_{X,D_1+D_2} = h_{X,D_1} + h_{X,D_2} + O_{X,D_1,D_2}(1).$$

(d) (Finiteness—Northcott Property) *Suppose that D is ample. Then for all $c, d > 0$, the set*

$$\Bigl\{P \in X(\bar{K}) : \bigl[K(P) : K\bigr] \le d \text{ and } h_{X,D}(P) \le c\Bigr\} \quad \text{is finite.}$$

[6] In other words, if we say that there is a quantity that is $O_{X,D}(1)$, then it is bounded by a quantity that depends on X and on the choice of Weil height function $h_{X,D}$ as described in Definition 4.7.

(e) (Algebraic Equivalence) *Let $D' \in \operatorname{Div}(X)$ be a divisor satisfying*

$$D' \equiv D,$$

i.e., the divisors D and D' are algebraically equivalent and suppose that D is ample. Then[7]

$$\lim_{\substack{P\in X(\bar K)\\ h_{X,D}(P)\to\infty}} \frac{h_{X,D'}(P)}{h_{X,D}(P)} = 1.$$

(f) (Positivity) *Let $D \in \operatorname{Div}(X)$ be an effective divisor.*[8] *Then there is a constant $C_3(D)$ so that*

$$h_{X,D}(P) \ge -C_3(D) \quad \textit{for all } P \in X(\bar K) \smallsetminus (\textit{base locus of } D).$$

Remark 4.9. One way to view the height machine is that it translates geometry into arithmetic. Picard groups, i.e., divisors and their linear equivalences, contain a lot of geometric information about varieties, heights contain a lot of arithmetic information about algebraic points on varieties, and the height maching translates relations between divisor classes into relations between heights.

Remark 4.10. A helpful reformulation of the Weil height machine is that it is the unique homomorphism

$$h_{X,\cdot} : \operatorname{Pic}(X) \longrightarrow \frac{\{\text{functions } X(\bar K) \to \mathbb{R}\}}{\{\text{bounded functions } X(\bar K) \to \mathbb{R}\}}$$

that satisfies:

- **Functoriality**. Let $\varphi : X \to Y$ be a morphisms and $D \in \operatorname{Div}(Y)$. Then $h_{X,\varphi^*D} = h_{Y,D} \circ \varphi + O(1)$.
- **Normalization**. Let h be the Weil height described in Definitions 4.4 and 4.5, and let $H \in \operatorname{Div}(\mathbb{P}^N)$ be a hyperplane. Then $h_{\mathbb{P}^N,H} = h + O(1)$.

4.1. **Heights over Function Fields.** For this subsection, we set the following notation:

k	a field
C/k	a smooth projective curve
$k(C)$	the function field of C

[7]With more work, one can prove the stronger estimate

$$h_{X,D'} = h_{X,D} + O_{X,D,D'}\bigl(\sqrt{|h_{X,D}|}\bigr).$$

[8]A divisor D is *effective* if $D = \sum n_i D_i$, where all of the D_i are irreducible codimension-1 subvarieties and all of the n_i are non-negative. The *base locus of D* is the intersection of the supports of all of the effective divisors that are linearly equivalent to D. For example, if D is very ample, then its base locus is empty.

The field $k(C)$ is very similar to a number field, so we can define height functions using $k(C)$. This can be done with absolute values as in Definition 4.1, but we will take a more geometric approach, using the fact that each element $\alpha \in k(C)$ is a rational function on C. More generally, any list of elements $\alpha_0, \ldots, \alpha_N \in k(C)$, not all 0, defines both a point

$$\boldsymbol{\alpha} = [\alpha_0, \ldots, \alpha_N] \in \mathbb{P}^N\bigl(k(C)\bigr)$$

and a k-morphism[9]

$$P_{\boldsymbol{\alpha}} : C \longrightarrow \mathbb{P}^N, \quad P_{\boldsymbol{\alpha}}(\gamma) = \bigl[\alpha_0(\gamma), \ldots, \alpha_N(\gamma)\bigr].$$

Definition 4.11. The *logarithmic Weil height* on $\mathbb{P}^N\bigl(k(C)\bigr)$ relative to the field $k(C)$ is the function[10]

$$h_{k(C)} : \mathbb{P}^n\bigl(k(C)\bigr) \longrightarrow [0,\infty), \quad h_{k(C)}(\boldsymbol{\alpha}) = \deg P_{\boldsymbol{\alpha}}^* \mathcal{O}_{\mathbb{P}^N}(1).$$

The *absolute logarithmic Weil height* on $\mathbb{P}^N\bigl(\overline{k(C)}\bigr)$ is the function

$$h : \mathbb{P}^N\bigl(\overline{k(C)}\bigr) \longrightarrow [0,\infty), \quad h(\boldsymbol{\alpha}) = \frac{1}{\bigl[k(C_\alpha) : k(C)\bigr]} h_{k(C_{\boldsymbol{\alpha}})}(\boldsymbol{\alpha}),$$

where $C_{\boldsymbol{\alpha}} \to C$ is any finite cover such that we can write $\boldsymbol{\alpha}$ using coordinates in $k(C_{\boldsymbol{\alpha}})$.

Example 4.12. The Weil height for the field $k(T) \cong k(\mathbb{P}^1)$ of rational functions is particularly easy to describe, since the PID $k[T]$ plays the same role for $k(T)$ that the PID $\mathbb{Z}$ plays for $\mathbb{Q}$. Thus given $\alpha_0(T), \ldots, \alpha_N(T) \in k(T)$, we can factor out common factors in their numerators and find a least common denominator for their denominators. In other words, we can find a non-zero $\beta(T) \in k(T)$ so that for $0 \le i \le N$ we have

$$\gamma_i(T) := \beta(T)\alpha_i(T) \in k[T], \quad \text{and} \quad \gcd\bigl(\gamma_0(T), \ldots, \gamma_N(T)\bigr) = 1.$$

Then

$$h_{k(T)}(P_{\mathbf{a}}) = \max_{0 \le i \le T} \deg \gamma_i(T).$$

In particular, the height of a rational function

$$\alpha(T) = \frac{a(T)}{b(T)} \quad \text{with } a(T), b(T) \in k[T] \text{ and } \gcd\bigl(a(T), b(T)\bigr) = 1$$

is

$$h\bigl(\alpha(T)\bigr) = \max\bigl\{\deg a(T), \deg b(T\bigr\}.$$

---

[9]In practice, we need to be a bit more careful when defining $P_{\boldsymbol{\alpha}}(\gamma)$. Thus we first choose an index i so that the quotient α_j/α_i is regular at γ for every j, and then we define $P_\alpha(\gamma)$ to be $\bigl[(\alpha_0/\alpha_i)(\gamma), \ldots, (\alpha_N/\alpha_i)(\gamma)\bigr]$.

[10]For those who prefer divisors to line bundles,we can also define $h_{k(C)}$ to be the intersection index $P_{\boldsymbol{\alpha}}(C) \cdot H$ of the image of C with a generic hyperplane H.

Example 4.13. Let $\alpha \in k(C)$, or equivalently, let $[\alpha, 1] \in \mathbb{P}^1\bigl(k(C)\bigr)$. Then α defines a map

$$\alpha : C \longrightarrow \mathbb{P}^1,$$

and the height of α is

$$h_{k(C)}(\alpha) = h_{k(C)}\bigl([\alpha, 1]\bigr) = \deg\Bigl(\alpha : C \to \mathbb{P}^1\Bigr).$$

Remark 4.14. If $K/\mathbb{Q}$ is a number field, then for any given bound B, Northcott's theorem (Theorem 4.8(d)) says that there are only finitely many points in $\mathbb{P}^N(K)$ whose height is smaller than c. This important finiteness property is false for $k(C)$ if k is an infinite field. Indeed, if $\alpha \in k$, then $[\alpha, 1] : C \to \mathbb{P}^1$ is the constant map, so

$$h\bigl([\alpha, 1]\bigr) = 0 \quad \text{for all } \alpha \in k.$$

Thus when working over a function field such as $k(C)$, proving that a set S of points is finite often requires two steps. First one shows that S is a set has bounded height. Then one proves that S has some sort of rigidity property that forces it to be finite.

Having defined the Weil height on $\mathbb{P}^N\bigl(k(C)\bigr)$, we can use Definition 4.7 to define the Weil height on smooth projective varieties X defined over $k(C)$, and then all of the properties of the Weil height described in Theorem 4.8 are true **except** that, as noted in Remark 4.14, the Northcott finiteness property may fail. In particular, we have the following relationship between heights and degrees of maps.

Proposition 4.15. *Let $\mathcal{X}$ be a projective variety, and let $\mathcal{X} \to C$ be a proper morphism whose generic fiber is a smooth projective variety $X/k(C)$. Let $D \in \operatorname{Div}(X)$, and let $\mathcal{D} \in \operatorname{Div}(\mathcal{X})$ be the closure of D. For each $k(C)$-rational point $\boldsymbol{\alpha} \in X\bigl(k(C)\bigr)$, we let $P_{\boldsymbol{\alpha}} : C \to \mathcal{X}$ be the associated section. Then the map*

$$X\bigl(k(C)\bigr) \longrightarrow \mathbb{R}, \quad \boldsymbol{\alpha} \longmapsto \deg P_{\boldsymbol{\alpha}}^*\mathcal{D},$$

is a Weil height function on $X(\bigl(k(C)\bigr)$ associated to the divisor D. Equivalently, if we fix a Weil height function $h_{X,D}$ on $X(\bigl(k(C)\bigr)$, then

$$\deg P_{\boldsymbol{\alpha}}^*\mathcal{D} = h_{A,D}(\boldsymbol{\alpha}) + O(1),$$

where the $O(1)$ is independent of $\boldsymbol{\alpha}$.

4.2. **Heights of (abelian) varieties.** Let X/K be a (smooth projective) variety defined over K. We would like to define the complexity, or height, of X/K. A naive definition would be that

$$h(X) = \begin{pmatrix} \text{the height of the coefficients of the} \\ \text{polynomials that are used to define } X \end{pmatrix}.$$

There are a number of problems with this definition. First, even if X is given with an embedding into projective space, there will generally be many choices for the polynomials that generate the ideal of X. Second, we'd prefer that the height of X not depend (too much) on the choice of an embedding. Third, we want isomorphic varieties to have more-or-less the same height. So here's a less naive definition that probably has the properties that we want:

$$h(X) = \min_{\substack{\text{embeddings}\\ \varphi:X\hookrightarrow\mathbb{P}^N}} \min_{\substack{\text{generators } f_1,\ldots,f_n\\ \text{for the ideal } I(\varphi(X))}} h\bigl([f_1,\ldots,f_n]\bigr).$$

Here we are writing $[f_1,\ldots,f_n]$ to denote the point in some large projective space whose coordinates are the coefficients of the polynomials $f_1,\ldots,f_n$. Unfortunately, although this definition of $h(X)$ captures the arithmetic complexity of X, it is difficult to use in practice.

Example 4.16 (The Height of an Elliptic Curve)**.** Every elliptic curve $E/\mathbb{Q}$ has a Weierstrass model of the form

$$E : y^2 = x^3+Ax+B \quad \text{with } A,B\in\mathbb{Z} \text{ and } \gcd(A^3,B^2) \text{ 12th-power-free.}$$

We can then define the height of $E/\mathbb{Q}$ to be $h\bigl([A,B,1]\bigr)$, although for homogeneity reasons, people often use

$$h_1(E/\mathbb{Q}) = h\bigl([A^3,B^2,1]\bigr).$$

More generally, for an elliptic curve E/K defined over a number field, we look at all Weierstrass equations having integral coefficients and define

$$h_1(E/K) = \min\left\{ h\bigl([A^3,B^2,1]\bigr) : \begin{array}{l} A,B\in R_K \text{ and } E \text{ is } K\text{-iso-}\\ \text{morphic to } y^2=x^3+Ax+B \end{array}\right\}. \tag{4.3}$$

Continuing to look at elliptic curves, we search for a more intrinsic measure of complexity. For example, we could use the height $h\bigl(j(E)\bigr)$ of the j-invariant of E. Unfortunately, this height doesn't have the desired finiteness property, since $j(E)$ only classifies E up to $\bar{K}$-isomorphism, so there are infinitely many K-isomorphism classes of elliptic curves with bounded j-height. This suggests adding a bit more data to the height, for example, the primes of bad reduction. One can show that if we define the height of E/K by

$$h_2(E/K) = h\bigl(j(E)\bigr) + \sum_{\substack{\text{primes } \mathfrak{p} \text{ such that}\\ E \text{ has bad reduction at } \mathfrak{p}}} \log \mathsf{N}_{K/\mathbb{Q}}(\mathfrak{p}), \tag{4.4}$$

then there there are only finitely many K-isomorphism classes of elliptic curves with bounded height. Indeed, one can show that there are positive constants so that these two heights are related by

$$C_4 h_1(E/K) \le h_2(E/K) \le C_5 h_1(E/K),$$

see Exercise 4.H.

We are going to generalize (4.4) to abelian varieties. Let $\mathcal{A}_g$ denote the moduli space of principally polarized abelian varieties of dimension g, fix an embedding $\mathcal{A}_g \hookrightarrow \mathbb{P}^N$ into some projective space, and let

$$h_{\mathcal{A}_g} : \mathcal{A}_g(\bar{\mathbb{Q}}) \longrightarrow [0,\infty) \tag{4.5}$$

be the Weil height associated to this embedding. We also write

$$j(A) \in \mathcal{A}_g(\bar{\mathbb{Q}}) \tag{4.6}$$

for the point in moduli space associated to a principally polarized abelian variety $A/\bar{\mathbb{Q}}$. (For example, if E is an elliptic curve, then $j(E) \in \mathbb{A}^1 \subset \mathbb{P}^1$ is the usual j-invariant of E.) It is tempting to define the height of A to be the height of $j(A)$, but $j(A)$ only captures the $\bar{\mathbb{Q}}$-isomorphism class of A. Hence all of the twists of A will have the same $j(A)$, which means that for a given number field $K/\mathbb{Q}$, there will be infinitely many A/K that are not K-isomorphic, yet have the same j-value. We can deal with the issue by also measuring the primes of bad reduction.

Definition 4.17. With the height $h_{\mathcal{A}_g}$ defined by (4.5) and the j-map defined by (4.6), we define the height of a principally polarized abelian variety A/K defined over a number field K to be

$$h(A/K) = h_{\mathcal{A}_g}\bigl(j(A)\bigr) + \sum_{\substack{0\ne\mathfrak{p}\in\operatorname{Spec}(R_K)\\ A\text{ has bad reduction at }\mathfrak{p}}} \log \mathsf{N}_{K/\mathbb{Q}}(\mathfrak{p}).$$

Supplementary Material 4.18 (**Faltings' height of an abelian variety**). Faltings [29, 30] used metrized line bundles to define a height function h_F on the space of abelian varieties. For A/K having everywhere semi-stable reduction, it satisfies

$$\Bigl|h_F(A) - h_{\mathcal{A}_g}(j(A))\Bigr| \le C_6 \log\Bigl(h_{\mathcal{A}_g}(j(A))\Bigr),$$

from which one sees that there are only finitely many semi-stable A/K with bounded $h_F(A)$. The Faltings' height function has nice functorial properties such as

$$h_F(A\times B) = h_F(A) + h_F(B) \quad\text{and}\quad h_F(\hat{A}) = h_F(A).$$

Faltings also proves a formula for the difference of the heights of isogenous abelian varieties in terms of the arithmetic properties of the isogeny. This formula plays a crucial role in his proof of Tate's isogeny conjecture, which in turn he uses to prove the Shafarevich and Mordell conjectures.

Exercises for Section 4.

Exercise 4.A. Let $P \in \mathbb{P}^n(\mathbb{Q})$, and write P as

$$P = [a_0, a_1, \ldots, a_n] \quad \text{with } a_0, \ldots, a_n \in \mathbb{Z} \text{ and } \gcd(a_0, \ldots, a_n) = 1.$$

Prove directly from the definition 4.11 that

$$h(P) = \log \max\bigl\{|a_0|,\, |a_1|,\, \ldots,\, |a_n|\bigr\}.$$

Exercise 4.B. (a) Prove the product formula (Proposition 4.3(a)), first for $K = \mathbb{Q}$, and then for arbitrary number fields.

(b) Prove the extension formula (Proposition 4.3(b)).

Exercise 4.C. Let C/K be a smooth projective curve, and let $D_1, D_2 \in \operatorname{Div}(C)$ be divisors with $\deg(D_1) \ge 1$. Prove that

$$\lim_{\substack{t \in C(\bar K)\\ h_C(t)\to\infty}} \frac{h_{C,D_2}(t)}{h_{C,D_1}(t)} = \frac{\deg D_2}{\deg D_1}.$$

Hint: Divisors on a curve C are algebraically equivalent if and only if they have the same degree.

Exercise 4.D. Let $\beta \in \bar{\mathbb{Q}}$ with $\beta \ne 0$, and fix a minimal polynomial

$$F_\beta(X) = a_0X^d + a_1X^{d-1} + \cdots + a_d \in \mathbb{Z}[X] \quad \text{with } \gcd(a_0, \ldots, a_d) = 1.$$

Factor F_β over $\mathbb{C}$ as

$$F_\beta(X) = a_0(X - \beta_1)(X - \beta_2)\ldots(X - \beta_d).$$

Prove that

$$h\bigl([\beta, 1]\bigr) = \frac{1}{d}\left(\log|a_0| + \sum_{i=1}^{d} \max\bigl\{|\beta_i|,\, 1\bigr\}\right).$$

Exercise 4.E. Let K be a number field, and let $F(X) \in K[X]$ be a polynomial of degree $d \ge 1$ that factors completely over K, say

$$F(X) = a_0X^d + a_1X^{d-1} + \cdots + a_d = (X - \alpha_1)(X - \alpha_2)\ldots(X - \alpha_d).$$

Prove that

$$-d \cdot \log(2) \le h\bigl([a_0, \ldots, a_d]\bigr) - \sum_{i=1}^{d} h(\alpha_i) \le (d-1)\log(2).$$

This gives an explicit estimate relating the height of the coefficients of a polynomial to the heights of its roots. *Hint*: Prove by induction on d a similar estimate for each $v \in M_K$, and then sum over $v \in M_K$.

Exercise 4.F. Let $m \ge 2$, let $D \in \mathbb{Z}$ be an integer that is mth-power-free, and let

$$P = [1, \alpha_1, \ldots, \alpha_N] \in \mathbb{P}^N(\bar{\mathbb{Q}})$$

be a point whose coordinates generate $\mathbb{Q}(D^{1/m})$, i.e.,

$$\mathbb{Q}(\alpha_1, \ldots, \alpha_N) = \mathbb{Q}(D^{1/m}).$$

Prove that

$$h(P) \ge \frac{1}{2m-2}\left(\frac{1}{m}\log|D| - m\log m\right).$$

Hint: Try $N = 1$ and/or $m = 2$ first.

Exercise 4.G. Let $K/\mathbb{Q}$ be a number field, and let $\Delta \in K^*$. We define the *m-power free height* of Δ to be

$$h_K^{(m)}(\Delta) := \min_{\beta \in K^*} h_K(\beta^m \cdot \Delta).$$

Prove that[11]

$$h_K^{(m)}(\Delta) \asymp \log \mathsf{N}_{K/\mathbb{Q}}\Bigl(\operatorname{Disc}\bigl(K(\Delta^{1/m})/K\bigr)\Bigr) \quad \text{for all } \Delta \in K^*,$$

where the implied constants may depend on K and m.

Exercise 4.H. Let h_1 and h_2 be the two height functions on the space of elliptic curves defined by (4.3) and (4.3). Prove that there are positive constants C_{11} and C_{12} so that for all number field E/K and all elliptic curves E/K we have

$$C_{11}h_1(E/K) \le h_2(E/K) \le C_{12}h_1(E/K).$$

Exercise 4.I. (a) Give an example of a dominant rational map

$$\varphi : \mathbb{P}^2 \dashrightarrow \mathbb{P}^2$$

such that functoriality (Theorem 4.8(b)) fails.

(b) Let $\varphi : \mathbb{P}^2 \dashrightarrow \mathbb{P}^2$ be a dominant rational map of degree d. Prove that there are constants $C_{13}(\varphi) > 0$ and $C_{14}(\varphi) \ge 0$ and a non-empty Zariski open set $U_\varphi \subset \mathbb{P}^2$ so that[12]

$$h\bigl(\varphi(P)\bigr) \ge C_{13}(\varphi)\cdot h(P) - C_{14}(\varphi) \quad \text{for all } P \in U(\bar{K}). \tag{4.7}$$

(c) Prove that the constant C_{13} in (b) may be chosen to depend only on the degree of φ.

(d) Generalize (c) to $\mathbb{P}^N$ with a constant $C_{13}(N,d)$ that depends only on the dimension of $\mathbb{P}^N$ and the degree d of the map φ.

(e) The *height expansion ratio for degree d maps of* $\mathbb{P}^N$ is, roughly speaking, the best possible value for the constant $C_{13}(N,d)$ in (d). More precisely, we define

$$\overline{\mu}_d(\mathbb{P}^N) = \inf_{\substack{\varphi:\mathbb{P}^N \dashrightarrow \mathbb{P}^N\\ \varphi \text{ dominant}\\ \deg(\varphi)=d}} \;\sup_{\emptyset \ne U \subset \mathbb{P}^N}\; \liminf_{\substack{P \in U(\bar{\mathbb{Q}})\\ h(P)\to\infty}} \frac{h\bigl(\varphi(P)\bigr)}{h(P)}.$$

---

[11] The notation $f(x) \asymp g(x)$ means that there are positive constants that may depend on f and g so that

$$C_7 f(x) - C_8 \le g(x) \le C_9 f(x) + C_{10} \quad \text{for all } x.$$

[12] If φ is a morphism, then Theorem 4.8 says that we may take $C_{13}(\varphi) = \deg(\varphi)$.

Let $d \ge 2$. Prove that

$$\overline{\mu}_d(\mathbb{P}^1) = d \quad \text{and} \quad \overline{\mu}_d(\mathbb{P}^N) \le \frac{1}{d^{N-1}} \quad \text{for } N \ge 2.$$

(f) Find a formula for $\overline{\mu}_d(\mathbb{P}^N)$ as a function of N and d. *Hint*: For $N \ge 2$, this is an open problem!! See [89].

5. Canonical heights: Construction and Basic Properties

Definition 5.1. Let A be an abelian variety, and let $D \in \operatorname{Div}(A)$ be a divisor. We say that

$$\begin{array}{lll} D \text{ is } \textit{symmetric} & \text{if} & [-1]^*D \sim D. \\ D \text{ is } \textit{anti-symmetric} & \text{if} & [-1]^*D \sim -D. \end{array}$$

Theorem 5.2. (Néron–Tate) *Set the following notation*:

K	*a number field or a function field, with algebraic closure* $\bar{K}$.
A/K	*an abelian variety defined over* K.
D	$\in \operatorname{Div}(A)$, *a divisor that is either symmetric or anti-symmetric.*
ρ	*the quantity* $\rho = \begin{cases} 2 & \text{if } D \text{ is symmetric,} \\ 1 & \text{if } D \text{ is anti-symmetric.} \end{cases}$
P	$\in A(\bar{K})$, *a* $\bar{K}$*-rational point on* A.

(a) *The following limit converges*:

$$\hat{h}_{A,D}(P) := \lim_{n\to\infty} \frac{1}{2^{\rho n}} h_{A,D}\bigl([2^n](P)\bigr). \tag{5.1}$$

(b) *There is a constant* $C(A,D)$ *so that limit defined in* (a) *satisfies*

$$\left|\hat{h}_{A,D}(P) - h_{A,D}(P)\right| \le C(A,D) \quad \textit{for all } P \in A(\bar{K}). \tag{5.2}$$

More precisely, the constant depends on the choice of the Weil height function $h_{A,D}$.

(c) *For all* $m \in \mathbb{Z}$, *the limit defined in* (a) *satisfies*

$$\hat{h}_{A,D}\bigl([m](P)\bigr) = m^\rho \hat{h}_{A,D}(P) \quad \textit{for all } P \in A(\bar{K}).$$

(d) *Let* $D, D' \in \operatorname{Div}(A)$. *Then*

$$D \sim D' \quad \Longrightarrow \quad \hat{h}_{A,D} = \hat{h}_{A,D'}.$$

In other words, the function $\hat{h}_{A,D}$ *depends only on the line bundle* $\mathcal{L}(D) \in \operatorname{Pic}(A)$.

Proof. (a) We give Tate's proof via a telescoping sum. The theorem of the cube (Corollary 3.8 with $m = 2$) tells us that in general we have $[2]^*D \sim 3D + [-1]^*D$, so the assumption that D is symmetric, respectively anti-symmetric, yields

$$[2]^*D \sim 2^\rho D. \tag{5.3}$$

We start with the estimate

$$\begin{aligned} h_{A,D}\circ[2] &= h_{A,[2]^*D}+O(1) && \text{functoriality of heights (Theorem 4.8(b)),}\\ &= h_{A,2^\rho D}+O(1) && \text{from (5.3) and the linear equivalence}\\ & && \text{property of heights (Theorem 4.8(a)),}\\ &= 2^\rho h_{A,D}+O(1) && \text{additivity of heights (Theorem 4.8(c))} \end{aligned}$$

We rewrite this big-O estimate as

$$\Bigl|h_{A,D}\bigl([2](Q)\bigr)-2^\rho h_{A,D}(Q)\Bigr|\le C(A,D)\quad\text{for all } Q\in A(K),\tag{5.4}$$

where the constant depends on A and D, but is independent of $Q\in A(K)$.

We are going to show that the sequence

$$\Bigl(2^{-\rho n}h_{A,D}\bigl([2^n](P)\bigr)\Bigr)_{n\ge0}\tag{5.5}$$

is a Cauchy sequence. To do this, we let $n\ge m\ge0$ and compute

$$\begin{aligned} &\left|\frac{1}{2^{\rho n}}h_{A,D}\bigl([2^n](P)\bigr)-\frac{1}{2^{\rho m}}h_{A,D}\bigl([2^m](P)\bigr)\right|\\ &\quad=\left|\sum_{i=m}^{n-1}\left(\frac{1}{2^{\rho(i+1)}}h_{A,D}\bigl([2^{i+1}](P)\bigr)-\frac{1}{2^{\rho i}}h_{A,D}\bigl([2^i](P)\bigr)\right)\right| \quad\text{telescoping sum,}\\ &\quad=\sum_{i=m}^{n-1}\frac{1}{2^{\rho(i+1)}}\Bigl|h_{A,D}\bigl([2^{i+1}](P)\bigr)-2^\rho h_{A,D}\bigl([2^i](P)\bigr)\Bigr| \quad\text{triangle inequality,}\\ &\quad\le\sum_{i=m}^{n-1}\frac{1}{2^{\rho(i+1)}}\cdot C(A,D)\quad\text{using (5.4) with } Q=[2^i](P),\\ &\quad\le\frac{C(A,D)}{2^{\rho m}(2^\rho-1)}. \end{aligned}\tag{5.6}$$

The upper bound $C(A,D)/2^m$ in (5.6) goes to 0 as $m\to\infty$ (with $n\ge m$), which completes the proof that the sequence (5.5) is Cauchy, and hence that it converges.

(b) Taking $m=0$ in (5.6) gives

$$\left|\frac{1}{2^{\rho n}}h_{A,D}\bigl([2^n](P)\bigr)-h_{A,D}(P)\right|\le\frac{C(A,D)}{2^\rho-1},\tag{5.7}$$

and then letting $n\to\infty$ gives (5.2).

(c) We compute

$$
\begin{aligned}
\hat{h}_{A,D}\bigl([m](P)\bigr)
&= \lim_{n\to\infty} \frac{1}{2^{\rho n}} h_{A,D}\bigl([2^n m](P)\bigr) \quad \text{by definition of } \hat{h}_{A,D},\\
&= \lim_{n\to\infty} \frac{1}{2^{\rho n}} h_{A,D}\bigl([m]\circ[2^n](P)\bigr)\\
&= \lim_{n\to\infty} \frac{1}{2^{\rho n}} \Bigl(h_{A,[m]^*D}\bigl([2^n](P)\bigr) + O_{A,D,m}(1)\Bigr) \quad \text{from Theorem 4.8(b)},\\
&= \lim_{n\to\infty} \frac{1}{2^{\rho n}} \Bigl(h_{A,m^\rho D}\bigl([2^n](P)\bigr) + O_{A,D,m}(1)\Bigr)\\
&\qquad\qquad \text{from Theorem 4.8(a) and Corollary 3.8},\\
&= \lim_{n\to\infty} \frac{1}{2^{\rho n}} \Bigl(m^\rho h_{A,D}\bigl([2^n](P)\bigr) + O_{A,D,m}(1)\Bigr) \quad \text{from Theorem 4.8(c)},\\
&= m^\rho \hat{h}_{A,D}(P) \quad \text{since } m \ge 2 \text{ and } O_{A,D,m}(1) \text{ is independent of } n.
\end{aligned}
$$

(d) Theorem 4.8(a) tells us that there is a constant $C(A,D,D')$ so that

$$\Bigl| h_{A,D}(Q) - h_{A,D'}(Q) \Bigr| \le C(A,D,D') \quad \text{for all } Q \in A(\bar{K}).$$

We substitute $Q = [2^n](P)$, divide by $2^{\rho n}$, and let $n \to \infty$, which gives the desired result $\hat{h}_{A,D}(P) = \hat{h}_{A,D'}(P)$. □

Definition 5.3. Let K be a number field or a function field, let A/K be an abelian variety defined over K, let $D \in \operatorname{Div}(A)$, and let $D^- = [-1]^*D$. We define the *canonical height*

$$\hat{h}_{A,D} : A(K) \longrightarrow \mathbb{R}$$

by

$$\hat{h}_{A,D}(P) := \frac{1}{2}\hat{h}_{A,D+D^-}(P) + \frac{1}{2}\hat{h}_{A,D-D^-}(P),$$

where $\hat{h}$ for the symmetric divisor $D+D^-$ is defined by (5.1) in Proposition 5.2(a) with $\rho = 2$, and $\hat{h}$ for the anti-symmetric divisor $D - D^-$ is defined by (5.1) in Proposition 5.2(a) with $\rho = 1$. We note that Theorem 5.2(d) implies that the canonical height function $\hat{h}_{A,D}$ depends only on the linear equivalence class of D, i.e., $\hat{h}_{A,D}$ depends only on the associated line bundle $\mathcal{L}(D)$

Theorem 5.4. (Néron–Tate) *Let K be a number field or a function field, let A/K be an abelian variety defined over K, let $D \in \operatorname{Div}(A)$ be a symmetric divisor, and let*

$$\hat{h}_{A,D} : A(\bar{K}) \longrightarrow \mathbb{R}$$

be the associated canonical height function.

(a) *The function* $\hat{h}_{A,D}$ *is a quadratic form*[13] *on* $A(\bar{K})$.

(b) **Number Field Case**. *Let* K *be a number field, and assume that* D *is an ample symmetric divisor.*

(b-1) *Let* $P \in A(K)$. *Then*

- $\hat{h}_{A,D}(P) \ge 0$.
- $\hat{h}_{A,D}(P) = 0$ *if and only if* $P \in A(K)_{\text{tors}}$.

(b-2) $\hat{h}_{A,D}$ *is a positive definite quadratic form on* $A(K)/A(K)_{\text{tors}}$.

(b-3) *More generally,* $\hat{h}_{A,D}$ *extends to a positive definite quadratic form*[14] *on* $A(K) \otimes \mathbb{R}$.

(c) **Function Field Case**. *Let* $K = k(C)$ *be a function field, let* $(B/k, \varphi)$ *be a* K/k*-trace*[15] *for* A/K, *and assume that* D *is an ample symmetric divisor.*

(c-1) *Let* $P \in A(K)$. *Then*

- $\hat{h}_{A,D}(P) \ge 0$.
- $\hat{h}_{A,D}(P) = 0$ *if and only if* $P \in \varphi(B(k)) + A(K)_{\text{tors}}$.

(c-2) $\hat{h}_{A,D}$ *is a positive definite quadratic form on*

$$A(K)/\bigl(\varphi(B(k)) + A(K)_{\text{tors}}\bigr);$$

(c-3) *More generally,* $\hat{h}_{A,D}$ *extends to a positive definite quadratic form on*

$$\bigl(A(K)/\varphi(B(k))\bigr) \otimes \mathbb{R}.$$

Proof. (a) We already know from Theorem 5.2(c) and our assumption that D is symmetric that $\hat{h}_{A,D}\bigl([m]P\bigr) = m^2\hat{h}_{A,D}(P)$. So it remains to prove bilinearity of the pairing on $A(K)$ defined by the formula

$$A(\bar{K}) \times A(\bar{K}) \longrightarrow \mathbb{R}, \quad (P,Q) \longmapsto \hat{h}_{A,D}(P+Q) - \hat{h}_{A,D}(P) - \hat{h}_{A,D}(Q).$$

[13]In general, if R is a commutative ring and M and N are R-modules, then a *quadratic form* from M to N is a function $Q : M \to N$ such that $Q(r\alpha) = r^2 Q(\alpha)$ and such that the following map is R-bilinear:

$$M \times M \longrightarrow N, \quad (\alpha, \alpha') = Q(\alpha + \alpha') - Q(\alpha) - Q(\alpha').$$

If $N = \mathbb{R}$, then Q is *positive definite* if $Q(\alpha) = 0$ implies that $\alpha = 0$.

[14]See Section 6 for details on how to extend the quadratic form $\hat{h}_{A,D}$ from $A(K)$ to $A(K) \otimes \mathbb{R}$.

[15]See Definition 3.22 in Section 3.2 for the definition of the trace.

By symmetry of the definition, it suffices to show that the pairing is linear in the first variable. We thus need to prove that

$$\begin{aligned}\hat h_{A,D}(P+Q+R) &- \hat h_{A,D}(P+Q) - \hat h_{A,D}(R)\\ &\overset{?}{=} \Bigl(\hat h_{A,D}(P+R) - \hat h_{A,D}(P) - \hat h_{A,D}(R)\Bigr)\\ &\quad + \Bigl(\hat h_{A,D}(Q+R) - \hat h_{A,D}(Q) - \hat h_{A,D}(R)\Bigr).\end{aligned}$$

Moving everything to one side of the equation, we need to prove that

$$\begin{aligned}\hat h_{A,D}(P+Q+R) - \hat h_{A,D}(P+Q) - \hat h_{A,D}(P+R) - \hat h_{A,D}(Q+R)&\\ + \hat h_{A,D}(P) + \hat h_{A,D}(Q) + \hat h_{A,D}(R) \overset{?}{=} 0.&\end{aligned} \tag{5.8}$$

We apply the theorem of the cube (Theorem 3.7) to the three projection maps

$$\pi_1, \pi_2, \pi_3 : A^3 \longrightarrow A.$$

This yields the divisor class relation

$$\begin{aligned}(\pi_1+\pi_2+\pi_3)^*D - (\pi_1+\pi_2)^*D - (\pi_1+\pi_3)^*D - (\pi_2+\pi_3)^*D&\\ + \pi_1^*D + \pi_2^*D + \pi_3^*D \sim 0.&\end{aligned} \tag{5.9}$$

For notational convenience, we let

$$\pi_{123} = \pi_1+\pi_2+\pi_3 \quad\text{and}\quad \pi_{ij} = \pi_i + \pi_j \quad \text{for } i,j \in \{1,2,3\}.$$

This allows us to rewrite (5.9) more succintly as

$$\pi_{123}^*D - \pi_{12}^*D - \pi_{13}^*D - \pi_{23}^*D + \pi_1^*D + \pi_2^*D + \pi_3^*D \sim 0. \tag{5.10}$$

Applying the linear equivalence part of the height machine (Theorem 4.8(a)) to the linear equivalence (5.10), we obtain

$$\begin{aligned}&h_{A,\pi_{123}^*D}(P,Q,R)\\ &\quad - h_{A,\pi_{12}^*D}(P,Q,R) - h_{A,\pi_{13}^*D}(P,Q,R) - h_{A,\pi_{23}^*D}(P,Q,R)\\ &\quad + h_{A,\pi_1^*D}(P,Q,R) + h_{A,\pi_2^*D}(P,Q,R) + h_{A,\pi_3^*D}(P,Q,R) = O(1).\end{aligned}$$

We stress that the $O(1)$ is independent of the point $(P,Q,R) \in A^3(K)$. We next use functoriality of heights (Theorem 4.8(b)) to get

$$\begin{aligned}&h_{A,D}\bigl(\pi_{123}(P,Q,R)\bigr)\\ &\quad - h_{A,D}\bigl(\pi_{12}(P,Q,R)\bigr) - h_{A,D}\bigl(\pi_{13}(P,Q,R)\bigr) - h_{A,D}\bigl(\pi_{23}(P,Q,R)\bigr)\\ &\quad + h_{A,D}\bigl(\pi_1(P,Q,R)\bigr) + h_{A,D}\bigl(\pi_2(P,Q,R)\bigr) + h_{A,D}\bigl(\pi_3(P,Q,R)\bigr) = O(1).\end{aligned}$$

Using the definition of the various π maps, this becomes

$$\begin{aligned}h_{A,D}(P+Q+R) - h_{A,D}(P+Q) - h_{A,D}(P+R) - h_{A,D}(Q+R)&\\ + h_{A,D}(P) + h_{A,D}(Q) + h_{A,D}(R) = O(1).&\end{aligned}$$

Now all that remains is to replace

$$(P, Q, R) \text{ with } \bigl([2^n](P), [2^n](Q), [2^n](R)\bigr),$$

divide by 4^n, and let $n \to \infty$. Then the Weil heights become canonical heights, while the $O(4^{-n})$ disappears, which completes the proof of the desired result (5.8).

(b-1) We choose an integer $m \geq 1$ so that mD is very ample, and we let $f_{mD} : A \hookrightarrow \mathbb{P}^{\ell(D)-1}$ be a projective embedding associated to mD as described in Definition 2.1. Then Definition 4.7 says that

$$h_{A,mD}(P) = h\bigl(f_D(P)\bigr)$$

is a Weil height associated to mD. Additivity of heights (Theorem 4.8(c)) gives

$$h_{A,D}(P) = \frac{1}{m} \cdot h\bigl(f_D(P)\bigr) + O(1),$$

and Proposition 4.6(b) tells us that $h\bigl(f_D(P)\bigr) \geq 0$. It follows that there is a constant $C_{15} \geq 0$ so that

$$h_{A,D}(P) \geq -C_{15} \quad \text{for all } P \in A(K).$$

Replacing P with $[2^n]P$, dividing by 4^n, and letting $n \to \infty$ yields

$$\hat{h}_{A,D}(P) = \lim_{n\to\infty} 4^{-n} h_{A,D}\bigl([2^n]P\bigr) \geq \lim_{n\to\infty} -4^{-n} C_{15} = 0.$$

This proves the first assertion that $\hat{h}_{A,D}$ is a non-negative function.

For the second assertion, we start with the computation

$$\begin{aligned} P \in A(K)_{\text{tors}} \quad &\Longrightarrow \quad [n]P = 0 \text{ for some } n \geq 1, \\ &\Longrightarrow \quad 0 = \hat{h}_{A,D}(0) = \hat{h}_{A,D}\bigl([n]P\bigr) = n^2 \hat{h}_{A,D}(P) \\ &\Longrightarrow \quad \hat{h}_{A,D}(P) = 0. \end{aligned}$$

Conversely, suppose that $\hat{h}_{A,D}(P) = 0$. Then for all $n \geq 0$ we have

$$\hat{h}_{A,D}\bigl([n]P\bigr) = n^2 \hat{h}_{A,D}(P) = 0.$$

We know that the canonical height $\hat{h}_{A,D}$ and the Weil height $h_{A,D}$ differ by a bounded amount (Theorem 5.2(b)), so we can find a constant C_{16} so that

$$h_{A,D}(Q) \leq \hat{h}_{A,D}(Q) + C_{16} \quad \text{for all } Q \in A(K).$$

In particular, putting $Q = [n]P$ yields

$$h_{A,D}\bigl([n]P\bigr) \leq C_{16} \quad \text{for all } n \geq 0.$$

Hence

$$\bigl\{[n]P : n \geq 0\bigr\} \subseteq \bigl\{Q \in A(K) : h_{A,D}(Q) \leq C_{16}\bigr\}. \tag{5.11}$$

We are given that the divisor D is ample, so the finiteness property of heights (Theorem 4.8(d)) tells us that sets of bounded D-height are finite. Applying this fact to (5.11) implies that

$$\bigl\{[n]P : n \ge 0\bigr\} \quad \text{is a finite set,}$$

so by the pigeon-hole principle, we can find integers $n_2 > n_1$ such that $[n_2]P = [n_1]P$. Hence

$$[n_2 - n_1]P = 0,$$

so $P \in A(K)_{\text{tors}}$.
(b-2) Let $P \in A$ and $T \in A_{\text{tors}}$, and choose $n \ge 1$ so that $[n]T = 0$. Then

$$\begin{aligned}\hat{h}_{A,D}(P+T) &= n^{-2}\hat{h}_{A,D}\bigl([n](P+T)\bigr)\\ &= n^{-2}\hat{h}_{A,D}\bigl([n]P + [n]T\bigr) = n^{-2}\hat{h}_{A,D}\bigl([n]P\bigr) = \hat{h}_{A,D}(P).\end{aligned}$$

This proves that $\hat{h}_{A,D}$ is well-defined on the quotient $A(K)/A(K)_{\text{tors}}$, and we know from (a) that $\hat{h}_{A,D}$ is a quadratic form. Further, it follows from (b-1) that that $\hat{h}_{A,D}$ is non-negative, and that when viewed as a function on $A(K)/A(K)_{\text{tors}}$, it vanished only at 0. Hence $\hat{h}_{A,D}$ is a positive definite quadratic form on $A(K)/A(K)_{\text{tors}}$.
(b-3) We refer the reader to [39, Proposition B.5.3(b)] for the proof. We also note that the postive definiteness of $\hat{h}_{A,D}$ on $A(K) \otimes \mathbb{R}$ is not an immediate consequence of its positive definiteness on $A(K)/A(K)_{\text{tors}}$; see Exercise 5.D.
(c) Much of the proof in the function field case mirrors the number field proof. The primary difference arises from the fact that the field K may have infinitely many elements of bounded height. Indeed, this will always be the case if the base field k is infinite, since elements of k have height 0; see Remark 4.14. However, one can show that on an abelian variety $A/k(C)$, infinite sets of bounded height come only from the presence of points "defined over the base field k." More precisely, if D is an ample divisor, then

$$\text{Image}\Bigl(\bigl\{P \in A(\bigl(k(C)\bigr) : h_{A,D}(P) \le T\bigr\} \longrightarrow A\bigl(k(C)\bigr)/\varphi\bigl(B(k)\bigr)\Bigr) \quad \text{is finite.} \tag{5.12}$$

The finiteness statement (5.12) is used in the proof of the function field Mordell–Weil theorem (Theorem 3.25) and in proving properties of the canonical height in the function field setting. We do not include a proof (5.12), but refer the reader to [47, Chapter 6] for a proof in a very general setting, and to [85, Chapter III] for the simpler elliptic surface case. □

Definition 5.5. Let K be a number field or a function field with algebraic closure $\bar{K}$, let A/K be an abelian variety defined over K, and let $D \in \operatorname{Div}(A)$ be a divisor. The *Néron–Tate pairing* (or *canonical height pairing*) is the function[16]

$$\begin{gathered}\langle\,\cdot\,,\,\cdot\,\rangle_{A,D} : A(\bar{K}) \times A(\bar{K}) \longrightarrow \mathbb{R}, \\ \langle P, Q\rangle_{A,D} = \frac{1}{2}\Bigl(\hat{h}_{A,D}(P+Q) - \hat{h}_{A,D}(P) - \hat{h}_{A,D}(Q)\Bigr).\end{gathered}$$

Definition 5.6. Let K be a number field or a function field $K = k(C)$, let A/K be an abelian variety defined over K, let $D \in \operatorname{Div}(A)$ be an ample symmetric divisor, and let $P_1, \ldots, P_r \in A(K)$ be a set of points. The *Néron–Tate regulator of* $P_1, \ldots, P_r$ is the quantity

$$\operatorname{Reg}_D(P_1, \ldots, P_r) = \det\Bigl(\langle P_i, P_j\rangle_{A,D}\Bigr)_{1\le i,j\le r}.$$

Let $P_1, \ldots, P_r \in A(K)$ be a basis for the appropriate torsion-free finitely generated abelian group, i.e.,

$$\begin{aligned}&\operatorname{Span}_{\mathbb{Z}}(P_1, \ldots, P_r) \\ &\qquad = \begin{cases} A(K)/A(K)_{\text{tors}} & \text{if } K \text{ is a number field,} \\ A(K)/\bigl(\varphi(B(k)) + A(K)_{\text{tors}}\bigr) & \text{if } K \text{ is a function field,}\end{cases}\end{aligned}$$

where $(B/k, \phi)$ is the K/k-trace in the function field case. We then define the *Néron–Tate regulator of* A/K to be

$$\operatorname{Reg}_D(A/K) = \operatorname{Reg}_D(P_1, \ldots, P_r).$$

(If $r = 0$, we set $\operatorname{Reg}_D(A/K) = 1$.) We note that Theorem 5.4(b) and (b$'$) tell us that $\operatorname{Reg}_D(A/K) > 0$.

Supplementary Material 5.7 (**Values of the canonical height**). The algebraic properties of canonical heights on abelian varieties over number fields are largely unknown. One might guess that if $P \in A(K)$ is a non-torsion point and D is an ample divisor, then $\hat{h}_{A,D}(P)$ and $\exp\bigl(\hat{h}_{A,D}(P)\bigr)$ are transcendental over $\mathbb{Q}$, but there does not seem to be even a single example in which it is known that one of these values is not in $\mathbb{Q}$. More generally, one might ask:

Question 5.8. Let $K/\mathbb{Q}$ be a number field, let $A_1/K, \ldots, A_n/K$ be geometrically simple non-isogenous abelian varieties, let $D_1, \ldots, D_n$ be ample divisors on $A_1, \ldots, A_n$, and for each $1 \le i \le n$, let $P_{i,1}, \ldots, P_{i,r_i} \in A_i(K)$ be $\mathbb{Z}$-linearly independent points. Is it true that the transcendence degree of

$$K\Bigl(\hat{h}_{A_i,D_i}(P_{i,j}) : 1 \le i \le n,\ 1 \le j \le r_i\Bigr) \text{ over } K \text{ equals } \sum_{i=1}^{n} r_i\ ?$$

Ditto with $\exp(\hat{h})$, or even adjoining both $\hat{h}$ and $\exp(\hat{h})$ values.

Supplementary Material 5.9 (**Canonical heights of subvarieties of an abelian variety**). We have defined the canonical height of a point on an abelian variety. Philippon [73, 74, 75] has shown how to define more generally the canonical height on an arbitrary subvariety of an abelian variety. Let $D \in \operatorname{Div}(A)$ be a very ample symmetric divisor with an associated embedding $\varphi_D : A \hookrightarrow \mathbb{P}^N$,

[16]We include the factor of $\frac{1}{2}$ so that $\langle P, P\rangle_{A,D} = \hat{h}_{A,D}(P)$.

let $X \subset A$ be a subvariety, and let $G_X = \{P \in A : X + P = X\}$ be the stabilizer of X. Then the limit

$$\hat{h}_{A,D}(X) := \lim_{n\to\infty} \frac{\#\bigl(A[n] \cap G_X\bigr)}{n^{2\dim(X)+2}} \cdot h\bigl(\varphi_D\bigl([n]X\bigr)\bigr)$$

exists, and the resulting function has the following properties:

(a) $\hat{h}_{A,mD}(X) = m^{\dim(X)+1}\hat{h}_{A,D}(X)$.
(b) $\hat{h}_{A,D}(X+T) = \hat{h}_{A,D}(X)$ for all $T \in A_{\text{tors}}$.
(c) Let $\alpha \in \operatorname{End}(A)$ satisfy $\alpha^*D \sim qD$. Then

$$\hat{h}_{A,D}\bigl(\alpha(X)\bigr) = \frac{q^{\dim(X)+1}}{\#\bigl(\ker(\alpha) \cap G_X\bigr)} \cdot \hat{h}_{A,D}(X).$$

(d) $\hat{h}_{A,D}(X) = 0$ if and only if there exist an abelian subvariety $B \subseteq A$ and a torsion point $T \in A_{\text{tors}}$ such that $X = B + T$.

See [39, Section F.2] for further information and pointers to the literature.

Supplementary Material 5.10 (**Polarized dynamical systems**). The theory of canonical heights can be generalized to iterates of maps on more general varieties [9]. A *polarized dynamical system* is a triple (X, φ, D) consisting of a smooth projective variety $X/\bar{\mathbb{Q}}$, a finite morphism $\varphi : X \to X$, and a divisor $D \in \operatorname{Div}(X) \otimes \mathbb{R}$ satisfying $\varphi^*D \sim \kappa D$ for some real number $\kappa > 1$, where $\sim$ denotes linear equivalence. The *canonical height* associated to φ and D is the function $\hat{h}_{X,\varphi,D} : X(\bar{\mathbb{Q}}) \to \mathbb{R}$ defined by the limit

$$\hat{h}_{X,\varphi,D}(P) = \lim_{n\to\infty} \frac{1}{\kappa^n} h_D\bigl(\varphi^{\circ n}(P)\bigr).$$

The telescoping sum argument described in Theorem 5.2(a) carries over to this more general setting, and the proofs of Theorem 5.2(b) and (c) yield

$$\hat{h}_{X,\varphi,D}(P) = h_{X,D}(P) + O_{\varphi,D}(1) \quad \text{and} \quad \hat{h}_{X,\varphi,D}\bigl(\varphi(P)\bigr) = (\deg\varphi)\hat{h}_{X,\varphi,D}(P) + O_{\varphi,D}(1)$$

If D is ample, then $\hat{h}_{X,\varphi,D}(P) = 0$ if and only if P is preperiodic for φ.

Exercises for Section 5.

Exercise 5.A. Let $P, Q \in A(K)$.

(a) Prove that

$$P - Q \in A(K)_{\text{tors}} \quad \Longrightarrow \quad \hat{h}_{A,D}(P) = \hat{h}_{A,D}(Q).$$

(b) Let $K/\mathbb{Q}$ be a number field. Is the converse to (a) true? I do not know any counterexamples!

Exercise 5.B. Let A/K be an abelian variety, and let $D \in \operatorname{Div}(A)$ be an anti-symmetric divisor, i.e., $[-1]^*D \sim -D$. Prove that the map

$$\hat{h}_{A,D} : A(\bar{K}) \longrightarrow \mathbb{R}$$

is linear, i.e., prove that $\hat{h}_{A,D}(P+Q) = \hat{h}_{A,D}(P) + \hat{h}_{A,D}(Q)$.

Exercise 5.C. Let A/K be an abelian variety, and let $D \in \operatorname{Div}(A)$ be a (not necessarily symmetric or anti-symmetric) divisor on A.

(a) Prove that the Néron–Tate pairing $\langle\,\cdot\,,\,\cdot\,\rangle_{A,D}$ as given in Definition 5.5 is a symmetric bilinear pairing.
(b) Let $D, D' \in \operatorname{Div}(A)$. Prove that

$$\langle\,\cdot\,,\,\cdot\,\rangle_{A,D+D'} = \langle\,\cdot\,,\,\cdot\,\rangle_{A,D} + \langle\,\cdot\,,\,\cdot\,\rangle_{A,D'}.$$

(c) Define the *symmetrization of* D to be

$$D^\sigma := \frac{1}{2}\bigl(D + [-1]^*D\bigr).$$

Prove that D^σ is symmetric, and that

$$\langle\,\cdot\,,\,\cdot\,\rangle_{A,D^\sigma} = \langle\,\cdot\,,\,\cdot\,\rangle_{A,D}.$$

(In fancier terminology, this shows that the Néron–Tate pairing for D depends on only the algebraic equivalence class of D, i.e., on the image of D in the Néron–Severi group $\operatorname{NS}(A)$.)

Exercise 5.D. . Consider the map

$$q : \mathbb{Z}[\sqrt{2}] \longrightarrow \mathbb{R}, q(a+b\sqrt{2}) = |a+b\sqrt{2}|^2,$$

where we view $\mathbb{Z}[\sqrt{2}]$ as a free $\mathbb{Z}$-module of rank 2.

(a) Prove that q is a positive definite quadratic form on $\mathbb{Z}[\sqrt{2}]$.
(b) Prove that the extension of q to $\mathbb{Z}[\sqrt{2}] \otimes_{\mathbb{Z}} \mathbb{R}$ is not positive definite.

Exercise 5.E. Let A/K and B/K be abelian varieties, let $\varphi : B \to A$ an isogeny, and let $D \in \operatorname{Div}(A)$. Prove that

$$\hat{h}_{A,D}\bigl(\varphi(P)\bigr) = \hat{h}_{B,\varphi^*D}(P) \quad \text{for all } P \in B(\bar{K}).$$

Exercise 5.F. We fix an ample divisor $H \in \operatorname{Div}(A)$, and we use H to define the Rosatti involution

$$\operatorname{End}(A)_{\mathbb{Q}} \longrightarrow \operatorname{End}(A)_{\mathbb{Q}}, \quad \alpha \longmapsto \alpha' := \varphi_H^{-1} \circ \hat{\alpha} \circ \varphi_H,$$

and a map

$$\operatorname{NS}(A)_{\mathbb{Q}} \longrightarrow \operatorname{End}(A)_{\mathbb{Q}}, \quad D \longmapsto \varphi_H^{-1} \circ \varphi_D;$$

see Definition 3.18. Prove that the canonical height pairing satisfies the following two formulas:

(a) $\bigl\langle \alpha(P),\, Q\bigr\rangle_{A,D} = \bigl\langle P,\, \alpha'(Q)\bigr\rangle_{A,H}$.
(b) $\langle P, Q\rangle_{A,D} = \bigl\langle P,\, \Phi_D(Q)\bigr\rangle_{A,H}$.

(These formulas are proven in [5, Proposition 3] and [43, Propositions 27 & 28], where they are then applied to solve various problems.)

Exercise 5.G. Learn how to compute canonical heights on elliptic curves using a computer algebra system such as Magma, Sage, or PARI-GP. *Hint*: If you ask ChatGPT "How do I compute the canonical height on an elliptic curve using XXX," it will give you some sample code that may or may not actually work.

Exercise 5.H. This exercise asks you to compute some canonical heights. Feel free to use a computer algebra system; see Exercise 5.G. We consider the elliptic curve

$$E : y^2 = x^3 + 17$$

and the points

$$P=(-2,3),\qquad Q=(2,5),\qquad R=(-1,4),$$
$$S=(8,23),\qquad T=(52,375),\qquad U=(5234,378661),$$
$$V=\left(\frac{94}{25},\frac{1047}{125}\right),\qquad W=\left(\frac{19}{25},\frac{522}{125}\right).$$

Verify that $P,Q,\dots,W$ are points in $E(\mathbb{Q})$.

(a) Compute the canonical heights of the points $P,Q,\dots,W$.

(b) Compute the ratios

$$\frac{\hat{h}_E(Q)}{\hat{h}_E(P)},\quad \frac{\hat{h}_E(R)}{\hat{h}_E(P)},\quad \frac{\hat{h}_E(S)}{\hat{h}_E(P)},\quad \frac{\hat{h}_E(W)}{\hat{h}_E(P)}.$$

Draw some conclusions and check that your conclusions are correct.

(c) Compute the 2-by-2 height pairing matrix for P and Q and take its determinant. What can you conclude?

(d) Compute the 3-by-3 height pairing matrix for P, Q, and U and take its determinant. What can you conclude? Verify your conclusion with an explicit algebraic formula.

6. Applications to Counting Points

For this section we fix the following notation:

$K/\mathbb{Q}$	a number field.
A/K	an abelian variety of dimension $g\ge 1$.
D	$\in \operatorname{Div}_K(A)$, an ample symmetric divisor.
$\hat{h}_{A,D}$	the absolute logarithmic canonical height on $A(\bar{K})$.
$\langle\,\cdot\,,\,\cdot\,\rangle_{A,D}$	the associated canonical height pairing.
r	the rank of the Mordell–Weil group $A(K)$.

The Mordell–Weil group $A(K)$ is a finitely generated abelian group, so we can create an $\mathbb{R}$-vector space by tensoring with $\mathbb{R}$. We denote this vector space by

$$A(K)_{\mathbb{R}} := A(K)\otimes\mathbb{R}\cong\mathbb{R}^r.$$

It has dimension r, where $r=\operatorname{rank}A(K)$. The canonical height pairing (5.5) may be extended to the $\mathbb{R}$-vector space $A(K)_{\mathbb{R}}$ as follows. Every element of $A(K)_{\mathbb{R}}$ is a linear combination of elements of the form $P\otimes a$ with $P\in A(K)$ and $a\in\mathbb{R}$. Then for any $P_i,Q_j\in A(K)$ and $a_i,b_j\in\mathbb{R}$, we define

$$\left\langle \sum_{i=1}^r P_i\otimes a_i,\ \sum_{j=1}^r Q_j\otimes b_j\right\rangle_{A,D} := \sum_{i=1}^r\sum_{j=1}^r a_ib_j\langle P_i,Q_j\rangle_{A,D}.$$

Alternatively, let $P_1,\dots,P_r\in A(K)$ be generators for the free group

$$A(K)/A(K)_{\mathrm{tors}}\cong\mathbb{Z}^r.$$

Then $\{P_1,\dots,P_r\}$ is a basis for the $\mathbb{R}$-vector space $A(K)_{\mathbb{R}}$, and we use the matrix of real numbers

$$\Bigl(\langle P_i,P_j\rangle_{A,D}\Bigr)_{1\le i,j\le r}$$

to define a bilinear form on $A(K)_{\mathbb{R}}$ relative to this basis.

It is important to stress that the canonical height pairing on $A(K)$ is defined intrinsically in terms of the geometry and arithmetic of $A(K)$ and D; its definition does not depend on a choice of basis. So the pairing on $A(K)_{\mathbb{R}}$ is similarly intrinsic and does not depend on the choice of basis.[17]

Theorem 5.4(b) tells us that the associated quadratic form on $A(K)_{\mathbb{R}}$ is positive definite, and thus that the image of $A(K)$ in $A(K)_{\mathbb{R}}$ is a discrete subgroup (lattice) relative to this quadratic form. We set the (non-standard) notation

$$A(K)_{\mathbb{Z}}:=\operatorname{Image}\Bigl(A(K)\longrightarrow A(K)_{\mathbb{R}}\Bigr)$$

for the *Mordell–Weil lattice*, so we have an exact sequence

$$0\longrightarrow A(K)_{\mathrm{tors}}\longrightarrow A(K)\longrightarrow A(K)_{\mathbb{Z}}\longrightarrow 0,$$

and we define

$$\|\cdot\|_{A,D}:A(K)_{\mathbb{R}}\longrightarrow\mathbb{R},\quad \|P\|_{A,D}=\sqrt{\langle P,P\rangle_{A,D}},$$

for the norm on the vector space $A(K)_{\mathbb{R}}$ associated to the canonical height pairing.

It's now time to take advantage of these constructions to give a strong counting result for the points in $A(K)$.

Theorem 6.1. (Néron) *Fix a Weil height $h_{A,D}$ associated to an ample symmetric divisor D, and define a point counting function by*

$$\mathcal{N}\bigl(A(K),h_{A,D},T\bigr):=\#\bigl\{P\in A(K):h_{A,D}(P)\le T\bigr\}.$$

Let $r=\operatorname{rank}A(K)$. Then there is a constant $\alpha(A/K,D)>0$ such that

$$\mathcal{N}\bigl(A(K),h_{A,D},T\bigr)=\alpha(A/K,D)T^{r/2}+O(T^{(r-1)/2})\quad\text{as } T\to\infty.$$

[17]In general, a bilinear form $B:V\times V$ on a vector space may be represented by a matrix $\bigl(B(\mathbf{v}_i,\mathbf{v}_j)\bigr)$ after we choose a basis $\mathbf{v}_1,\dots,\mathbf{v}_r$ for V. But B doesn't depend on the choice of basis; and conversely, a matrix only determines a bilinear form on V after we specify a basis.

More precisely, we may take

$$\alpha(A/K,D) = \#A(K)_{\mathrm{tors}} \cdot \frac{\pi^{r/2}\Gamma(r/2+1)}{\mathrm{Reg}_D(A/K)^{1/2}},$$

where we note that $\pi^{r/2}\Gamma(r/2+1)$ *is the volume of the unit ball in* $\mathbb{R}^r$.

Proof Sketch. We first note that the counting functions for the given Weil height $h_{A,D}$ and the associated canonical height $\hat{h}_{A,D}$ are closely related, since Theorem 5.2(b) tells us that the difference $\hat{h}_{A,D} - h_{A,D}$ is bounded. It follows from Exercise 6.A that it thus suffices to prove the theorem for the counting function $\mathcal{N}\bigl(A(K),\hat{h}_{A,D},T\bigr)$ associated to the canonical height.

The kernel of the map

$$A(K) \longrightarrow A(K)_{\mathbb{Z}}$$

is $A(K)_{\mathrm{tors}}$, and we know from Exercise 5.A that

$$P \equiv Q \pmod{A(K)_{\mathrm{tors}}} \quad \Longrightarrow \quad \hat{h}_{A,D}(P) = \hat{h}_{A,D}(Q).$$

It follows that each point in $A(K)_{\mathbb{Z}}$ corresponds to $\#A(K)_{\mathrm{tors}}$ points having the same canonical height in $A(K)$, so we find that

$$\begin{aligned}\mathcal{N}\bigl(A(K),\hat{h}_{A,D},T\bigr) &= \#A(K)_{\mathrm{tors}} \cdot \mathcal{N}\bigl(A(K)_{\mathbb{Z}},\hat{h}_{A,D},T\bigr)\\ &= \#A(K)_{\mathrm{tors}} \cdot \mathcal{N}\bigl(A(K)_{\mathbb{Z}},\|\cdot\|_{A,D}^2,T\bigr)\\ &= \#A(K)_{\mathrm{tors}} \cdot \mathcal{N}\bigl(A(K)_{\mathbb{Z}},\|\cdot\|_{A,D},T^{1/2}\bigr).\end{aligned}$$

In summary, we are given the following material:

$$\begin{aligned} A(K)_{\mathbb{R}} &= \text{an } r\text{-dimensional real vector space.}\\ A(K)_{\mathbb{Z}} &= \text{a rank } r\text{-lattice in } A(K)_{\mathbb{R}}.\\ \|\cdot\|_{A,D} &= \hat{h}_{A,D}(\,\cdot\,)^{1/2} = \text{a Euclidean norm on } A(K)_{\mathbb{R}}.\\ \mathrm{Reg}_D(A/K)^{1/2} &= \text{the volume of a fundamental}\\ &\ \text{domain for } A(K)_{\mathbb{Z}} \text{ in } A(K)_{\mathbb{R}}.\end{aligned}$$

And we want to count the number of points in the lattice whose norm is smaller than $T^{1/2}$. This is now a lattice point counting problem that has nothing to do with abelian varieties or height functions, so the following result (Proposition 6.2) completes the proof. □

Proposition 6.2. *Let* V *be an* r*-dimensional* $\mathbb{R}$*-vector space, let* $\|\cdot\|$ *be a Euclidean norm on* V*, let* $L \subset V$ *be a lattice with fundamental domain* $\mathcal{F}_L$*, and let* $\mathcal{B} := \bigl\{\mathbf{v} \in V : \|\mathbf{v}\| \le 1\bigr\}$ *be the unit ball relative* *to the given norm. Then*

$$\#\{\mathbf{v} \in L : \|\mathbf{v}\| \le T\} = \frac{\mathrm{Vol}(\mathcal{B})}{\mathrm{Vol}(\mathcal{F}_L)} \cdot T^r + O(T^{r-1}),$$

where Vol *is the Lebesgue measure on* V *for some identification* $V \cong \mathbb{R}^r$*.*

Proof Idea. Let $\mathcal{B}(T) = \{\mathbf{v} \in V : \|\mathbf{v}\| \le T\}$ be the ball of radius T, so we want to count

$$\mathcal{N}(L,T) := \#\bigl(\mathcal{B}(T) \cap L\bigr).$$

We note that the vector space V is the disjoint union of the translated fundamental domains $\mathcal{F}_L + \mathbf{w}$ as $\mathbf{w}$ ranges over L. Let $d = \sup_{\mathbf{v}\in\mathcal{F}_L} \|\mathbf{v}\|$ be the length of the longest vector in the fundamental domain $\mathcal{F}_L$. Then one can check (left as an exercise for the reader) that

$$\mathcal{B}(T-d) \subseteq \bigcup_{\mathbf{w}\in L,\, \|\mathbf{w}\|\le T} \mathcal{F}_L + \mathbf{w} \subseteq \mathcal{B}(T+d). \tag{6.1}$$

Taking the volume of both sides and using the disjointedness of the translated fundamental domains gives

$$\mathrm{Vol}\bigl(\mathcal{B}(T-d)\bigr) \le \sum_{\mathbf{w}\in L,\, \|\mathbf{w}\|\le T} \mathrm{Vol}(\mathcal{F}_L + \mathbf{w}) \le \mathrm{Vol}\bigl(\mathcal{B}(T+d)\bigr).$$

The volume form is homogeneous and translation invariant, so

$$\mathrm{Vol}\bigl(\mathcal{B}(1)\bigr)(T-d)^r \le \mathcal{N}(L,T) \cdot \mathrm{Vol}(\mathcal{F}_L) \le \mathrm{Vol}\bigl(\mathcal{B}(1)\bigr)(T+d)^r.$$

Hence

$$\mathcal{N}(L,T) = \frac{\mathrm{Vol}\bigl(\mathcal{B}(1)\bigr)}{\mathrm{Vol}(\mathcal{F}_L)} T^r + O(T^{r-1}). \qquad \square$$

Exercises for Section 6.

Exercise 6.A. This exercise is about abstract counting functions. Let S be a set, let $h : S \to \mathbb{R}$ a non-negative valued function, and let

$$\mathcal{N}(S,h,T) = \#\{\alpha \in S : h(\alpha) \le T\}.$$

be the associated counting function. Suppose that there are real numbers $A > 0$ and $n \ge 1$ such that

$$\mathcal{N}(S,h,T) = AT^n + O(T^{n-1}) \quad \text{as } T \to \infty.$$

Now let $h' : S \to \mathbb{R}$ be another non-negative valued function such that there is a constant $C = C(S,h,h')$ with the property that

$$\bigl|h(\alpha) - h'(\alpha)\bigr| \le C \quad \text{for all } \alpha \in S.$$

Prove that

$$\mathcal{N}(S,h',T) = AT^n + O(T^{n-1}) \quad \text{as } T \to \infty,$$

where A and n are the same as for h, but the big-O constant will change and is allowed to depend on C.

Exercise 6.B. Prove the inclusions (6.1).

7. Local Canonical Heights

We recall that the height of a point $P = [\alpha_0, \ldots, \alpha_n] \in \mathbb{P}^n(K)$ in projective space is defined (Definition 4.4) as a sum

$$h_K\bigl([\alpha_0, \ldots, \alpha_n]\bigr) = \sum_{v \in M_K} [K_v : \mathbb{Q}_v] \cdot \log \max\bigl\{|\alpha_0|_v,\ |\alpha_1|_v,\ \ldots,\ |\alpha_n|_v\bigr\}$$

over the absolute values on K. Thus we might say that the "global height" $h(P)$ is a weighted sum of "local heights"

$$\log |P|_v = \log \max\bigl\{|\alpha_0|_v,\ |\alpha_1|_v,\ \ldots,\ |\alpha_n|_v\bigr\}.$$

However, we see immediately that there is a problem with this definition, because $\log |P|_v$ is not a function of P. It depends on a choice of homogeneous coordinates for P. It is only after taking the sum that the product formula (Proposition 4.3) eliminates the dependence on that choice.

So we take a different approach via measuring the v-adic distance from points to hypersurfaces. Let $D \subset \mathbb{P}^n$ be an irreducible divisor given by the vanishing $F(\mathbf{x}) = 0$ of a homogeneous polynomial of degree d. Then for $v \in M_K$ and for points $P \in \mathbb{P}^n(K)$, we define a *v-adic distance from P to D* using the formula[18]

$$\operatorname{dist}_v(P, D) = \min\left\{ \left|\frac{F(P)}{x_0(P)^d}\right|_v, \left|\frac{F(P)}{x_1(P)^d}\right|_v, \cdots, \left|\frac{F(P)}{x_n(P)^d}\right|_v \right\}. \tag{7.1}$$

We note that $\operatorname{dist}_v(P, D)$ does not depend on the choice of homogeneous coordinates for P, and although it does depend on the choice of the polynomial F, this dependence is not too severe. Indeed, if also $D = \{F' = 0\}$, then there are positive constants such that

$$C_{17} \operatorname{dist}_v(P, F) \le \operatorname{dist}_v(P, F') \le C_{18} \operatorname{dist}_v(P, F) \quad \text{for all } P. \tag{7.2}$$

Finally, since our heights are logarithmic, we define a v-adic local height associated to the divisor D by the formula

$$\lambda_{\mathbb{P}^n,D,v} : \mathbb{P}^n(K_v) \longrightarrow \mathbb{R} \cup \{\infty\}, \quad \lambda_{\mathbb{P}^n,D,v}(P) = -\log \operatorname{dist}_v(P, D).$$

Thus $\lambda_{\mathbb{P}^n,D,v}(P)$ is large when P is v-adically close to D, and one can check that

$$h_{\mathbb{P}^n,D}(P) = \frac{1}{[K : \mathbb{Q}]} \sum_{v \in M_K} [K_v : \mathbb{Q}_v] \cdot \lambda_{\mathbb{P}^n,D,v}(P) + O(1) \qquad \text{for } P \notin \operatorname{Support}(D). \tag{7.3}$$

[18]If $x_i(P) = 0$ and $F(P) \ne 0$, then we formally set $\bigl|F(P)/x_i(P)\bigr|_v = \infty$, while if $P \in D$, i.e., if $F(P) = 0$, then we set $\operatorname{dist}_v(P, D) = 0$.

With more work one can define v-adic distances and v-adic local heights for all smooth varieties X. These local heights have many nice properties, similar to the properties of global heights described by the Height Machine (Theorem 4.8), although just as with global Weil heights, they are only well-defined up to bounded functions. We refer the reader to [39, Theorem B.8.1] or [47, Chapter 10] for details.

Néron used the group law on an abelian variety to construct canonical local heights get rid of the indeterminate bounded functions. We summarize his main results.

Theorem 7.1. (Néron [69]) *For notational clarity in the statement of this theorem, we write $|D|$ for the support of the divisor D. Let A/K be an abelian variety. There exists a unique collection of functions*

$$\hat{\lambda}_{A,D,v} : A(\bar{K}_v) \smallsetminus |D| \longrightarrow \mathbb{R},$$

indexed by divisors $D \in \operatorname{Div}_K(A)$ and absolute values $v \in M_K$, so that the following are true:

(a) *The map $\hat{\lambda}_{A,D,v}$ is continuous, where we give $A(\bar{K}_v)$ the v-adic topology.*

(b) *For all $D, D' \in \operatorname{Div}_K(A)$ and all $v \in M_K$,*

$$\hat{\lambda}_{A,D+D',v} = \hat{\lambda}_{A,D,v} + \hat{\lambda}_{A,D',v} \quad \text{on } A(\bar{K}_v) \smallsetminus \bigl(|D| \cup |D'|\bigr).$$

(c) *For all morphisms $\varphi : A \to B$ of abelian varieties over K and all $D \in \operatorname{Div}_K(B)$,*

$$\hat{\lambda}_{A,\varphi^* D,v} = \hat{\lambda}_{B,D,v} \circ \varphi \quad \text{on } A(\bar{K}_v) \smallsetminus |\varphi^* D|.$$

(d) *For all rational functions $f \in K(A)$ and all $v \in M_K$, there is a constant $\gamma_{f,v}$ such that*

$$\hat{\lambda}_{A,\operatorname{div}(f),v} = v \circ f + \gamma_{f,v} \quad \text{on } A(\bar{K}_v) \smallsetminus \bigl|\operatorname{div}(f)\bigr|.$$

Further, there are only finitely many $v \in K$ with $\gamma_{f,v} \neq 0$.

(e) (Normalization) *For all $D \in \operatorname{Div}_K(A)$ and all $v \in M_K$, we have*[19]

$$\lim_{N\to\infty} N^{-2g} \sum_{\substack{P \in A[N] \\ P \notin |D|}} \hat{\lambda}_{A,D,v}(P) = 0. \tag{7.4}$$

[19]Without this normalization, which Néron did not impose in his original formulation, the function $\hat{\lambda}_{A,D,v}$ is only well-defined up to a constant, although that constant will be 0 for all but finitely many v. We also mention that if the absolute value on K is archimedean, then the normalization condition is equivalent to $\int_{A(\bar{K}_v)} \hat{\lambda}_{A,D,v}(P)\,d\mu(P) = 0$, where μ is Haar measure on $A(\bar{K}_v) \cong A(\mathbb{C})$.

(f) (Local-Global Decomposition) *There is a constant $\kappa(A,D)$ so that for all finite extensions L/K and all $P \in A(L) \smallsetminus |D|$,*

$$\hat{h}_{A,D}(P) = \frac{1}{[L:K]} \sum_{w \in M_L} [L_w : K_w] \cdot \hat{\lambda}_{A,D,w}(P) - \kappa(A,D).$$

Proof. See [47, Chapter 11] for a proof. There is also a discussion in [39, Theorem B.9.3]. □

Remark 7.2. Néron also gives explicit formulas for the local canonical height. For non-archimedean places of good reduction, the formula is given in terms of an intersection on the special fiber of the Néron model; indeed, Néron's original motivation for constructing what we now call the Néron model was for precisely this application to the theory of heights. For non-archimedean places of bad reduction, the intersection index is supplemented with a correction factor that causes many difficulties when trying to prove lower bounds for canonical heights. For archimedean places, the explicit formula for the local height uses complex analysis and is given in terms of the theta function associated to the divisor D. At the end of this section we briefly give supplemental material describing these formulas, with some notation and terminology left undefined.

Remark 7.3. The limit formula (7.4) is used to normalize the canonical local height functions $\hat{\lambda}_{A,D,v}$, but as noted in Theorem 7.1(f), the sum of the normalized local canonical heights may then differ from the global canonical height by a constant that we have denoted $\kappa(A,D)$. If $\dim(A) = 1$, i.e., if A is an elliptic curve, then one can prove that $\kappa(A,D) = 0$. However, in dimension 2 and greater, it is possible to have $\kappa(A,D) \neq 0$. See [36] for a discussion, and [53] for an application that could be significantly improved if we had better knowledge of this mysterious κ-constant.

Supplementary Material 7.4 (**A Soupçon of History**). In a short address at the 1958 ICM [67], Néron conjectured the existence of what is now known as the canonical, or Néron–Tate, height on abelian varieties. Tate constructed $\hat{h}_{A,D}$ using the telescoping sum argument that we have described. Néron then constructed local canonical heights $\hat{\lambda}_{A,D,v}$ using a variety of methods. This provided a new proof of the existence of $\hat{h}_{A,D}$ as the sum of the $\hat{\lambda}_{A,D,v}$, as well as giving much finer insight into $\hat{h}_{A,D}$. A first announcement of these results was provided by Lang [44] at a Séminaire Bourbaki talk in 1964. The review of Lang's talk by Cassels [11] details the sometimes convoluted dissemination and publication processes of the time. Cassels writes that Lang's note gives "*an account of two papers containing fundamental results in the theory of heights of points on algebraic varieties defined over global fields; the first, by John Tate ('non publié, comme d'habitude'), has already been partly published by proxy* (*by Manin* [55]); *. . . In the other paper, by Néron, which will doubtless be published in due course in the conventional way,. . .* ." Néron's paper [69] appeared in 1965.

Supplementary Material 7.5 (**Explicit Formulas for Local Canonical Heights**). As noted in Remark 7.2, Néron gave explicit formulas for local canonical heights. We summarize some of those formulas in the following theorem.

Theorem 7.6. *Let A/K be an abelian variety, let $D \in \operatorname{Div}(A)$, let $\mathcal{A}/R_K$ be the Néron model of A, and let $v \in M_K$.*

(a) (Good Reduction) *If $v \in M_K^\circ$ and A has good reduction at v, then*[20]

$$\hat{\lambda}_{A,D,v}(P) = \langle \overline{D} \cdot \overline{P} \rangle_{\mathcal{A},v} \tag{7.5}$$

is the intersection index in the fiber over v of the closures of D and P in $\mathcal{A}$. In particular, in the good reduction case we have

$$\hat{\lambda}_{A,D,v}(P) \ge 0 \quad \text{for all } P \in A(\bar{K}_v) \smallsetminus \operatorname{Support}(D).$$

(b) (Bad Reduction) *Let*

$$j_v : A(K) \longrightarrow (\mathcal{A}/\mathcal{A}^\circ)_v(k_v)$$

be the homomorphism that sends a point to its image in the group of components of the Néron model over v. Then there is a function[21]

$$\mathbb{B}_{D,v} : (\mathcal{A}/\mathcal{A}^\circ)_v(k_v) \longrightarrow \mathbb{R}$$

so that for all $P \in A(\bar{K}_v) \smallsetminus \operatorname{Support}(D)$, we have

$$\hat{\lambda}_{A,D,v}(P) = \langle \overline{D} \cdot \overline{P} \rangle_{\mathcal{A},v} + \mathbb{B}_{D,v}\bigl(j_v(P)\bigr) - \kappa(A,D,v). \tag{7.6}$$

Here $\kappa(A,D,v)$ is a constant that is chosen so that (7.4) holds.

(c) (Archimedean Absolute Values) *We let $\mathfrak{H}_g$ denote the Siegel upper half space consisting of all g-by-g matrices τ with complex entries such that $\operatorname{Im}(\tau)$ is positive definite. Each $\tau \in \mathfrak{H}_g$ determines a rank $2g$ lattice $\Lambda_\tau = \mathbb{Z}^g + \tau\mathbb{Z}^g \subset \mathbb{C}^g$, a complex torus $A_\tau = \mathbb{C}^g/\Lambda_\tau$, and a theta function*

$$\theta(\mathbf{z},\tau) = \sum_{\mathbf{m}\in\mathbb{Z}^g} \exp\bigl(\pi i \cdot {}^t\,\mathbf{m}\cdot\tau\cdot n + 2\pi i \cdot {}^t\,\mathbf{m}\cdot\mathbf{z}\bigr).$$

The theta function is not Λ_τ-invariant, but for $\boldsymbol{\lambda} \in \Lambda_\tau$, the ratio $\theta(\mathbf{z}+\boldsymbol{\lambda},\tau)/\theta(\mathbf{z},\tau)$ is a non-vanishing function, so

$$\Theta_\tau := \bigl\{\mathbf{z} \in A_\tau : \theta(\mathbf{z}) = 0\bigr\} \in \operatorname{Div}(A_\tau).$$

One can show that Θ_τ is a principal polarization of A_τ. Then the local canonical height on $A_\tau(\mathbb{C})$ is more-or-less $\log|\theta|$, but we need a small correction since θ is not itself a function on $A_\tau(\mathbb{C})$. Néron proved that

$$\begin{gathered}\hat{\lambda}_{A_\tau,\Theta_\tau} : A_\tau(\mathbb{C}) \smallsetminus |\Theta_\tau| \longrightarrow \mathbb{R}, \\ \hat{\lambda}_{A_\tau,\Theta_\tau}(\mathbf{z}) = -\log\bigl|\theta(\mathbf{z},\tau)\bigr| + \pi \cdot {}^t\,(\operatorname{Im}\mathbf{z})\cdot(\operatorname{Im}\tau)^{-1}\cdot(\operatorname{Im}\mathbf{z}) + \kappa_\tau,\end{gathered} \tag{7.7}$$

where κ_τ is chosen so that (7.4) holds.

Theta functions on elliptic curves have product expansions that lead to alternative formulas such as the following: Let $\tau \in \mathfrak{H}_1$ with associated elliptic curve $E_\tau(\mathbb{C}) \cong \mathbb{C}/(\mathbb{Z}+\tau\mathbb{Z})$. Set $q = e^{2\pi i\tau}$ and $u = e^{2\pi i z}$, and let

$$\mathbb{B}_2(x) = x^2 - x + \frac{1}{6}$$

be the second Bernoulli polynomial. Then

$$\hat{\lambda}_{E_\tau,(0)}(z) = -\frac{1}{2}\mathbb{B}_2\left(\frac{\operatorname{Im} z}{\operatorname{Im}\tau}\right)\log|q| - \log|1-u| - \sum_{n\ge1}\log\bigl|(1-q^n u)(1-q^n u^{-1}\bigr|. \tag{7.8}$$

---

[20]Néron proved that (7.5) is true up to a constant. See [8] for a proof that the average of the intersection multiplicities over torsion points goes to 0, which implies that the constant vanishes.

[21]Néron further proved that the values of $\mathbb{B}_{D,v}$ are rational numbers with denominators dividing $2\#(\mathcal{A}/\mathcal{A}^\circ)_v(k_v)$.

Exercises for Section 7.

Exercise 7.A. (a) Prove (7.2), which says that different equations for the divisor D give distance functions that are bounded by multiples of one another, where dist_v is defined by (7.1).

(b) Prove (7.3), which says that the global Weil height on $\mathbb{P}^n$ is the sum of the local heights.

Exercise 7.B. Aside from (a), this "exercise" is really a research question. Let A/K be a geometrically simple abelian variety, and let $D \in \text{Div}(A)$ be an ample effective divisor that gives a principal polarization on A. And even (a) is non-trivial to prove.

(a) If $\dim(A) = 1$, prove that $\kappa(A, D) = 0$.

(b) Is it true that $\kappa(A, D) \geq 0$? If not, is there a constant $C_{19}(g) > 0$ so that

$$\kappa(A,D) \geq -C_{19}(g) \quad \text{for all } A/K \text{ with } \dim(A) = g \text{ and for all ample effective divisors } D \in \text{Div}(A) \text{ that give a principal polarization on } A.$$

(c) If $\dim(A) \geq 2$, is it always true that $\kappa(A, D) \neq 0$?

(d) Let $\mathcal{A}_g$ be the moduli space of principally polarized abelian varieties of dimension g, and let $h_{\mathcal{A}_g}$ be a Weil height function on $\mathcal{A}_g$ associated to an ample divisor. Is $\kappa(A, D)$ related to the height $h_{\mathcal{A}_g}(A, D)$ of the point in the $\mathcal{A}_g(K)$ corresponding to the pair (A, D)? For example, might it be true that there are positive constants so that

$$C_{20}(g,\epsilon) \cdot h_{\mathcal{A}_g}(A,D)^{1-\epsilon} \leq \kappa(A,D) \leq C_{21}(g,\epsilon) \cdot h_{\mathcal{A}_g}(A,D)^{1+\epsilon}?$$

(e) Let C/K be a curve of genus 2, let J/K be the Jacobian variety of C, and let $\Theta \in \text{Div}(J)$ be the Θ-divisor (or more prosaically, let Θ be a copy of C sitting in J). What can you say about $\kappa(J, \Theta)$?

8. Lower Bounds for Canonical Heights

Let A/K be an abelian variety defined over a number field and let $D \in \text{Div}(A)$ be an ample symmetric divisor. The canonical height $\hat{h}_{A,D}(P)$ of a point $P \in A(\bar{K})$ is a measure of the arithmetic complexity of P. We proved in Theorem 5.4(b-1) that

$$\hat{h}_{A,D}(P) = 0 \quad \Longleftrightarrow \quad P \in A_{\text{tors}},$$

so in some sense, torsion points have minimal complexity. It is thus of interest to understand the complexity of the non-torsion points; in particular, to study how small can it be.

There are two directions that this may take us, both of which lead to many interesting theorems and open problems.

(1) We can fix the abelian variety A/K and take non-torsion points $P \in A(\bar{K})$ defined over larger and larger fields.

(2) We can fix the field K and vary over all abelian varieties A/K and all non-torsion points $P \in A(K)$.

We start with (1), for which we first briefly review the much-studied classical case where the abelian variety A is replaced by the multiplicative group $\mathbb{G}_m$. But (2) has no $\mathbb{G}_m$-analogue, since there is only one $\mathbb{G}_m$, so we jump straight to the abelian variety case.

8.1. **Height Lower Bounds as the Field Varies: Multiplicative Group.** We start with the classical (absolute logarithmic) Weil height

$$h : \bar{\mathbb{Q}} \longrightarrow \mathbb{R}, \quad h(\alpha) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M_K} [K_v : \mathbb{Q}] \cdot \max\{|\alpha|_v, 1\},$$

where $K/\mathbb{Q}$ is any finite extension with $\alpha \in K$. The Weil height might be said to be canonical relative to the multiplicative group $\mathbb{G}_m(\bar{\mathbb{Q}}) = \bar{\mathbb{Q}}^*$, since it satisfies

$$h(\alpha^n) = |n| h(\alpha) \quad \text{for all } n \in \mathbb{Z}.$$

It also clearly satisfies

$$h(\alpha) \ge 0 \quad \text{for all } \alpha \in \bar{\mathbb{Q}}.$$

Theorem 8.1. (Kronecker) *Let $\alpha \in \bar{\mathbb{Q}}^*$. Then*

$$h(\alpha) = 0 \quad \Longleftrightarrow \quad \alpha \text{ is a root of unity.}$$

Proof. If $\alpha^n = 1$ for some $n > 0$, then

$$0 = h(1) = h(\alpha^n) = n \cdot h(\alpha),$$

so $h(\alpha) = 0$. Conversely, suppose that $h(\alpha) = 0$. Then

$$h(\alpha^n) = |n| \cdot h(\alpha) = 0 \quad \text{for all } n \in \mathbb{Z},$$

so $\{\alpha^n : n \in \mathbb{Z}\}$ is a set of bounded height that is contained in the number field $\mathbb{Q}(\alpha)$. It follows from Northcott's theorem (Remark 4.14) that the set is finite. Hence we can find distinct integers $n > m$ such that $\alpha^n = \alpha^m$. Thus $\alpha^{n-m} = 1$ and $n - m \neq 0$, so α is a root of unity. (It is instructive to compare this proof with the analogous result for abelian varieties which states that $\hat{h}_{A,D}(P) = 0$ if and only if $P \in A(K)_{\text{tors}}$; see Theorem 5.4(b-1).) □

In view of Kronecker's theorem, it is natural to ask how small the height can be for non-roots of unity. The answer is arbitrarily small, since for example

$$h(2^{1/n}) = \frac{\log 2}{n}. \tag{8.1}$$

However, the number $2^{1/n}$ generates a field of degree n, so we can rewrite (8.1) as

$$h(2^{1/n}) = \frac{\log 2}{\bigl[\mathbb{Q}(2^{1/n}) : \mathbb{Q}\bigr]}. \tag{8.2}$$

Conjecture 8.2 (Lehmer Conjecture for $\mathbb{G}_m$)**.** *There is an absolute constant $C > 0$ such that if $\alpha \in \bar{\mathbb{Q}}^*$ is not a root of unity, then*

$$h(\alpha) \ge \frac{C}{\bigl[\mathbb{Q}(\alpha) : \mathbb{Q}\bigr]}.$$

Remark 8.3. Extensive computation [64] suggests that the best constant in Lehmer's conjecture is $\log(\alpha_0)$, where $\alpha_0 \approx 1.1762808$ is a real root of the polynomial

$$x^{10} + x^9 - x^7 - x^6 - x^5 - x^4 - x^3 + x + 1,$$

which was discovered by Mahler in 1933.

Theorem 8.4 (Dobrowolski [26])**.** *For every $\epsilon > 0$ the is a constant $C(\epsilon) > 0$ such that for all $\alpha \in \bar{\mathbb{Q}}^*$ that are not roots of unity, we have*[22]

$$h(\alpha) \ge \frac{C(\epsilon)}{\bigl[\mathbb{Q}(\alpha) : \mathbb{Q}\bigr]^{1+\epsilon}}.$$

8.2. **Height Lower Bounds as the Field Varies: Abelian Varieties.** Lehmer's conjecture has been extended to other settings, including elliptic curves, abelian varieties, and dynamical systems. We start by noting that Theorem 5.4(b-1), which we restate here, is the abelian variety analogue of Kronecker's theorem; see Theorem 8.1

Theorem 8.5. *Let $K/\mathbb{Q}$ be a number field, let A/K be an abelian variety, and let $D \in \operatorname{Div}_K(A)$ be an ample symmetric divisor. Then*

$$\hat{h}_{A,D}(P) = 0 \quad \Longleftrightarrow \quad P \textit{ is a torsion point.}$$

Example 8.6. We note that the example in (8.2) generalizes to abelian varieties. Thus fix a non-torsion point $P \in A(K)$ and, for each $n \ge 1$, let $P_n \in A(\bar{K})$ be a point satisfying $[n]P_n = P$. The Galois conjugates of P_n are contained in the set $P_n + A[n]$, and for large n we expect that

$$\bigl[K(P_n) : K\bigr] \approx \#A[n] = n^{2g}.$$

[22]Dobrowolski's result is stronger than this. Writing $d = \bigl[\mathbb{Q}(\alpha) : \mathbb{Q}\bigr]$ for notational convenience, he proves that there is an absolute constant $C > 0$ such that

$$h(\alpha) \ge \frac{C}{d} \cdot \left(\frac{\log\log d}{\log d}\right)^3.$$

This leads to the estimate

$$\hat{h}_{A,D}(P_n) = \frac{1}{n^2}\hat{h}_{A,D}\bigl([n]P_n\bigr) = \frac{1}{n^2}\hat{h}_{A,D}(P) \approx \frac{\hat{h}_{A,D}(P)}{\bigl[K(P_n):K\bigr]^{1/g}},$$

and thus to the following generalization of Lehmer's conjecture (Conjecture 8.2) to the setting of abelian varieties.

Conjecture 8.7 (Lehmer Conjecture for Abelian Varieties [22, 58])**.** *Let $K/\mathbb{Q}$ be a number field, let A/K be a geometrically simple abelian variety, and let $D \in \operatorname{Div}_K(A)$ be an ample symmetric divisor. Then there is a constant $C_{22}(A/K,D) > 0$ such that*

$$\hat{h}_{A,D}(P) \ge \frac{C_{22}(A/K,D)}{\bigl[K(P):K\bigr]^{1/g}} \quad \textit{for all non-torsion points } P \in A(\bar{K}).$$

We now state, without proof, the best general result currently known for abelian varieties, after which we describe some stronger results for certain classes of elliptic curves.

Theorem 8.8. *Let $K/\mathbb{Q}$ be a number field, let A/K be an abelian variety of dimension g, and let $D \in \operatorname{Div}_K(A)$ be an ample symmetric divisor.*

(a) (Masser [56, 58], see also [20, 32]) *For every $\epsilon > 0$ there is a constant $C_{23}(A/K,D,\epsilon) > 0$ such that*

$$P \in A(\bar{K}) \smallsetminus A_{\text{tors}} \quad \Longrightarrow \quad \hat{h}_{A,D}(P) \ge \frac{C_{23}(A/K,D,\epsilon)}{\bigl[K(P):K\bigr]^{2g+1+\epsilon}}. \tag{8.3}$$

(b) (David–Hindry [22], Ratazzi [76]) *If A has complex multiplication,*[23] *then the exponent in* (8.3) *can be replaced by $1+\epsilon$.*

For elliptic curves, i.e., for $\dim(A) = 1$, there were earlier proofs of both parts of Theorem 8.8. The following theorem describes those results and gives an intermediate case.

Theorem 8.9. *Let $K/\mathbb{Q}$ be a number field with ring of integers R_K, let E/K be an elliptic curve, and let $\hat{h}_E$ be the canonical height on E relative to the divisor* (0)*. Then for every $\epsilon > 0$ there is a constant $C_{24}(E/K,\epsilon) > 0$ such that*

$$P \in E(\bar{K}) \smallsetminus E_{\text{tors}} \quad \Longrightarrow \quad \hat{h}_E(P) \ge \frac{C_{24}(E/K,\epsilon)}{\bigl[K(P):K\bigr]^{\ell+\epsilon}}$$

[23] An abelian variety A of dimension g has *complex multiplication* if $\operatorname{End}(A)_{\mathbb{Q}}$ is a CM field, i.e., an imaginary quadratic extension of a totally field of degree g over $\mathbb{Q}$.

for the values of ℓ given in the following table and under the stated restrictions on E:

ℓ	E/K	Reference
$\ell = 3$	E/K arbitrary	Masser [59]
$\ell = 2$	E/K with $j(E) \notin R_K$	Hindry–Silverman [38]
$\ell = 1$	E/K has complex multiplication	Laurent [51]

Remark 8.10. The CM results in Theorems 8.8 and 8.9 for CM abelian varieties are direct analogues of Dobrowolski's result for $\mathbb{G}_m$. And just as in Dobrowolski's theorem, the $\bigl[K(P):K\bigr]^{-\epsilon}$ may be replaced by a logarithmic factor as described in the footnote to Theorem 8.4.

Proof Sketch of Theorem 8.9*: Two Ideas.* The two methods that have been used to prove estimates such as those described in Theorems 8.8 and 8.9 may be informally described as the "transcendence theory method" and the "Fourier averaging method." The proofs are too intricate for us to give in detail, but we make a few brief remarks about each. In the following, all constants are postive and may depend on A/K and D.

The Transcendence Theory Method. Rather than trying to prove that the height of an individual point cannot be too small, we instead estimate the number of points of small height defined over a finite extension L/K. Let

$$A(L;B) := \bigl\{P \in A(L) : \hat{h}_{A,D}(P) \le B\bigr\}.$$

Then the goal is to prove that there are positive constants C_{25}, C_{26} so that

$$\#A\bigl(L;\, C_{25}/[L:K]\bigr) \le C_{26} \cdot [L:K]^g \cdot \bigl(\log[L:K]\bigr)^g. \tag{8.4}$$

This estimate gives something a bit better than is stated in Theorem 8.8(a), as well as providing a non-trivial estimate for the size of $A(L)_{\text{tors}}$; see Exercise 8.B.

In general, sets of the form $A(L;B)$ have no additive structure, but we can take advantage of the group law to form linear combinations without increasing the height too much. Thus the set

$$A(L;B)^{(g)} := \bigl\{P_1 + \cdots + P_g : P_1, \dots, P_g \in A(L;B)\bigr\}$$

contains $\#A(L;B)^g/g!$ points[24] whose heights are bounded by g^2B. One then constructs a non-zero "small" function on A that vanishes to high order at the points in $A(L;B)^{(g)}$. Next one uses Cauchy's theorem to get an upper bound, and a zero-estimate from transcendence theory

[24] Yes, we're ignoring linear combinations that sum to 0; in practice, there are lots of such matters to be dealt with.

to get a lower bound, for the values of various partial derivatives of the function evaluated at the points in $A(L;B)^{(g)}$. If $A(L;B)$ is sufficiently large, these bounds contradict one another.[25]

The Fourier Averaging Method. The Fourier averaging method has been used obtain Lehmer-type bounds in the case of elliptic curves, although see [53] for some initial work on abelian surfaces. So we take E/K to be an elliptic curve, we let $\hat{h}_E$ be the canonical height relative to the divisor (0), and we consider the decomposition of $\hat{h}_{A,D}$ as a sum of local heights as described in Theorem 7.1(f). The explicit formulas (7.6), (7.7), (7.8) for the local heights give a homomorphism and a "nice" periodic fuction

$$t_w : E(\bar{K}) \longrightarrow \mathbb{R}/\mathbb{Z} \quad \text{and} \quad g_w : \mathbb{R}/\mathbb{Z} \longrightarrow \mathbb{R}$$

so that we can write $\hat{\lambda}_{E,w}$ as (we're cheating a bit here)

$$\hat{\lambda}_{E,w} = g_w \circ t_w + (\text{non-negative function}). \tag{8.5}$$

The idea is to look at at a sum of the heights of some multiples of P, weighted by the Fejér kernel. On the one hand, the fact that $\hat{h}_E$ is a quadratic form yields

$$\begin{aligned} \sum_{m=1}^{M} \underbrace{\left(1 - \frac{m}{M+1}\right)}_{\text{This is the Fejér kernel.}} \cdot \hat{h}_E\big([m]P\big) &= \left(\sum_{m=1}^{M} \left(1 - \frac{m}{M+1}\right) \cdot m^2\right) \hat{h}_E(P) \\ &= \frac{M(M+1)(M+2)}{12} \cdot \hat{h}_E(P). \end{aligned} \tag{8.6}$$

On the other hand, when we expand (8.6) using the local decomposition from Theorem 7.1(f), apply the formula (8.5), and use the fact that t_w is a homomorphism, we end up needing to find a lower bound for sums of the form

$$\sum_{m=1}^{M} \left(1 - \frac{m}{M+1}\right) \cdot g(mt), \quad \begin{array}{l}\text{where } g : \mathbb{R}/\mathbb{Z} \to \mathbb{R} \text{ is even, has } g(0) \text{ as a} \\ \text{maximum, and satisfies } \int_{\mathbb{R}/\mathbb{Z}} g(t)\,dt = 0.\end{array} \tag{8.7}$$

We write the Fourier series of g as[26]

$$g(t) = \sum_{n \neq 0} c_n e^{2\pi i n t} = \sum_{n=1}^{\infty} 2c_n \cos(2\pi n t),$$

---

[25]We emphasize again that we're cheating in many ways in order to give the basic idea of the proof. For example, the "function" that vanishes at the points in $A(L;B)^{(g)}$ is really a theta function, i.e., a section to a line bundle on A.

[26]Note that $c_0 = \int_{\mathbb{R}/\mathbb{Z}} g(t)\,dt = 0$, and $c_{-n} = c_n$ since g is even.

and then (8.7) becomes

$$\sum_{m=1}^{M}\left(1-\frac{m}{M+1}\right)g(mt)$$

$$= \sum_{m=1}^{M}\left(1-\frac{m}{M+1}\right)\cdot\sum_{n=1}^{\infty}2c_n\cos(2\pi nmt)$$

$$= \sum_{n=1}^{\infty}c_n\underbrace{\left(2\sum_{m=1}^{M}\left(1-\frac{m}{M+1}\right)\cos(2\pi mnt)\right)}_{\text{This is is }\ge -1\text{ for all }t\in\mathbb{R}\text{; see Exercise 8.C(a).}}$$

$$\ge -\sum_{n=1}^{\infty}c_n \quad \text{provided that the Fourier coefficients of } g \text{ are non-negative.} \tag{8.8}$$

For example, in the non-archimedean totally multiplicative case, the function $g(t)$ is essentially the periodic 2nd-Bernoulli polynomials

$$\mathbb{B}_2(t) = t^2 - t + \frac{1}{6} = \sum_{n=1}^{\infty}\frac{1}{\pi^2 n^2}\cos(2\pi nt)$$

whose Fourier coefficients are visibly positive.

The magic of (8.8) is that there are M terms in the sum, and in principle every $\cos(2\pi nmt)$ could be negative, but the lower bound in (8.8) says that this cannot happen. The underlying reason is because $\cos(0) = 1$ and the $2\pi nmt \bmod \mathbb{Z}$ values can't all cluster in the region where cos is negative.

Unfortunately, the lower bound in (8.8) is negative; all that the Fourier averaging yields is a lower bound that's less negative than it could be. So we also need to find some positive contributions to the sum. One way to do this is via a naive pigeon-hole argument; cf. [80]. Another method, which is used in [38], is described in Exercise 8.C(b). □

Supplementary Material 8.11 (**A Stronger Lehmer Conjecture**). In the setting of Conjecture 8.7, it has been conjectured [23, Conjecture 1.4] that we can replace K by the field $K(A_{\text{tors}})$ generated by all of the torsion points of A. Thus

$$\hat{h}_{A,D}(P) \ge \frac{C(A/K,D)}{\bigl[K(A_{\text{tors}})(P) : K(A_{\text{tors}})\bigr]^{1/g}} \quad \text{for all non-torsion points } P\in A(\bar{K}).$$

Supplementary Material 8.12 (**A Weaker/Stronger Lehmer-Type Result**). If we restrict the allowable field extensions, then it may be possible to prove a Lehmer-type estimate, or even something stronger. As an example of such a result, if we restrict P to be defined over the maximal abelian extension K^{ab} of K, then it is known [4] that

$$\hat{h}_{A,D}(P) \ge C(A/K,D) > 0 \quad \text{for all non-torsion points } P\in A(K^{\text{ab}}). \tag{8.9}$$

Note that the lower bound does not decrease as the field of definition of P increases. For an elliptic curve E/K, the same is true if we replace K^{ab} with $K(E_{\mathrm{tors}})$; see [34].

8.3. **Height Lower Bounds as the Abelian Variety Varies.** For this section we fix a number field $K/\mathbb{Q}$ and discuss lower bounds for the canonical heights of non-torsion points on abelian varieties defined over K. The general philosophy is that if the abelian variety has high complexity, then its (non-torsion) points should have high complexity.

Measuring the complexity of an abelian variety using the height $h(A/K)$ given in Definition 4.17, our general philosophy translates into the following conjecture.

Conjecture 8.13 (Dem'janenko–Lang Height Conjecture)**.** [25], [45, page 92], [82] *Let $K/\mathbb{Q}$ be a number field, and let $g \ge 1$. As described in Section 4.2, we fix a height function h on the set of abelian varieties of dimension g defined over K. Then there are constants*

$$C_{27}(K,g) > 0 \quad \textit{and} \quad C_{28}(K,g) \ge 0$$

so that for all abelian varieties A/K with $\dim(A) = g$ and all ample symmetric divisors $D \in \operatorname{Div}(A)$, we have

$$\hat{h}_{A,D}(P) \ge C_{27}(K,g) \cdot h(A/K) - C_{28}(K,g) \qquad \textit{for all } P \in A(K) \textit{ such that } \mathbb{Z}\cdot P \textit{ is Zariski dense in } A.$$

Remark 8.14. One might even hope that Conjecture 8.13 is true with constants C_{27} and C_{28} that depend only on $[K:\mathbb{Q}]$ and g.

Versions of Conjecture 8.13 are known if one restricts to a subset of the set of abelian varieties. In particular, the conjecture is true for abelian varieties A such that the distance of the moduli point of A to the boundary of moduli space satisfies some condition. We give two examples of such results, with the caveat that we have not stated the strongest versions from the cited papers.

Theorem 8.15. *Let $\mathcal{A}_g$ denote the moduli space of principally polarized abelian varieties of dimension g, let $\overline{\mathcal{A}}_g$ be the closure of $\mathcal{A}_g$ for some projective embedding, and let $\mathcal{A}_g^{\mathrm{simp}} \subset \mathcal{A}_g$ be the locus of geometrically simple abelian varieties. Then Conjecture 8.13*

$$\hat{h}_{A,D}(P) \ge C_{27}(K,g) \cdot h(A/K) - C_{28}(K,g)$$

is true for all $A \in \mathcal{A}_g(K)$ satisfying one of the following conditions:

(a) (David [21]) *There is an archimedean absolute value $v \in M_K^\infty$ such that*

$$h(A/K) \le C_{29}(K,g) \cdot \operatorname{dist}_v\bigl(A, (\overline{\mathcal{A}}_g \smallsetminus \mathcal{A}_g)(K_v)\bigr).$$

(b) (Pazuki [72]) *Assume that* $\dim(A) = 2$ *and that for all archimedean absolute values* $v \in M_K^\infty$*, we have*

$$\operatorname{dist}_v\bigl(A, (\overline{\mathcal{A}}_g \smallsetminus \mathcal{A}_g^{\mathrm{simp}})(K_v)\bigr) \ge C_{30}(K,g) > 0.$$

For elliptic curves, more is known. We state two results, restricting attention to $\mathbb{Q}$ for ease of exposition. In order to state these results, we set the following notation.

Definition 8.16. Let $E/\mathbb{Q}$ be an elliptic curve. A *quasi-minimal Weierstrass equation for* E is an equation of the form

$$E : y^2 = x^3 + Ax + B \quad \text{with } A, B \in \mathbb{Z}, \text{ and with } \gcd(A^3, B^2) \text{ a 12th-power-free integer.}$$

The *minimal discriminant, conductor, and Szpiro-ratio of* E are, respectively, the quantities[27]

$$\mathsf{D}_E^{\mathrm{qm}} = 4A^3 + 27B^2, \quad \mathsf{N}_E^{\mathrm{qm}} = \prod_{p \mid \mathsf{D}_E^{\mathrm{qm}}} p \cdot \prod_{p \mid \gcd(A,B)} p, \quad \mathsf{S}_E^{\mathrm{qm}} = \frac{\log |\mathsf{D}_E^{\mathrm{qm}}|}{\log \mathsf{N}_E^{\mathrm{qm}}},$$

where A and B are the coefficients of a quasi-minimal Weierstrass equation.

Theorem 8.17. (Hindry–Silverman [37, 80]) *Let* $E/\mathbb{Q}$ *be an elliptic curve. Then for all non-torsion points* $P \in E(\mathbb{Q})$*, we have*

$$\hat{h}_E(P) \ge C_{27} \cdot \max\Bigl\{\log|\Delta_E|,\, h\bigl(j(E)\bigr)\Bigr\} - C_{28}$$

where either:

(a) *the constants* C_{27} *and* C_{28} *depend only on the number of distinct primes dividing the denominator of* $j(E)$*, or*
(b) *the constants* C_{27} *and* C_{28} *depend only on the Szpiro-ratio* $\mathsf{S}_E^{\mathrm{qm}}$ *of* E.

Remark 8.18. A conjecture of Szpiro asserts that for any $\epsilon > 0$, there are only finitely many elliptic curves $E/\mathbb{Q}$ whose Szpiro ratio satisfies

$$\mathsf{S}_E^{\mathrm{qm}} > 6 + \epsilon.$$

The *ABC*-conjecture of Masser and Osterlé is more-or-less equivalent to Szpiro's conjecture; see for example [88, VIII §11]. Thus a combination of Theorem 8.17 and the *ABC*/Szpiro conjecture imply the validity of the height conjecture (Conjecture 8.13) for $\dim(A) = 1$.

[27] We are cheating here as to the powers of 2 and 3 that appear in the minimal discriminant and the conductor, hence our use of the "qm" superscript to indicate that our discriminant and conductor are relative to a quasi-minimal Weierstrass equation. However, since our cheating involves at most a small bounded power of 2 and 3, we can absorb our transgression into the constants!

We state and sketch the proof of a weak (but still interesting) version of the Dem'janenko–Lang conjecture (Conjecture 8.13) in which we vary over the cyclic twists of a fixed abelian variety A.

Theorem 8.19. ([82]) *Fix the following quanitites:*

A/K	*an abelian variety defined over a number field.*
D	*an ample symmetric divisor* $D \in \operatorname{Div}_K(A)$.
$m \geq 2$	*an integer such that* $\boldsymbol{\mu}_m \subseteq \operatorname{Aut}(A/K)$ *as* $\operatorname{Gal}(\bar{K}/K)$-*modules.*
A_Δ/K	*the cyclic twist of* A/K *by* $\Delta \in K^*$ *as described in Example* 3.29, *where we recall that* A_Δ *and* $A_{\Delta'}$ *are* K-*isomorphic if and only if* $\Delta/\Delta' \in K^{*m}$.
ξ_Δ	*a* $\bar{K}$-*isomorphism* $\xi_\Delta : A_\Delta \to A$.
D_Δ	*the divisor* $D_\Delta = \xi_\Delta^* D$.

Then there are constants

$$C_{31} = C_{31}(K, A, D, m) > 0 \quad \textit{and} \quad C_{32} = C_{32}(K, A, D, m) \geq 0$$

so that for all $\Delta \in K^*$ *and all* $P \in A_\Delta(K)$, *either*

$$[\zeta]P = P \textit{ for some } 1 \neq \zeta \in \boldsymbol{\mu}_m, \tag{8.10}$$

or else

$$\hat{h}_{A_\Delta, D_\Delta}(P) \geq C_{31} \sum_{\substack{0 \neq \mathfrak{p} \in \operatorname{Spec}(R_K) \\ m \nmid \operatorname{ord}_\mathfrak{p}(\Delta)}} \log \mathsf{N}_{K/\mathbb{Q}}(\mathfrak{p}) - C_{32}. \tag{8.11}$$

Remark 8.20. With notation as in Theorem 8.19, one can show that there is a finite set of prime $S(A/K, m)$ such that if $\mathfrak{p} \notin S(A/K, m)$, then for all $\Delta \in K^*$ we have

$$A_\Delta \text{ has bad reduction at } \mathfrak{p} \quad \Longleftrightarrow \quad m \nmid \operatorname{ord}_\mathfrak{p}(\Delta).$$

See Exercise 8.H. Hence the lower bound (8.11) for $\hat{h}_{A_\Delta, D_\Delta}(P)$ more-or-less measures the bad reduction of A_Δ, which in turn is more-or-less the height of A_Δ. Thus Theorem 8.19 is a version of Conjecture 8.13 for the cyclic twists of a fixed abelian variety.

The proof of Theorem 8.19 relies on the useful observation that the Weil height of a point is bounded below by a function that grows log-linearly as a function of the discriminant of the field generated by the point's coordinates.

Proposition 8.21. *Let $P \in \mathbb{P}^N(\bar{\mathbb{Q}})$, let K_P be the field of definition*[28] *of P, and let $d_P = [K : \mathbb{Q}]$. Then*

$$h(P) \ge \frac{1}{2d_P - 2}\left(\frac{1}{d_P} \cdot \log\bigl|\mathrm{Disc}(K_P/\mathbb{Q})\bigr| - d_P \log(d_P)\right).$$

Proof. Earlier we posed as an exercise the special case of Proposition 8.21 that will be used in the proof of Theorem 8.19; see Exercise 4.F. Proposition 8.21 for $\mathbb{P}^1$ was proven by Mahler [54] (using somewhat different terminology), and it was proven for $\mathbb{P}^N$ using height terminology in [82, Theorem 2]. □

Proof sketch of Theorem 8.19. We choose $\delta \in \bar{K}^*$ satisfying $\delta^m = \Delta$ so that for all $\sigma \in \mathrm{Gal}(\bar{K}/K)$, the isomorphism ξ_Δ satisfies

$$\xi_\Delta^{-1} \circ \sigma(\xi_\Delta) = [\delta^\sigma/\delta] \in \boldsymbol{\mu}_m \subseteq \mathrm{Aut}(A/K). \tag{8.12}$$

Then for any $P \in A_\Delta(K)$ and $\sigma \in \mathrm{Gal}(\bar{K}/K)$, we have

$$\begin{aligned} \sigma\bigl(\xi_\Delta(P)\bigr) &= \sigma(\xi_\Delta)\bigl(\sigma(P)\bigr) \\ &= \sigma(\xi_\Delta)(P) \quad \text{since } P \in A_\Delta(K), \\ &= \xi_\Delta\bigl([\delta^\sigma/\delta](P)\bigr) \quad \text{from (8.12).} \end{aligned} \tag{8.13}$$

In particular, if we further assume that the point $P \in A_\Delta(K)$ satisfies

$$P \ne [\zeta]P \quad \text{for all } 1 \ne \zeta \in \boldsymbol{\mu}_m, \tag{8.14}$$

then we can compute

$$\begin{aligned} \sigma\bigl(\xi_\Delta(P)\bigr) = \xi_\Delta(P) & \\ \iff& \quad \xi_\Delta(P) = \xi_\Delta\bigl([\delta^\sigma/\delta](P)\bigr) \quad \text{from (8.13).} \\ \iff& \quad P = [\delta^\sigma/\delta](P) \quad \text{since } \xi_\Delta \text{ is an isomorphism,} \\ \iff& \quad \delta^\sigma/\delta = 1 \quad \text{from (8.14).} \end{aligned}$$

Hence the field of definition of $\xi_\Delta(P)$ is

$$\begin{aligned} K_{\xi_\Delta(P)} &= \text{Fixed field of } \bigl\{\sigma \in \mathrm{Gal}(\bar{K}/K) : \sigma\bigl(\xi_\Delta(P)\bigr) = \xi_\Delta(P)\bigr\} \\ &= \text{Fixed field of } \bigl\{\sigma \in \mathrm{Gal}(\bar{K}/K) : \delta^\sigma = \delta\bigr\} \\ &= K(\delta). \end{aligned} \tag{8.15}$$

In particular, we note that

$$[K_{\xi_\Delta(P)} : \mathbb{Q}] = \bigl[K(\delta) : \mathbb{Q}\bigr] \le m \cdot [K : \mathbb{Q}] \tag{8.16}$$

is bounded by a quantity that does not depend on P.

[28] The *field of definition* of a point $P = [\alpha_0, \ldots, \alpha_N] \in \mathbb{P}^N(\bar{\mathbb{Q}})$ is obtained by choosing an index i with $\alpha_i \ne 0$, setting $\beta_j = \alpha_j/\alpha_i$, and taking $K_P = \mathbb{Q}(\beta_0, \ldots, \beta_N)$. Alternatively, K_P is the fixed field of $\bigl\{\sigma \in \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) : \sigma(P) = P\bigr\}$.

We compute

$$\begin{aligned}
&\hat{h}_{A_\Delta,D_\Delta}(P)\\
&\quad = \hat{h}_{A,D}\bigl(\xi_\Delta(P)\bigr) \quad \text{from Exercise 5.E,}\\
&\quad = h_{A,D}\bigl(\xi_\Delta(P)\bigr) + O(1) \quad \text{from the fact (Theorem 5.2(b)) that } \hat{h}_{A,D} - h_{A,D} \text{ is bounded,}\\
&\quad = \frac{1}{\nu} h_{\mathbb{P}^N}\bigl(\psi_{\nu D}\circ\xi_\Delta(P)\bigr) + O(1) \quad \text{where we select an integer } \nu \ge 1 \text{ so that } \nu D \text{ is very ample, and then fix an associated embedding } \psi_{\nu D} : A \hookrightarrow \mathbb{P}^N,\\
&\quad \ge C_{33}\log\Bigl|\operatorname{Disc}\bigl(K_{\xi_\Delta(P)}/\mathbb{Q}\bigr)\Bigr| - C_{34} \quad \text{from Proposition 8.21 and the fact (8.16) that the degree of the extension } K_{\xi_\Delta(P)}/\mathbb{Q} \text{ is bounded independently of } P,\\
&\quad = C_{33}\log\Bigl|\operatorname{Disc}\bigl(K(\delta)/\mathbb{Q}\bigr)\Bigr| - C_{34} \quad \text{from (8.15).}
\end{aligned}$$

It remains to observe that there is a finite set of prime $S(K,m)$ in $\operatorname{Spec}(R_K)$ such that for primes $\mathfrak{p} \notin S(K,m)$, we have

$$\begin{aligned}
m \nmid \operatorname{ord}_{\mathfrak{p}}(\Delta) \quad &\Longrightarrow \quad \mathfrak{p} \text{ is ramified in } K(\delta),\\
&\Longrightarrow \quad \mathfrak{p} \mid \operatorname{Disc}\bigl(K(\delta)/K\bigr).
\end{aligned}$$

It follows that

$$\log\Bigl|\operatorname{Disc}\bigl(K(\delta)/\mathbb{Q}\bigr)\Bigr| \ge C_{35}(K,m) \sum_{\substack{0\ne\mathfrak{p}\in\operatorname{Spec}(R_K)\\ m\nmid\operatorname{ord}_{\mathfrak{p}}(\Delta)}} \log \mathsf{N}_{K/\mathbb{Q}}(\mathfrak{p}) - C_{36}(K,m),$$

which completes the proof of Theorem 8.19. □

Supplementary Material 8.22 (**A Soupçon of History**). Dem'janenko stated and proved Conjecture 8.13 for a certain collection of elliptic curves. Lang expressed some doubt about Dem'janenko's proof, while also extending the conjecture to all elliptic curves. Silverman subsequently further extended the conjecture to abelian varieties, with the Zariski density condition being an observation of Bertrand.

Supplementary Material 8.23 (**Lower Bounds for $E/\mathbb{Q}$**). Lang's original formulation of Conjecture 8.13 for an elliptic curve $E/\mathbb{Q}$ with minimal discriminant $\Delta_E \in \mathbb{Z}$ had the form[29]

$$\hat{h}_E(P) \ge C_{27}\cdot\log|\Delta_E| - C_{28}. \tag{8.17}$$

We note that (8.17) implies that

$$\hat{h}_{\mathbb{Q}}^{\min} := \inf_{\substack{E/\mathbb{Q}\\ P\in E(\mathbb{Q})\smallsetminus E_{\text{tors}}}} \hat{h}_E(P) > 0 \quad\text{and}\quad \hat{\ell}_{\mathbb{Q}}^{\min} := \inf_{\substack{E/\mathbb{Q}\\ P\in E(\mathbb{Q})\smallsetminus E_{\text{tors}}}} \frac{\hat{h}_E(P)}{\log|\Delta_E|} > 0.$$

[29]Assuming a deep conjecture of M. Hall, we do not get a stronger statement if we include $h\bigl(j(E)\bigr)$ in lower bound in (8.17).

An early table giving upper bounds for $\hat{h}_{\mathbb{Q}}^{\min}$ and $\hat{\ell}_{\mathbb{Q}}^{\min}$ appears in [81, Table 2], and Elkies conducted a more extensive search giving a much better upper bound for $\hat{h}_{\mathbb{Q}}^{\min}$; see `people.math.harvard.edu/~elkies/low_height.html`. A search of the elliptic curves $E/\mathbb{Q}$ in the LMFDB [52] having conductor at most 10^6 found that the following curve in this set gives the smallest values for both $\hat{h}_{\mathbb{Q}}^{\min}$ and $\hat{\ell}_{\mathbb{Q}}^{\min}$:

$$y^2+xy+y=x^3+x^2-125615x+61201397 \qquad \text{[LMFDB Label = 3990.v1]}$$
$$\operatorname{Disc}_{E/\mathbb{Q}}=-2^{11}3^45^77^519^3, \quad P=(7107,\,594946),$$
$$\frac{\log(N_E)}{\log|\operatorname{Disc}_{E/\mathbb{Q}}|}=5.047, \quad \hat{h}_E(P)=0.00891, \quad \frac{\hat{h}_E(P)}{\log|\operatorname{Disc}_{E/\mathbb{Q}}|}=0.00021$$

Exercises for Section 8.

Exercise 8.A. With notation as in the statement of Lehmer's conjecture for abelian varieties (Conjecture 8.7), prove that for every $\epsilon>0$, there are infinitely many non-torsion points $P\in A(\bar{K})$ satisfying

$$\hat{h}_{A,D}(P)\le\frac{1}{\bigl[K(P):K\bigr]^{1/g-\epsilon}}.$$

Exercise 8.B. With notation as in the proof sketch of Theorem 8.9, use (8.4) to prove the following.

(a) Prove that

$$\#A(L)_{\text{tors}}\le C_{26}\cdot[L:K]^g\cdot\bigl(\log[L:K]\bigr)^g.$$

(b) Prove that

$$\hat{h}_{A,D}(P)\ge\frac{C_{37}(A/K,D)}{[L:K]^{2g+1}\cdot\bigl(\log[L:K]\bigr)^{2g}}\quad\text{for all } P\in A(L)\smallsetminus A_{\text{tors}}.$$

Note that this is a bit stronger than the estimate in Theorem 8.9. *Hint*: Apply (8.4) to the set of multiples of P.

Exercise 8.C. Let $t\in\mathbb{R}$.

(a) Prove that

$$\sum_{m=1}^{M}\left(1-\frac{m}{M+1}\right)\cos(mt)=-\frac12+\frac{1}{2(M+1)}\left|\sum_{k=0}^{M}e^{ikt}\right|^2.\ge-\frac12.$$

(b) Suppose that there is an integer $N\ge1$ such that $Nt\in\mathbb{Q}$. Prove that

$$\sum_{n=1}^{\infty}\frac{1}{n^2}\sum_{m=1}^{M}\left(1-\frac{m}{M+1}\right)\cos(2\pi nmt)\ge\frac{\pi^2}{6}(M+1)\left(\frac{M+1}{N}-1\right).$$

Exercise 8.D. Let K be a number field or a function field, let $\varphi:\mathbb{P}^N\to\mathbb{P}^N$ be a morphism of degree $d\ge2$ defined over K, and let $\varphi^{\circ n}=\varphi\circ\varphi\cdots\circ\varphi$ denote the n-fold iterate of φ. Let $h:\mathbb{P}^N(\bar{K})\to\mathbb{R}$ be the Weil height on $\mathbb{P}^N$ as described in Definition 4.5.

(a) Let $P\in\mathbb{P}^N(\bar{K})$. Prove that the following limit exists:

$$\hat{h}_\varphi(P):=\lim_{n\to\infty}d^{-n}\cdot h\bigl(\varphi^{\circ n}(P)\bigr).$$

(b) The φ-*orbit* of a point $P \in \mathbb{P}^N(\bar{K})$ is the image under the set of forward iterates of φ,

$$\mathrm{Orbit}_\varphi(P) := \bigl\{\varphi^{\circ n}(P) : n \ge 0\bigr\}.$$

A point P is said to be *preperiodic for f* if its orbit $\mathrm{Orbit}_\varphi(P)$ is finite. Suppose that K is a number field. Prove that

$$P \text{ is preperiodic} \quad\Longleftrightarrow\quad \hat{h}_\varphi(P) = 0.$$

This is a dynamical analogue of Theorems 8.1 and 8.5.

Exercise 8.E. Let G be a group, and for each $m \ge 1$, let

$$\varphi_m : G \longrightarrow G, \quad \varphi_m(g) = g^m$$

be the mth power map. Let $g \in G$. Prove that the following are equivalent.

(a) g is an element of finite order.
(b) There exists an $m \ge 2$ such that the set $\bigl\{\varphi_m^n(g) : n \ge 1\bigr\}$ is finite. (Here φ_m^n denotes the nth iterate of the map φ_m.)
(c) For all $m \ge 2$, the set $\bigl\{\varphi_m^n(g) : n \ge 1\bigr\}$ is finite.

Exercise 8.F. We consider the Dem'janenko–Lang conjecture (Conjecture 8.13).

(a) Prove that Conjecture 8.13 is false for $g \ge 2$ if we require only that $P \in A(K)$ is a non-torsion point.
(b) Prove that if we make the additional assumption that A is a simple abelian variety, then Conjecture 8.13 is equivalent the same statement with the requirement on $P \in A(K)$ being relaxed to the assumption that P is not a torsion point.

Exercise 8.G. With notation as in Theorem 8.19, prove that

$$\bigl\{\Delta \in K^*/K^{*m} : A_\Delta(K)_{\mathrm{tors}} \not\subseteq A[m]\bigr\}$$

is a finite set, i.e., prove that there are only finitely many distinct twists of A having a K-rational torsion point whose order deos not divide m. *Hint*: A starting point is Exercise 3.F, which says that if $[\zeta]P = P$ for some $1 \ne \zeta \in \boldsymbol{\mu}_m$, then $[m]P = 0$. Then use Theorem 8.19.

Exercise 8.H. With notation as in Theorem 8.19, let $S(A/K, m)$ be the set of primes $0 \ne \mathfrak{p} \in \mathrm{Spec}(R_K)$ satisfying any one of the following conditions:

(a) The characteristic p of $\mathfrak{p}$ satisfies $p \le [K : \mathbb{Q}]$.
(b) $\mathfrak{p} \mid m$.
(c) A/K has bad reduction at $\mathfrak{p}$.

Prove that for all $\Delta \in K^*$ and all primes $\mathfrak{p} \notin S(A/K, m)$, we have

$$A_\Delta \text{ has bad reduction} \quad\Longrightarrow\quad m \nmid \mathrm{ord}_\mathfrak{p}(\Delta).$$

Hint: The criterion of Néron–Ogg–Shafarevich [79] can be used to determine whether A_Δ has bad reduction at $\mathfrak{p}$.

Exercise 8.I. The list of elliptic curves $E/\mathbb{Q}$ in the LMFDB includes curves of conductor far larger than 10^6. Extend the search described in Supplementary Material 8.23 and see if you can find smaller (non-zero) values of $\hat{h}_E(P)$ and $\hat{h}_E(P)/\log|\operatorname{Disc}_{E/\mathbb{Q}}|$.

9. Canonical Heights in Families and Specialization Theorems

Suppose that we have a family of abelian varieties and a family of points. We might ask how the canonical heights of those points vary. In this section we give an answer to that question, discuss various ways in which that answer might be strengthened, and give a specialization application. For the remainder of Section 9, we fix the notation described in Table 1.

K	a number field.
C/K	a smooth projective curve C/K.
$A/K(C)$	an abelian variety defined over $K(C)$.
$(\mathcal{A},\pi)$	a smooth projective family of varieties $\pi:\mathcal{A}\to C$ defined over K whose generic fiber is A. This implies that there a non-empty Zariski open set $C^\circ\subset C$ so that $\mathcal{A}^\circ=\pi^{-1}(C^\circ)$ is an abelian group scheme† over C°.
P	a point $P\in A\bigl(K(C)\bigr)$.
$\mathcal{P}$	the section $\mathcal{P}:C\to\mathcal{A}$ associated to P. For notational convenience, we write $\mathcal{P}_t$ instead of $\mathcal{P}(t)$ for the image of a point $t\in C(\bar{K})$.
D	a divisor $D\in\operatorname{Div}(A/K)$.
$\mathcal{D}$	the divisor $\mathcal{D}\in\operatorname{Div}(\mathcal{A}/K)$ obtained by "thickening" D.‡
h_C	a Weil height function on $C(\bar{K})$ associated to a divisor of degree 1.

† In other words, the addition and inversion maps on A extend to morphisms $\mathcal{A}^\circ\times_{C^\circ}\mathcal{A}^\circ\to\mathcal{A}^\circ$ and $\mathcal{A}^\circ\to\mathcal{A}^\circ$ that commute with π and that make $\mathcal{A}^\circ$ into an abelian group scheme over $\mathcal{C}^\circ$.

‡ In other words, D is a divisor on A/K, where $A/K\cong\mathcal{A}\times\operatorname{Spec}K(C)$ is the generic fiber of $\mathcal{A}$, and such that $\mathcal{D}$ is the Zariski closure of D in $\mathcal{A}$.

Table 1. Notation for Section 9.

Theorem 5.2(b) says that the canonical height $\hat{h}_{A,D}$ on an abelian variety differs from the Weil height $h_{A,D}$ by a bounded amount, but in general the bound will depend on A and D. The following lemma quantifies this result by saying, roughly, that the the difference is bounded in terms of the equations that define A and D.

Lemma 9.1. *Fix a Weil height function $h_{\mathcal{A},\mathcal{D}}$ on $\mathcal{A}$. Then there are constants $C_{38} > 0$ and $C_{39} \ge 0$ such that for all $x \in \mathcal{A}^\circ(\bar{K})$,*

$$\left|\hat{h}_{\mathcal{A}_{\pi(x)},\mathcal{D}_{\pi(x)}}(x) - h_{\mathcal{A},\mathcal{D}}(x)\right| \le C_{38} h_C\bigl(\pi(x)\bigr) + C_{39}.$$

Proof. For ease of exposition, and because it is the case with the most applications, we will give the proof in the case that D is symmetric. We leave the anti-symmetric case to the reader, and the general case then follows by linearity.

We let

$$h_{\mathcal{A},\mathcal{D}} : \mathcal{A}(\bar{K}) \longrightarrow \mathbb{R}$$

be a Weil height function on $\mathcal{A}(\bar{K})$ associated to the divisor $\mathcal{D}$. We consider the duplication morphism $[2]_A : A \to A$, and we extend it to a dominant rational map[30]

$$[2]_{\mathcal{A}} : \mathcal{A} \dashrightarrow \mathcal{A}$$

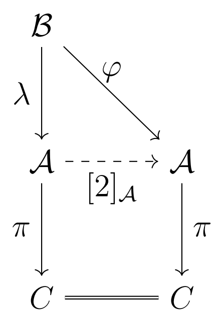
We further note that $[2]_{\mathcal{A}}$ is a morphism when when restricted to $\mathcal{A}^\circ$.

Height functions do not behave nicely for rational maps, but we know that it is always possible to turn a rational map into a morphism by blowing up the indeterminacy locus. Thus we can find a projective variety $\mathcal{B}/K$, a birational morphism $\lambda : \mathcal{B} \to \mathcal{A}$ that is an isomorphism over $\mathcal{A}^\circ$, and a morphism $\varphi : \mathcal{B} \to \mathcal{A}$, so that the following diagram commutes:

[30]We know that $\mathcal{A}^\circ \to C^\circ$ is a group scheme, so its multiplication map

$$\mu_{\mathcal{A}^\circ} : \mathcal{A}^\circ \times_{C^\circ} \mathcal{A}^\circ \longrightarrow \mathcal{A}^\circ$$

is a morphism. It follows that the duplication may $[2]_{\mathcal{A}^\circ}$ is a morphism on $\mathcal{A}^\circ$, since $[2]_{\mathcal{A}^\circ}$ is the composition of the diagonal embedding followed by $\mu_{\mathcal{A}^\circ}$. But when we try to extend $\mu_{\mathcal{A}^\circ}$ to all of $\mathcal{A}$, there may be a non-trivial indeterminacy locus, so in particular, the duplication map $[2]_{\mathcal{A}}$ may only be a rational map.



The fact that φ is a morphism allows us to apply the height machine to conclude that

$$h_{\mathcal{A},\mathcal{D}} \circ \varphi = h_{\mathcal{B},\varphi^*\mathcal{D}} + O(1). \tag{9.1}$$

We know from Corollary 3.8 that $[2]_A^* D \sim 4D$ on the generic fiber of $\mathcal{A}$.[31] Further,the duplication map is a morphism on $\mathcal{A}^\circ$, so we (more-or-less) have[32]

$$[2]_{\mathcal{A}^\circ}^* \mathcal{D}^\circ \sim 4\mathcal{D}^\circ \quad \text{in } \operatorname{Div}(\mathcal{A}^\circ).$$

Extending to all of $\mathcal{A}$, we find that

$$\varphi^*\mathcal{D} \sim 4\lambda^*\mathcal{D} + \mathcal{F} \tag{9.2}$$

for some divisor $\mathcal{F} \in \operatorname{Div}(\mathcal{B})$ that is supported above the indeterminacy locus of $[2]_\mathcal{A}$, and thus that satisfies

$$\operatorname{Support}(\mathcal{F}) \subseteq \lambda^{-1}(\mathcal{A} \smallsetminus \mathcal{A}^\circ) = (\pi \circ \lambda)^{-1}(C \smallsetminus C^\circ). \tag{9.3}$$

We compute

$$\begin{aligned}
h_{\mathcal{A},\mathcal{D}} \circ \varphi &= h_{\mathcal{B},\varphi^*\mathcal{D}} + O(1) && \text{from (9.1), since } \varphi \text{ is a morphism,} \\
&= h_{\mathcal{B},4\lambda^*\mathcal{D}+\mathcal{F}} + O(1) && \text{from (9.2),} \\
&= 4h_{\mathcal{B},\lambda^*\mathcal{D}} + h_{\mathcal{B},\mathcal{F}} + O(1) \\
&= 4h_{\mathcal{A},\mathcal{D}} \circ \lambda + h_{\mathcal{B},\mathcal{F}} + O(1) && \text{since } \lambda \text{ is a morphism.}
\end{aligned} \tag{9.4}$$

We use the fact (9.3) that the support of $\mathcal{F}$ is contained in a finite number of fibers of $\pi \circ \lambda$ to choose an effective divisor $E \in \operatorname{Div}(C)$ satisfying

$$\operatorname{Support}(E) \subset C \smallsetminus C^\circ \quad \text{and} \quad -(\pi \circ \lambda)^*(E) \le \mathcal{F} \le (\pi \circ \lambda)^*(E). \tag{9.5}$$

[31]This is where we use the assumption that D is symmetric. If it were anti-symmetric, we would instead have $[2]_A^* D \sim 2D$.

[32]More precisely, we can cover C° by finitely many Zariski open subsets so that we get this relation above each subset. But for ease of exposition, we will ignore such issues and leave it to the reader to make the necessary adjustments in the proof.

We use this to estimate

$$
\begin{aligned}
|h_{\mathcal{B},\mathcal{F}}| &\le h_{\mathcal{B},(\pi\circ\lambda)^*(E)} + O(1) && \text{away from the support of } (\pi\circ\lambda)^*(E), \\
&&& \text{using (9.5) and the fact that the} \\
&&& \text{height associated to effective divisor is} \\
&&& \text{bounded below off of its base locus,} \\
&= h_{C,E}\circ\pi\circ\lambda + O(1) && \text{on } C^\circ \text{, i.e., away from the support of } E, \\
&&& \text{since } \lambda \text{ and } \pi \text{ are morphisms,} \\
&= O(h_C\circ\pi\circ\lambda) && \text{since } h_C \text{ is a height attaced to a degree 1,} \\
&&& \text{and thus ample, divisor; see Exercise 4.C.}
\end{aligned}
\tag{9.6}
$$

Using (9.6) in (9.4), we find that

$$
h_{\mathcal{A},\mathcal{D}}\circ\varphi = 4h_{\mathcal{A},\mathcal{D}}\circ\lambda + O(h_C\circ\pi\circ\lambda) \quad \text{on } \mathcal{B}^\circ. \tag{9.7}
$$

Now let $x \in \mathcal{A}^\circ(\bar{K})$. The restriction of the map λ to $\mathcal{B}^\circ$ is an isomorphism from $\mathcal{B}^\circ$ to $\mathcal{A}^\circ$, so there is a unique point $y = \lambda^{-1}(x) \in \mathcal{B}(\bar{K})$. Evaluating (9.7) at y gives

$$
\begin{aligned}
h_{\mathcal{A},\mathcal{D}}\circ\varphi(y) &= 4h_{\mathcal{A},\mathcal{D}}\circ\lambda(y) + O\bigl(h_C\circ\pi\circ\lambda(y)\bigr) \\
h_{\mathcal{A},\mathcal{D}}\bigl(\varphi\circ\lambda^{-1}(x)\bigr) &= 4h_{\mathcal{A},\mathcal{D}}\bigl(\lambda\circ\lambda^{-1}(x)\bigr) + O\Bigl(h_C\circ\pi\circ\lambda\circ\lambda^{-1}(x)\Bigr) \\
h_{\mathcal{A},\mathcal{D}}\bigl([2]_{\mathcal{A}}(x)\bigr) &= 4h_{\mathcal{A},\mathcal{D}}(x) + O\Bigl(h_C\bigl(\pi(x)\bigr)\Bigr).
\end{aligned}
\tag{9.8}
$$

For each $n \ge 0$, we replace x in (9.8) with $[2^n]_{\mathcal{A}}(x)$, and we note that

$$
\pi\bigl([2^n]_{\mathcal{A}}(x)\bigr) = \pi(x),
$$

since $[2]_{\mathcal{A}}$ respects the fibration by π. This gives

$$
h_{\mathcal{A},\mathcal{D}}\bigl([2^{n+1}]_{\mathcal{A}}(x)\bigr) - 4h_{\mathcal{A},\mathcal{D}}\bigl([2^n]_{\mathcal{A}}(x)\bigr) = O\Bigl(h_C\bigl(\pi(x)\bigr)\Bigr). \tag{9.9}
$$

The key observation is that the big-O constant in (9.9) depends on neither x nor n. So we can use the same argument that we used in the existence proof of the canonical height to deduce

$$
\left|\frac{1}{4^n}h_{\mathcal{A},\mathcal{D}}\bigl([2^n]_{\mathcal{A}}(x)\bigr) - h_{\mathcal{A},\mathcal{D}}(x)\right| = O\Bigl(h_C\bigl(\pi(x)\bigr)\Bigr). \tag{9.10}
$$

See the proof of Theorem 5.2(a), especially the telescoping sum calculation in (5.6) and the special case with $m = 0$ in (5.7). Letting $n \to \infty$ in (9.10) yields

$$
\bigl|\hat{h}_{\mathcal{A},\mathcal{D}}(x) - h_{\mathcal{A},\mathcal{D}}(x)\bigr| = O\Bigl(h_C\bigl(\pi(x)\bigr)\Bigr).
$$

which completes the proof of Lemma 9.1. □

Remark 9.2. Lemma 9.1 holds more generally for families of abelian varieties $\mathcal{A} \to T$ where the base variety has higher dimension. See Exercise 9.A.

The next result improves on the estimate in Lemma 9.1 in the case that the points on $\mathcal{A}$ vary in an algebraic family.

Theorem 9.3. *Continuing with the notation described in Table* 1*, we have*

$$\lim_{\substack{t\in C(\bar{K})\\ h_C(t)\to\infty}} \frac{\hat{h}_{\mathcal{A}_t,\mathcal{D}_t}(\mathcal{P}_t)}{h_C(t)} = \hat{h}_{A,D}(P). \tag{9.11}$$

Proof. We use the triangle inequality to break the difference

$$\left|\hat{h}_{\mathcal{A}_t,\mathcal{D}_t}(\mathcal{P}_t) - \hat{h}_{A,D}(P)\cdot h_C(t)\right|$$

into three pieces. Thus

$$\begin{aligned}
&\left|\hat{h}_{\mathcal{A}_t,\mathcal{D}_t}(\mathcal{P}_t) - \hat{h}_{A,D}(P)\cdot h_C(t)\right| && (9.12)\\
&\qquad\le \left|\hat{h}_{\mathcal{A}_t,\mathcal{D}_t}(\mathcal{P}_t) - h_{\mathcal{A},\mathcal{D}}(\mathcal{P}_t)\right| && (9.13)\\
&\qquad\quad + \left|h_{\mathcal{A},\mathcal{D}}(\mathcal{P}_t) - h_{A,D}(P)\cdot h_C(t)\right| && (9.14)\\
&\qquad\quad + \left|h_{A,D}(P) - \hat{h}_{A,D}(P)\right|\cdot h_C(t). && (9.15)
\end{aligned}$$

We can use Lemma 9.1 with $x = \mathcal{P}_t$ and $\pi(x) = \pi(\mathcal{P}_t) = t$ to bound (9.13) by

$$\left|\hat{h}_{\mathcal{A}_t,\mathcal{D}_t}(\mathcal{P}_t) - h_{\mathcal{A},\mathcal{D}}(\mathcal{P}_t)\right| \le C_{38}h_C(t) + C_{39}. \tag{9.16}$$

For (9.14) we use the fact that the canonical height and the Weil height differ by a bounded amount (Theorem 5.2(b)), where we're working in the function field setting with $P \in A\bigl(K(C)\bigr)$. This yields

$$\left|h_{A,D}(P) - \hat{h}_{A,D}(P)\right|\cdot h_C(t) \le C_{40}h_C(t). \tag{9.17}$$

We note that the constants in (9.16) and (9.17) are independent of both $P \in A\bigl(K(C)\bigr)$ and $t \in C(\bar{K})$. We will not be quite that lucky when we bound (9.14).

The key to bounding (9.14) is to note that the map $\mathcal{P} : T \to \mathcal{A}$ is a morphism.[33] Hence

$$h_{\mathcal{A},\mathcal{D}}(\mathcal{P}_t) = h_{C,\mathcal{P}^*\mathcal{D}}(t) + O_{\mathcal{P}}(1),$$

[33]If $f : X \dashrightarrow Y$ is a rational map and X is a non-singular variety, then the indeterminacy locus of f has codimension at least 2. In particular, if X is a smooth curve, then f is a morphism; see [35, Lemma V.5.1].

which we rewrite as

$$\left| h_{\mathcal{A},\mathcal{D}}(\mathcal{P}_t) - h_{C,\mathcal{P}^*\mathcal{D}}(t) \right| \le C_{41}(P), \tag{9.18}$$

where we stress that the constant $C_{41}(P)$ depends on P, but does not depend on t.

We divide the inequality (9.12)–(9.15) by $h_C(t)$ and substitute in the estimates provided by (9.16)–(9.18). This yields

$$\begin{aligned}\left| \frac{\hat{h}_{\mathcal{A}_t,\mathcal{D}_t}(\mathcal{P}_t)}{h_C(t)} - \hat{h}_{A,D}(P) \right| &\le \left(C_{38} + \frac{C_{39}}{h_C(t)} \right) \\ &\quad + \left(\frac{h_{C,\mathcal{P}^*\mathcal{D}}(t)}{h_C(t)} + \frac{C_{41}(P)}{h_C(t)} - h_{A,D}(P) \right) + C_{42}. \end{aligned} \tag{9.19}$$

We next do a calculation that will be useful in dealing with (9.19). Thus

$$\begin{aligned} \lim_{\substack{t\in C(\bar{K}) \\ h_C(t)\to\infty}} \frac{h_{C,\mathcal{P}^*\mathcal{D}}(t)}{h_C(t)} &= \deg(\mathcal{P}^*\mathcal{D}) && \text{from Theorem 4.8(e), or more precisely, the special case described in Exercise 4.C, plus the fact that } h_C \text{ is the height on } C \text{ relative to a divisor of degree 1,} \\ &= h_{A,D}(P) + O(1) && \text{from Proposition 4.15.} \end{aligned} \tag{9.20}$$

We take the limsup of (9.19) as $h_C(t) \to \infty$, using (9.20) and the observation that although the constant $C_{41}(P)$ depends on $\mathcal{P}$, it does not depend on t, so $C_{41}(P)/h_C(t)$ vanishes as $h_C(t) \to \infty$. This yields

$$\limsup_{\substack{t\in C(\bar{K}) \\ h_C(t)\to\infty}} \left| \frac{\hat{h}_{\mathcal{A}_t,\mathcal{D}_t}(\mathcal{P}_t)}{h_C(t)} - \hat{h}_{A,D}(P) \right| \le C_{43}, \tag{9.21}$$

where C_{43} is the sum of C_{38} and C_{42} and a bound for the $O(1)$ appearing in (9.20).

We observe that the constant C_{43} appearing in (9.21) does not depend on the point $P \in A\bigl(k(C)\bigr)$. So we are free to replace P with $[m](P)$ for any integer $m \ge 1$. So we do that and exploit the fact that the canonical heights are quadratic forms,

$$\hat{h}_{\mathcal{A}_t,\mathcal{D}_t}\bigl([m](\mathcal{P})_t\bigr) = m^2 \cdot \hat{h}_{\mathcal{A}_t,\mathcal{D}_t}(\mathcal{P}_t) \quad \text{and} \quad \hat{h}_{A,D}\bigl([m](P)\bigr) = m^2 \cdot \hat{h}_{A,D}(P).$$

Hence evaluating (9.20) at $[m](P)$ and dividing both sides by m^2 yields

$$\limsup_{\substack{t\in C(\bar{K}) \\ h_C(t)\to\infty}} \left| \frac{\hat{h}_{\mathcal{A}_t,\mathcal{D}_t}(\mathcal{P}_t)}{h_C(t)} - \hat{h}_{A,D}(P) \right| \le \frac{C_{44}}{m^2}. \tag{9.22}$$

The estimate (9.22) holds for all $m \ge 1$, so we can let $m \to \infty$ to deduce that

$$\limsup_{\substack{t\in C(\bar K)\\ h_C(t)\to\infty}} \left| \frac{\hat h_{\mathcal{A}_t,\mathcal{D}_t}(\mathcal{P}_t)}{h_C(t)} - \hat h_{A,D}(P) \right| = 0.$$

This concludes that proof of the limit formula (9.11), and with it the proof of Theorem 9.3 □

Definition 9.4. Continuing with the notation in Table 1, we note that each point $Q \in A\bigl(\bar K(C)\bigr)$ extends to a unique section

$$\sigma_Q : C \longrightarrow \mathcal{A}$$

whose restriction to the generic fiber is Q. Then for each point $t \in C^\circ(\bar K)$, we define an associated *specialization map*

$$\mathsf{S}_t : A\bigl(\bar K(C)\bigr) \longrightarrow \mathcal{A}_t(\bar K), \quad \mathsf{S}_t(Q) = \sigma_Q(t).$$

Remark 9.5. The specialization map is a well-defined homomorphism from $A\bigl(\bar K(C)\bigr)$ to $\mathcal{A}_t(\bar K)$; see Exercise 9.B

Corollary 9.6 (Specialization Theorem)**.** *We continue with the notation in Table* 1 *and Definition* 9.4*, and we let* $(B/K,\varphi)$ *be the* $K(C)/K$*-trace of* A *as described in Definition* 3.22*. Then the set*

$$C^\circ_{\text{noninj}}(\bar K) := \left\{ t \in C^\circ(\bar K) : \begin{array}{l} \textit{the specialization map} \\ \mathsf{S}_t : A\bigl(\bar K(C)\bigr)/\varphi(B)(\bar K) \\ \qquad\qquad \to \mathcal{A}_t(\bar K)\varphi(B_t)(\bar K) \\ \textit{is not injective} \end{array} \right\} \qquad (9.23)$$

is a set of bounded height. In particular, for any finite extension L/K*, the set* $C^\circ_{\text{noninj}}(L)$ *is finite, i.e., there are only finitely many points* $t \in C^\circ(L)$ *for which the specialization map* S_t *fails to be injective.*

Proof of Corollary 9.6*.* For ease of exposition, we will assume that the $K(C)/K$-trace of A is trivial, where we we refer the reader to Definition 3.22 for the description of the trace. We recall from Definition 5.5 that there is a canonical height pairing on $A\bigl(\bar K(C)\bigr)$ and that there are canonical height pairings on each $\mathcal{A}_t(\bar K)$. Let $P, Q \in A\bigl(\bar K(C)\bigr)$. We use Theorem 9.3 to compute

$$\begin{aligned}
\lim_{\substack{t\in C(\bar K)\\ h_C(t)\to\infty}} \frac{\langle P_t, Q_t\rangle_{\mathcal{A}_t,\mathcal{D}_t}}{h_C(t)} &= \frac{1}{2}\lim_{\substack{t\in C(\bar K)\\ h_C(t)\to\infty}} \frac{\hat h_{\mathcal{A}_t,\mathcal{D}_t}(\mathcal{P}_t+\mathcal{Q}_t) - \hat h_{\mathcal{A}_t,\mathcal{D}_t}(\mathcal{P}_t) - \hat h_{\mathcal{A}_t,\mathcal{D}_t}(\mathcal{Q}_t)}{h_C(t)} \\
&= \frac{1}{2}\Bigl(\hat h_{A,D}(P+Q) - \hat h_{A,D}(P) - \hat h_{A,D}(Q)\Bigr) \\
&\qquad\qquad \text{from Theorem 9.3,} \\
&= \langle P, Q\rangle_{A,D}. \qquad (9.24)
\end{aligned}$$

Let r be the rank of the finitely generated group $A\bigl(\bar K(C)\bigr)$,[34] and choose points $P_1,\ldots,P_r \in A\bigl(\bar K(C)\bigr)$ that are generators for the free part of $A\bigl(\bar K(C)\bigr)/A\bigl(\bar K(C)\bigr)_{\text{tors}}$. Theorem 5.4(b′) says that $\hat h_{A,D}$ extends to a positive definite quadratic form on $A\bigl(\bar K(C)\bigr)\otimes\mathbb{R}$, so the associated height regulator (Definition 5.6) for the independent points $P_1,\ldots,P_r$ satisfies

$$\operatorname{Reg}_{A,D}(P_1,\ldots,P_r) := \det\Bigl(\langle P_i,P_j\rangle_{A,D}\Bigr)_{1\le i,j\le r} > 0.$$

Applying (9.24) to each pair (P_i,P_j) and using the fact that the determinant is a sum of r-fold products of its entries, we obtain the formula

$$\lim_{\substack{t\in C(\bar K)\\ h_C(t)\to\infty}} \frac{\operatorname{Reg}_{\mathcal{D}_t}\bigl(\mathsf{S}_t(P_1),\ldots,\mathsf{S}_t(P_r)\bigr)}{h_C(t)^r} = \operatorname{Reg}_D(P_1,\ldots,P_r) > 0. \qquad (9.25)$$

We consider the subgroup

$$\operatorname{Span}_{\mathbb{Z}}(P_1,\ldots,P_r) \subseteq A\bigl(\bar K(C)\bigr)$$

generated by $P_1,\ldots,P_r$. Restricting the specialization map to this subgroup, we see for $t\in C^\circ(\bar K)$ that

$$\begin{aligned}
&\mathsf{S}_t \text{ restricted to } \operatorname{Span}_{\mathbb{Z}}(P_1,\ldots,P_r) \text{ is not injective} \\
&\quad\iff \sigma_t(P_1),\ldots,\sigma_t(P_r) \text{ are } \mathbb{Z}\text{-linearly dependent in } \mathcal{A}(\bar K) \\
&\quad\iff \operatorname{Reg}_{\mathcal{D}_t}\bigl(\sigma_t(P_1),\ldots,\sigma_t(P_r)\bigr) = 0 \\
&\qquad\qquad \text{since } \langle\,\cdot\,,\,\cdot\,\rangle_{\mathcal{A}_t,\mathcal{D}_t} \text{ is positive definite on } \mathcal{A}_t(\bar K)\otimes\mathbb{R}. \qquad (9.26)
\end{aligned}$$

[34] This is one place that we use the assumption that $\operatorname{Trace}_{K(C)/K}(A)=0$, since in general it is only the quotient $A\bigl(\bar K(C)\bigr)/\varphi\bigl(B(\bar K)\bigr)$ that is finitely generated, where $(B,\varphi)=\operatorname{Trace}_{K(C)/K}(A)$. See Theorem 3.25.

Combining (9.25) and (9.26) yields

$$\bigl\{t \in C^\circ(\bar{K}) : \mathsf{S}_t \text{ restricted to } \operatorname{Span}_{\mathbb{Z}}(P_1,\ldots,P_r) \text{ is not injective}\bigr\} \quad \text{is a set of bounded height.} \tag{9.27}$$

We can thus find a constant C_{45} so that

$$h_C(t) \ge C_{45} \quad\Longrightarrow\quad \mathsf{S}_t : \operatorname{Span}_{\mathbb{Z}}(P_1,\ldots,P_r) \to \mathcal{A}_t(\bar{K}) \text{ is injective.} \tag{9.28}$$

We next observe that for any non-zero point $Q \in A\bigl(\bar{K}(C)\bigr)$, the set

$$\bigl\{t \in C^\circ(\bar{K}) : \mathsf{S}_t(Q) = 0\bigr\} \quad \text{is a finite set,}$$

since it is the intersection of distinct 1-dimensional subvarieties of $\mathcal{A}$. Applying this to the finitely many non-zero points in the torsion subgroup $A\bigl(\bar{K}(C)\bigr)_{\text{tors}}$, we see that there is a constant C_{46} so that

$$h_C(t) \ge C_{46} \quad\Longrightarrow\quad \mathsf{S}_t : A\bigl(\bar{K}(C)\bigr)_{\text{tors}} \to \mathcal{A}_t(\bar{K}) \text{ is injective.} \tag{9.29}$$

The group $A\bigl(K(C)\bigr)$ is the direct sum of its torsion and free pieces, i.e.,

$$A\bigl(K(C)\bigr) = A\bigl(K(C)\bigr)_{\text{tors}} + \operatorname{Span}_{\mathbb{Z}}(P_1,\ldots,P_r);$$

and we know from (9.28) and (9.29) that S_t is injective on each piece for all points $t \in C^\circ(\bar{K})$ satisfying

$$h_C(t) \ge C_{47} := \max\{C_{45}, C_{46}\}.$$

It follows that

$$h_C(t) \ge C_{47} \quad\Longrightarrow\quad \mathsf{S}_t : A\bigl(K(C)\bigr) \to \mathcal{A}_t(\bar{K}) \text{ is injective,}$$

which completes the proof of the first part of Corollary 9.6. The second part is an immediate consequence, since for any finite extension L/K, a set of bounded height in $C(L)$ is finite. □

Example 9.7. We consider the family of elliptic curves and points

$$E_T : y^2 = x^2 - (2T^2 - T - 1)x + T^2, \quad P_T = (0, T).$$

The point P_T has infinite order in $E_T\bigl(\mathbb{Q}(T)\bigr)$. We give two proofs of this assertion. First, if $P_T \in E_T\bigl(\mathbb{Q}(T)\bigr)_{\text{tors}}$, then all multiples of P_T would have polynomial coordinates (this is a function field version of the Nagell–Lutz theorem), but we compute

$$[2]P_T = \left(\tfrac{4T^4-4T^3-3T^2+2T+1}{4T^2}, \tfrac{8T^6-12T^5-14T^4+11T^3+3T^2-3T-1}{8T^3}\right).$$

Second,if $P_T \in E_T\bigl(\mathbb{Q}(T)\bigr)_{\text{tors}}$, then for all $t \in \mathbb{Q}$ we would have $P_t \in E_t(\mathbb{Q})_{\text{tors}}$. We use a computer algebra system to compute

$$\hat{h}_{E_2}(P_2) = 0.7983,$$

which shows that P_2 is not a torsion point, and thus that P_T is not a torsion point. We next do a brief search for values of $t \in \mathbb{Q}$ for which $P_t \in E_t(\mathbb{Q})_{\text{tors}}$. We find the values

$$[2]P_0 = 0, \quad [3]P_1 = [3]P_{-1/2} = 0, \quad \text{and} \quad [4]P_{-1} = 0.$$

So the specialization map fails to be injective for $t \in \{0, \pm 1, -\frac{1}{2}\}$. I suspect that these are the only $t \in \mathbb{Q}$ for which P_t becomes a torsion point, but have not proven it.

Supplementary Material 9.8 (**Beyond the Specialization Theorem**). We briefly discuss some of the many directions that one can go starting from Corollary 9.6.

Remark 9.9 (Higher Dimensional Bases)**.** Suppose that instead of taking a family of abelian varieties over a curve, we consider more generally a family $\mathcal{A} \to T$ of abelian varieties over a base of higher dimension. For each $t \in T(\bar{K})$ we still get a specialization map

$$\mathsf{S}_t : A\bigl(\bar{K}(T)\bigr) \dashrightarrow \mathcal{A}_t(\bar{K}),$$

although the map S_t may not be a morphism even if $\mathcal{A}_t$ is smooth. So we let $T^\circ \subseteq T$ be the Zariski open subset on which the fibers are smooth and S_t is a morphism. Using the notation described in (9.23), one might then ask whether the set $T^\circ_{\text{noninj}}(\bar{K})$ is a set of bounded height. A simple dimension count suggests that this may not be true in general. Thus let $P, Q \in A(\bigl(\bar{K}(T)\bigr)$ be distinct points. We note that

$$\sigma_P(t) = \sigma_Q(t) \quad \Longrightarrow \quad \mathsf{S}_t(P) = \mathsf{S}_t(Q) \quad \Longrightarrow \quad t \in T^\circ_{\text{noninj}},$$

so $\sigma_P(T) \cap \sigma_Q(T) \subseteq T^\circ_{\text{noninj}}$. On the other hand, it is reasonable to expect that if $\dim(T) \ge \dim(A)$, then

$$\begin{aligned}
&\dim\Bigl(\sigma_P(T) \cap \sigma_Q(T)\Bigr) \\
&\quad = \dim\bigl(\sigma_P(T)\bigr) + \dim\bigl(\sigma_Q(T)\bigr) - \dim(\mathcal{A}) \quad && \text{from the “generic” formula} \\
& && \operatorname{codim}(X \cap Y) = \operatorname{codim}(X) + \operatorname{codim}(Y), \\
&\quad = 2\dim(T) - \Bigl(\dim(A) + \dim(T)\Bigr) && \text{since } \dim \sigma_P(T) = \dim \sigma_Q(T) = \dim(T), \\
&\quad = \dim(T) - \dim(A).
\end{aligned}$$

In particular, if $\dim(T) > \dim(A)$, then $\sigma_P(T) \cap \sigma_Q(T)$ is likely to be positive dimensional, so its $\bar{K}$ points do not have bounded height, and hence $T^\circ_{\text{noninj}}(\bar{K})$ is not a set of bounded height.

One solution to this problem is to relax the conclusion. The first general result on abelian variety specialization was proven by Néron [68]. It predates Corollary 9.6 and was used by Néron to prove that there are infinitely many elliptic curves $E/\mathbb{Q}$ with rank $E(\mathbb{Q}) \ge 11$. We won't give Néron's most general result, but the following statement conveys the flavor: Let $\mathcal{A} \to \mathbb{P}^n$ be a family of abelian varieties defined over a number field K. Then $(\mathbb{P}^n)^\circ_{\text{noninj}}(K)$ is a thin set.[35] In particular, using the Weil height H on $\mathbb{P}^n(K)$,

$$\lim_{B \to \infty} \frac{\#\bigl\{t \in (\mathbb{P}^n)^\circ_{\text{noninj}}(K) : H(t) \le B\bigr\}}{\#\bigl\{t \in (\mathbb{P}^n)^\circ(K) : H(t) \le B\bigr\}} = 0.$$

For a more refined statement that deals with all $\bar{K}$ points and implies that $T^\circ_{\text{noninj}}(\bar{K})$ is small, see for example [60].

Remark 9.10 (Rank Jumps)**.** For a 1-dimensional family of abelian varieties, Corollary 9.6 says in particular that $C^\circ_{\text{noninj}}(K)$ is a finite set. Hence for points in the complementary set

$$C^\circ_{\text{inj}}(K) := C(K) \smallsetminus C^\circ_{\text{noninj}}(K),$$

[35] Roughly speaking, a thin subset of $X(K)$ is a finite union of images $Y(K) \to X(K)$ via finite maps $Y \to X$ of degree at least 2. They are the sorts of exceptional sets that appear in Hilbert's irreducibility theorem.

we have

$$\operatorname{rank} \mathcal{A}_t(K) \ge \operatorname{rank} A\bigl(K(C)\bigr) \quad \text{for all } t \in C^\circ_{\text{inj}}(K).$$

This raises the question of when the specialized rank is the same as the generic rank, and when it jumps. Further, since the sign of the functional equation affects the parity of the rank (according to the conjecture of Birch and Swinnerton-Dyer), we expect rank jumps of 1 to occur fairly frequently. A reasonable guess might be that the set

$$\left\{ t \in C^\circ_{\text{inj}}(K) : \operatorname{rank} \mathcal{A}_t(K) = \operatorname{rank} A\bigl(K(C)\bigr) + \begin{pmatrix} \text{1 if mandated by the sign} \\ \text{of the functional equation} \end{pmatrix} \right\}$$

is "large" in an appropriate sense. See for example [16, 78], and the papers that they cite and are cited by, for work on rank jumps. We also note that even if the rank doesn't jump, there is the question of whether $\mathsf{S}\bigl(A\bigl(K(C)\bigr)\bigr)$ equals $\mathcal{A}_t(K)$, or whether some of the points in the image of the specialization map become further divisible in $\mathcal{A}_t(K)$. See for example [83] for work on the indivisibility problem.

Remark 9.11 (Unlikely Intersections)**.** Let $\pi : \mathcal{E} \to C$ be an elliptic surface, i.e., a family of elliptic curves, and let $\mathcal{P} : C \to \mathcal{E}$ be a section of infinite order, all defined over a number field K. Corollary 9.6 tells us that the set

$$C^\circ_{\text{noninj}}(\mathcal{P}) := \bigl\{ t \in C^\circ \bar{K}) : \mathcal{P}_t \in \mathcal{E}_t(\bar{K})_{\text{tors}} \bigr\}$$

is a set of bounded height, but it is unlikely to be a finite set. The reason is that for each $n \ge 1$, the image $[n]\mathcal{P}(C)$ will intersect the image of the zero-section $\mathcal{O}(C)$ in roughly n^2 points (admittedly counted with multiplicity). So for each n we get roughly $O(n^2)$ points $t \in C(\bar{K})$ with $[n]\mathcal{P}_t = \mathcal{O}_t$, and thus $O(n^2)$ points in $C^\circ_{\text{noninj}}(\mathcal{P})$. We stress that this occurs because both $[n]\mathcal{P}(C)$ and $\mathcal{O}(C)$ are curve lying in the surface $\mathcal{E}$, so it is likely that they intersect. Further the curve $[n]\mathcal{P}(C)$ has increasing degree as $n \to \infty$, so we expect to see more-and-more intersection points as $n \to \infty$.

However, suppose that we consider two independent sections $\mathcal{P}, \mathcal{Q} : C \to \mathcal{A}$. Then

$$C^\circ_{\text{noninj}}(\mathcal{P}) \cap C^\circ_{\text{noninj}}(\mathcal{Q})$$

is the intersection of two "unrelated" subsets of $C(\bar{K})$ of bounded height, so one might guess that it is a finite set. Similarly, if we take a family $\pi : \mathcal{A} \to C$ of abelian varieties of dimension g with $g \ge 2$, then $[n]\mathcal{P}(C)$ and $\mathcal{O}(C)$ are curves sitting in a variety of dimension $g+1 \ge 3$, so we might guess that $C^\circ_{\text{noninj}}(\mathcal{P})$ is already finite (subject to an appropriate non-degeneracy condition). These sorts of ideas lead to the theory of *unlikely intersections*, a currently very active field of arithmetic geometry. An influential paper [57] of Masser and Zannier proved that for the following family of elliptic curve and points

$$\mathcal{E}_T : y^2 = x(x-1)(x-T), \quad \mathcal{P}_T = \bigl(2, \sqrt{2(2-T)}\,\bigr), \quad \mathcal{Q}_T = \bigl(3, \sqrt{6(3-T)}\,\bigr),$$

the set

$$\{ t \in \mathbb{C} : \mathcal{P}_t \text{ and } \mathcal{Q}_t \text{ are both in } \mathcal{E}_t(\mathbb{C})_{\text{tors}} \} \quad \text{is finite.}$$

We refer the reader to a recent survey article by Capuano [10] that describes problems and recent progress on unlikely intersection on families of abelian varieties, as well as to an older, but longer and more comprehensive, monograph of Zannier [92] that contains an overview of the larger field of unlikely intersections.

Supplementary Material 9.12 (**Stronger Height Specialization Estimates**). Theorem 9.3 has been strengthened in various ways under various hypotheses. We gather some of these improvements in the following theorem.

Theorem 9.13. *We use the notation described in Table* 1*, and we let* $h_C^+(t) = \max\bigl\{h_C(t), 1\bigr\}$.

(a) (Ingram, special case of [41])

$$\hat{h}_{\mathcal{A}_t,\mathcal{D}_t}(\mathcal{P}_t) = \hat{h}_{A,D}(P) \cdot h_C(t) + \begin{cases} O\bigl(h_C^+(t)^{2/3}\bigr) & \textit{in general,} \\ O\bigl(h_C^+(t)^{1/2}\bigr) & \textit{if } C \cong \mathbb{P}^1. \end{cases}$$

(b) (Tate [90]) *Assume that* $\dim(A) = 1$, *i.e.,* A *is an elliptic curve. Then* $t \mapsto \hat{h}_{\mathcal{A}_t,\mathcal{D}_t}(\mathcal{P}_t)$ *is a Weil height on* C *associated to a divisor in* $\mathrm{Div}(C)\otimes\mathbb{Q}$ *of degree* $\hat{h}_{A,D}(P)$. *In particular,*

$$\hat{h}_{\mathcal{A}_t,\mathcal{D}_t}(\mathcal{P}_t) = \hat{h}_{A,D}(P)\cdot h_C(t) + \begin{cases} O\bigl(h_C^+(t)^{1/2}\bigr) & \text{in general,}\\ O(1) & \text{if } C \cong \mathbb{P}^1.\end{cases}$$

(c) (Silverman, special case of [84, 86, 87]) *Assume that* $\dim(A) = 1$ *and that* $C \cong \mathbb{P}^1$, *so* $K(C) = K(T)$ *is the field of rational functions. Then there is a modulus* m *and, for each* $a \in \mathbb{Z}/m\mathbb{Z}$, *a power series* $f_a(z)$ *converging in some neighborhood of* 0 *such that*

$$\hat{h}_{\mathcal{A}_t,\mathcal{D}_t}(\mathcal{P}_t) = \hat{h}_{A,D}(P)\cdot\log|t| + f_{(t \bmod m)}\bigl(|t|^{-1}\bigr) \quad \text{for all sufficiently large } t\in\mathbb{Z}.$$

Exercises for Section 9.

Exercise 9.A. Let T/K be a smooth projective variety, and let $\pi : \mathcal{A} \to T$ be a morphism such that the generic fiber of π is an abelian variety over $K(T)$. Prove an analogue of Lemma 9.1. More precisely, let h_T be a Weil height function on T relative to an ample variety, let $\mathcal{D} \in \mathrm{Div}(A)$ be a (symmetric) divisor, and let $\mathcal{A}^\circ \subseteq \mathcal{A}$ be a subvariety such that π makes $\mathcal{A}^\circ$ into an abelian scheme over an appropriate subvariety of T. Prove that

$$\bigl|\hat{h}_{\mathcal{A},\mathcal{D}}(x) - h_{\mathcal{A},\mathcal{D}}(x)\bigr| = O\Bigl(h_T\bigl(\pi(x)\bigr)\Bigr) \quad \text{for all } x \in \mathcal{A}^\circ(\bar{K}),$$

where the big-O constant does not depend on x.

Exercise 9.B. Let $t \in C^\circ(\bar{K})$. Prove that the specialization map described in Definition 9.4 is a well-defined homomorphism from $A\bigl(\bar{K}(C)\bigr)$ to $\mathcal{A}_t(\bar{K})$.

Exercise 9.C. Feel free to use a computer algebra system for this problem; see Exercise 5.G. We consider the elliptic curve

$$E_T : y^2 = x^3 + x + T^2 \quad \text{and point} \quad P_T = (0,T) \in E_T\bigl(\mathbb{Q}(T)\bigr)$$

defined over $\mathbb{Q}(T)$.

(a) Compute $[m](P_T)$ for $1 \le m \le 12$ and make a table of the values

$$\frac{1}{m^2}\deg x\bigl([m](P_T)\bigr). \tag{9.30}$$

Use your data to guess the value of $\hat{h}_{E_T,2(0)}(P_T)$, which is the limit as $m \to \infty$ of (9.30).

(b) Make a table of values of

$$\frac{\hat{h}_{E_t,2(0)}(P_t)}{\log|t|}$$

for a sequence of large integers t. Does it seem to be converging to the value of $\hat{h}_{E_T,2(0)}(P_T)$ from (a)?

(c) Make a table of values of the difference

$$\hat{h}_{E_t,2(0)}(P_t) - \hat{h}_{E_T,2(0)}(P_T)\cdot h(t) \tag{9.31}$$

for a sequence of large integers, and more generally for some $t \in \mathbb{Q}$ with large numerators and/or denominators. Does the difference (9.31) seem to be converging? If so, can you guess a likely value? (This example illustrates Theorem 9.13(c).)

Exercise 9.D. Consider the family of elliptic curves and points

$$E_T : y^2 = x^3 + (3 - T^2)x + T^2, \quad P_T = (0,T), \quad Q_T = (2,1).$$

(a) Find a value of $t \in \mathbb{Q}$ such that P_t and Q_t generate a rank 2 subgroup of $E_t(\mathbb{Q})$.
(b) Use (a) to conclude that that P_T and Q_T generate a rank 2 subgroup of $E_T\bigl(\mathbb{Q}(T)\bigr)$.
(c) Search for some values of $t \in \mathbb{Q}$ such that P_t and Q_t generate a rank 1 subgroup of $E_t(\mathbb{Q})$. *Hint*: I found 5 such values, but there could be more.

10. Further Topics for Abelian Varieties over Global Fields

These notes cover only a small fraction of the theory of abelian varieties defined over number fields and more general global fields. Among the many important topics that have been omitted, we mention:

- The rank of the Mordell–Weil group $A(K)$: How to compute it? What is the average value (in families)? Is it bounded or unbounded for fixed K and $\dim(A)$?
- The size of the torsion subgroup $A(K)_{\text{tors}}$: Is it uniformly bounded in terms of $[K : \mathbb{Q}]$ and $\dim(A)$?
- Analytic theory: Does $L(A/K, s)$ have an analytic continuation? Is the conjecture of Birch and Swinnerton-Dyer true?
- Unlikely intersections: Unlikely intersections on abelian varieties. Unlikely intersections in moduli spaces of abelian varieties.
- Special values: Generating abelian extension (explicit class field theory) via the theory of complex multiplication.
- Transcendence theory: Trancendental values of periods and other quantities associated to abelian varieties defined over $\bar{\mathbb{Q}}$.

11. Some Places to Read About Abelian Varieties

We list some reference books and other sources that discuss the theory of abelian varieties, concentrating mostly on the geometry.

[6] Birkenhake, Christina and Lange, Herbert, *Complex abelian varieties*, 2004
[18] Conrad, Brian, *Abelian varieties* (course notes), 2017.
[19] Cornell, Gary and Silverman, Joseph H. (editors), *Arithmetic geometry*, 1986
  - Michael Rosen, Abelian varieties over $\mathbb{C}$ (pp. 79–101)

- J. S. Milne, Abelian varieties (pp. 103—150)
- J. S. Milne, Jacobian varieties (pp. 167-–212)

[24] Debarre, Olivier, *Complex tori and abelian varieties*, 2005

[33] Griffiths, Phillip and Harris, Joseph, *Principles of algebraic geometry*, 1994 (reprint of 1978 edition)
- Complex tori and abelian varieties, (chapter 2, section 6)
- Curves and their Jacobians, (chapter 2, section 7)

[39] Hindry, Marc and Silverman, Joseph H., *Diophantine geometry*, 2000
- The Geometry of Curves and Abelian Varieties (Part A)

[46] Lang, Serge, *Abelian varieties*, 1983 (reprint of 1959 edition)

[50] Lange, Herbert, *Abelian varieties over the complex numbers*, 2023

[65] Mumford, David, *Abelian varieties*, 2008 (reprint of 1974 edition)

[66] Murty, V. Kumar, *Introduction to abelian varieties*, 1993

[70] Orr, Martin, Abelian varieties blog covering many topics.

[85] Silverman, J.H., *Advanced topics in the arithmetic of elliptic curves*, 1994

[88] Silverman, J.H., *The arithmetic of elliptic curves* (2nd ed), 2009

References

[1] F. Amoroso, S. David, and U. Zannier. On fields with Property (B). *Proc. Amer. Math. Soc.*, 142(6):1893–1910, 2014.

[2] F. Amoroso and D. Masser. Erratum: Lower bounds for the height in Galois extensions (Bull. London Math. Soc. 48 (2016) 1008–1012) [MR3608945]. *Bull. Lond. Math. Soc.*, 48(6):1050, 2016.

[3] F. Amoroso and D. Masser. Lower bounds for the height in Galois extensions. *Bull. Lond. Math. Soc.*, 48(6):1008–1012, 2016.

[4] M. H. Baker and J. H. Silverman. A lower bound for the canonical height on abelian varieties over abelian extensions. *Math. Res. Lett.*, 11(2-3):377–396, 2004.

[5] D. Bertrand. Minimal heights and polarizations on group varieties. *Duke Math. J.*, 80(1):223–250, 1995.

[6] C. Birkenhake and H. Lange. *Complex abelian varieties*, volume 302 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 2004.

[7] E. Bombieri and U. Zannier. A note on heights in certain infinite extensions of $\mathbb{Q}$. *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl.*, 12:5–14, 2001.

[8] J. Boxall. Une propriété des hauteurs locales de Néron-Tate sur les variétés abéliennes. *J. Théor. Nombres Bordeaux*, 7(1):111–119, 1995. Les Dix-huitièmes Journées Arithmétiques (Bordeaux, 1993).

[9] G. S. Call and J. H. Silverman. Canonical heights on varieties with morphisms. *Compositio Math.*, 89(2):163–205, 1993.

[10] L. Capuano. An overview on problems of unlikely intersections in families of abelian varieties. *Expo. Math.*, 41(3):603–617, 2023.

[11] J. Cassels. Review of the article "Les formes bilinéaires de Néron et Tate" by S. Lang. *Mathematical Reviews*, MR0176984, 1964. `mathscinet.ams.org/mathscinet/article?mr=176984`.

[12] M. Çiperiani and E. Ozman. Local to global trace questions and twists of genus one curves. *Proc. Amer. Math. Soc.*, 143(9):3815–3826, 2015.

[13] S. Checcoli and A. Fehm. On the Northcott property and local degrees. *Proc. Amer. Math. Soc.*, 149(6):2403–2414, 2021.

[14] S. Checcoli and A. Fehm. Constructing totally p-adic numbers of small height. *Int. J. Number Theory*, 18(3):501–514, 2022.

[15] S. Checcoli and M. Widmer. On the Northcott property and other properties related to polynomial mappings. *Math. Proc. Cambridge Philos. Soc.*, 155(1):1–12, 2013.

[16] J.-L. Colliot-Thélène. Point générique et saut du rang du groupe de Mordell-Weil. *Acta Arith.*, 196(1):93–108, 2020.

[17] B. Conrad. Chow's K/k-image and K/k-trace, and the Lang-Néron theorem. *Enseign. Math. (2)*, 52(1-2):37–108, 2006.

[18] B. Conrad. Abelian varieties, 2017. AMS Open Math Notes, `www.ams.org/open-math-notes/omn-view-listing?listingId=110657`.

[19] G. Cornell and J. H. Silverman, editors. *Arithmetic geometry*. Springer-Verlag, New York, 1986. Papers from the conference held at the University of Connecticut, Storrs, Connecticut, July 30–August 10, 1984.

[20] S. David. Fonctions thêta et points de torsion des variétés abéliennes. *Compositio Math.*, 78(2):121–160, 1991.
[21] S. David. Minorations de hauteurs sur les variétés abéliennes. *Bull. Soc. Math. France*, 121(4):509–544, 1993.
[22] S. David and M. Hindry. Minoration de la hauteur de Néron-Tate sur les variétés abéliennes de type C. M. *J. Reine Angew. Math.*, 529:1–74, 2000.
[23] S. David and P. Philippon. Minorations des hauteurs normalisées des sous-variétés des puissances des courbes elliptiques. *Int. Math. Res. Pap. IMRP*, (3):Art. ID rpm006, 113, 2007.
[24] O. Debarre. *Complex tori and abelian varieties*, volume 11 of *SMF/AMS Texts and Monographs*. American Mathematical Society, Providence, RI; Société Mathématique de France, Paris, 2005. Translated from the 1999 French edition by Philippe Mazaud.
[25] V. A. Dem'janenko. Tate height, and the representation of numbers by binary forms. *Izv. Akad. Nauk SSSR Ser. Mat.*, 38:459–470, 1974.
[26] E. Dobrowolski. On a question of Lehmer and the number of irreducible factors of a polynomial. *Acta Arith.*, 34:391–401, 1979.
[27] A. Dubickas. The size of algebraic integers with many real conjugates. *Ars Math. Contemp.*, 14(1):165–176, 2018.
[28] R. Dvornicich and U. Zannier. On the properties of Northcott and of Narkiewicz for fields of algebraic numbers. *Funct. Approx. Comment. Math.*, 39:163–173, 2008.
[29] G. Faltings. Endlichkeitssätze für abelsche Varietäten über Zahlkörpern. *Invent. Math.*, 73(3):349–366, 1983.
[30] G. Faltings. Finiteness theorems for abelian varieties over number fields. In *Arithmetic geometry (Storrs, Conn., 1984)*, pages 9–27. Springer, New York, 1986. Translated from the German original [Invent. Math. **73** (1983), no. 3, 349–366; MR0718935; ibid. **75** (1984), no. 2, 381] by Edward Shipz.
[31] A. Fehm. Three counterexamples concerning the Northcott property of fields. *Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl.*, 29(2):309–314, 2018.
[32] É. Gaudron and G. Rémond. Nombre de petits points sur une variété abélienne, Décembre 2022. preprint, `https://lmbp.uca.fr/~gaudron/art21.pdf`.
[33] P. Griffiths and J. Harris. *Principles of algebraic geometry*. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1994. Reprint of the 1978 original.
[34] P. Habegger. Small height and infinite nonabelian extensions. *Duke Math. J.*, 162(11):2027–2076, 2013.
[35] R. Hartshorne. *Algebraic geometry*. Graduate Texts in Mathematics, No. 52. Springer-Verlag, New York-Heidelberg, 1977.
[36] M. Hindry. Sur les hauteurs locales de néron sur les variétés abéliennes. Prépublications Mathématiques de l'U.R.A. 212 'Théories Géométriques' no. 51, Université Paris 7 1993, `webusers.imj-prg.fr/~marc.hindry/Neron.pdf`.
[37] M. Hindry and J. H. Silverman. The canonical height and integral points on elliptic curves. *Invent. Math.*, 93(2):419–450, 1988.
[38] M. Hindry and J. H. Silverman. On Lehmer's conjecture for elliptic curves. In *Séminaire de Théorie des Nombres, Paris 1988–1989*, volume 91 of *Progr. Math.*, pages 103–116. Birkhäuser Boston, Boston, MA, 1990.

[39] M. Hindry and J. H. Silverman. *Diophantine geometry*, volume 201 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000. An introduction.
[40] G. Höhn. On a theorem of Garza regarding algebraic numbers with real conjugates. *Int. J. Number Theory*, 7(4):943–945, 2011.
[41] P. Ingram. Variation of the canonical height in a family of polarized dynamical systems, 2021. `arXiv:2104.12877`.
[42] S. Kamienny. Torsion points on elliptic curves and q-coefficients of modular forms. *Invent. Math.*, 109(2):221–229, 1992.
[43] S. Kawaguchi and J. H. Silverman. Dynamical canonical heights for Jordan blocks, arithmetic degrees of orbits, and nef canonical heights on abelian varieties. *Trans. Amer. Math. Soc.*, 368(7):5009–5035, 2016.
[44] S. Lang. Les formes bilinéaires de Néron et Tate. In *Séminaire Bourbaki, 16e année: 1963/64, Fasc. 3*, pages Exposé 274, 11. Secrétariat mathématique, Paris, 1964.
[45] S. Lang. *Elliptic curves: Diophantine analysis*, volume 231 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin-New York, 1978.
[46] S. Lang. *Abelian varieties*. Springer-Verlag, New York-Berlin, 1983. Reprint of the 1959 original.
[47] S. Lang. *Fundamentals of Diophantine geometry*. Springer-Verlag, New York, 1983.
[48] S. Lang. *Algebraic number theory*, volume 110 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1994.
[49] S. Lang and A. Néron. Rational points of abelian varieties over function fields. *Amer. J. Math.*, 81:95–118, 1959.
[50] H. Lange. *Abelian varieties over the complex numbers—a graduate course*. Grundlehren Text Editions. Springer, Cham, 2023.
[51] M. Laurent. Minoration de la hauteur de Néron-Tate. In *Séminaire de Théorie des Nombres*, Progress in Mathematics, pages 137–151. Birkhäuser, 1983. Paris 1981–1982.
[52] T. LMFDB Collaboration. The L-functions and modular forms database. `https://www.lmfdb.org`, 2023. [Online; accessed 8 December 2023].
[53] N. R. Looper and J. H. Silverman. A lehmer-type height lower bound for abelian surfaces over function fields. *Trans. Amer. Math. Soc.*, 2024. `doi.org/10.1090/tran/9024`, `arxiv.org/abs/2108.09577`.
[54] K. Mahler. An inequality for the discriminant of a polynomial. *Michigan Math. J.*, 11:257–262, 1964.
[55] J. I. Manin. The Tate height of points on an Abelian variety, its variants and applications. *Izv. Akad. Nauk SSSR Ser. Mat.*, 28:1363–1390, 1964.
[56] D. Masser. Unpublished letter to Daniel Bertrand, 17th Nov. 1986.
[57] D. Masser and U. Zannier. Torsion anomalous points and families of elliptic curves. *Amer. J. Math.*, 132(6):1677–1691, 2010.
[58] D. W. Masser. Small values of the quadratic part of the Néron-Tate height on an abelian variety. *Compositio Math.*, 53(2):153–170, 1984.
[59] D. W. Masser. Counting points of small height on elliptic curves. *Bull. Soc. Math. France*, 117(2):247–265, 1989.
[60] D. W. Masser. Specializations of finitely generated subgroups of abelian varieties. *Trans. Amer. Math. Soc.*, 311(1):413–424, 1989.

[61] B. Mazur. Modular curves and the Eisenstein ideal. *Inst. Hautes Études Sci. Publ. Math.*, (47):33–186 (1978), 1977. With an appendix by Mazur and M. Rapoport.

[62] J. Mello and M. Sha. On the properties of Northcott and Narkiewicz for elliptic curves. *Int. J. Number Theory*, 18(10):2129–2144, 2022.

[63] L. Merel. Bornes pour la torsion des courbes elliptiques sur les corps de nombres. *Invent. Math.*, 124(1-3):437–449, 1996.

[64] M. J. Mossinghoff, G. Rhin, and Q. Wu. Minimal Mahler measures. *Experiment. Math.*, 17(4):451–458, 2008.

[65] D. Mumford. *Abelian varieties*, volume 5 of *Tata Institute of Fundamental Research Studies in Mathematics*. Published for the Tata Institute of Fundamental Research, Bombay; by Hindustan Book Agency, New Delhi, 2008. With appendices by C. P. Ramanujam and Yuri Manin, Corrected reprint of the second (1974) edition.

[66] V. K. Murty. *Introduction to abelian varieties*, volume 3 of *CRM Monograph Series*. American Mathematical Society, Providence, RI, 1993.

[67] A. Néron. Valeur asymptotique du nombre des points rationnels de hauteur bornée sur une courbe elliptique. short communication in Proc. Internat. Congr. Mathematicians (August, 1958, Edinburgh), Cambridge Univ. Press, Cambridge, 1960.

[68] A. Néron. Problèmes arithmétiques et géométriques rattachés à la notion de rang d'une courbe algébrique dans un corps. *Bull. Soc. Math. France*, 80:101–166, 1952.

[69] A. Néron. Quasi-fonctions et hauteurs sur les variétés abéliennes. *Ann. of Math. (2)*, 82:249–331, 1965.

[70] M. Orr. Abelian varieties. Blog covering many topics, `www.martinorr.name/blog/category/abelian-varieties/`.

[71] J. Park, B. Poonen, J. Voight, and M. M. Wood. A heuristic for boundedness of ranks of elliptic curves. *J. Eur. Math. Soc. (JEMS)*, 21(9):2859–2903, 2019.

[72] F. Pazuki. Minoration de la hauteur de Néron-Tate sur les surfaces abéliennes. *Manuscripta Math.*, 142(1-2):61–99, 2013.

[73] P. Philippon. Sur des hauteurs alternatives. I. *Math. Ann.*, 289(2):255–283, 1991.

[74] P. Philippon. Sur des hauteurs alternatives. II. *Ann. Inst. Fourier (Grenoble)*, 44(4):1043–1065, 1994.

[75] P. Philippon. Sur des hauteurs alternatives. III. *J. Math. Pures Appl. (9)*, 74(4):345–365, 1995.

[76] N. Ratazzi. Intersection de courbes et de sous-groupes et problèmes de minoration de hauteur dans les variétés abéliennes C.M. *Ann. Inst. Fourier (Grenoble)*, 58(5):1575–1633, 2008.

[77] G. Rémond. Généralisations du problème de Lehmer et applications à la conjecture de Zilber-Pink. In *Around the Zilber-Pink conjecture/Autour de la conjecture de Zilber-Pink*, volume 52 of *Panor. Synthèses*, pages 243–284. Soc. Math. France, Paris, 2017.

[78] C. Salgado. On the rank of the fibers of rational elliptic surfaces. *Algebra Number Theory*, 6(7):1289–1314, 2012.

[79] J.-P. Serre and J. Tate. Good reduction of abelian varieties. *Ann. of Math. (2)*, 88:492–517, 1968.

[80] J. H. Silverman. Lower bound for the canonical height on elliptic curves. *Duke Math. J.*, 48(3):633–648, 1981.

[81] J. H. Silverman. *The Néron–Tate height on elliptic curves*. ProQuest LLC, Ann Arbor, MI, 1982. Thesis (Ph.D.)–Harvard University.

[82] J. H. Silverman. Lower bounds for height functions. *Duke Math. J.*, 51(2):395–403, 1984.

[83] J. H. Silverman. Divisibility of the specialization map for families of elliptic curves. *Amer. J. Math.*, 107(3):555–565, 1985.

[84] J. H. Silverman. Variation of the canonical height on elliptic surfaces. I. Three examples. *J. Reine Angew. Math.*, 426:151–178, 1992.

[85] J. H. Silverman. *Advanced topics in the arithmetic of elliptic curves*, volume 151 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1994.

[86] J. H. Silverman. Variation of the canonical height on elliptic surfaces. II. Local analyticity properties. *J. Number Theory*, 48(3):291–329, 1994.

[87] J. H. Silverman. Variation of the canonical height on elliptic surfaces. III. Global boundedness properties. *J. Number Theory*, 48(3):330–352, 1994.

[88] J. H. Silverman. *The arithmetic of elliptic curves*, volume 106 of *Graduate Texts in Mathematics*. Springer, Dordrecht, second edition, 2009.

[89] J. H. Silverman. Height estimates for equidimensional dominant rational maps. *J. Ramanujan Math. Soc.*, 26(2):145–163, 2011.

[90] J. Tate. Variation of the canonical height of a point depending on a parameter. *Amer. J. Math.*, 105(1):287–294, 1983.

[91] M. Widmer. On certain infinite extensions of the rationals with Northcott property. *Monatsh. Math.*, 162(3):341–353, 2011.

[92] U. Zannier. *Some problems of unlikely intersections in arithmetic and geometry*, volume 181 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2012. With appendixes by David Masser.

[93] J. G. Zarhin. A remark on endomorphisms of abelian varieties over function fields of finite characteristic. *Izv. Akad. Nauk SSSR Ser. Mat.*, 38:471–474, 1974.

List of Notation

K	a field, e.g., number field, local field, algebraically closed field, 2
$\bar{K}$	an algebraic closure of K, 2
X/K	a smooth projective algebraic variety, defined over K, 2
$K(X)$, $\bar{K}(X)$	the function field of X over K, respectively over $\bar{K}$, 2
$\operatorname{Div}(X)$	the group of geometric divisors of X, i.e., divisors defined over $\bar{K}$, 2
$\operatorname{div}(f)$	the divisor of a function $f \in \bar{K}(X)$, 2
$\sim$	linear equivalence, $D \sim D'$ if $D - D' = \operatorname{div}(f)$ for some $f \in \bar{K}(X)$, 2
$\mathcal{L}(D)$	$= H^0\bigl(X, \mathcal{O}_X(D)\bigr) = \bigl\{f \in \bar{K}(X) : \operatorname{div}(f) + D \ge 0\bigr\}$, 2
$\ell(D)$	$= \dim_{\bar{K}} \mathcal{L}(D)$, 2
$\equiv$	algebraic equivalence of divisors, 2
$\operatorname{Pic}(X)$	$= \operatorname{Div}(X)/\sim$, the Picard group of X, 2
$\operatorname{Div}^0(X)$	$= \{D \in \operatorname{Div}(X) : D \equiv 0\}$, 2
$\operatorname{Pic}^0(X)$	$= \operatorname{Div}^0(X)/\sim$, 2
$\operatorname{NS}(X)$	$= \operatorname{Div}(X)/\equiv$, the Neron–Severi group of X, 2
$\rho(X)$	$= \operatorname{rank} \operatorname{NS}(X)$, 2
$\operatorname{End}(X)$	the ring of endomorphisms $X \to X$, 2
$\operatorname{Aut}(X)$	the group of automorphisms $X \to X$, i.e., $\operatorname{Aut}(X) = \operatorname{End}(X)^*$, 2
φ_D	embedding $f_D : X \hookrightarrow \mathbb{P}^{\ell(D)-1}$ associated to a very ample D, 3
T_Q	translation-by-Q map, 5
$\operatorname{Hom}(A, B)$	the group of isogenies from A to B, 5
$\operatorname{End}(A)$	the ring of endomorphisms of A, 5
φ_D	the map $\varphi_D : A \to \hat{A}$ induced by the divisor D, 6
$\operatorname{End}(A)_{\mathbb{Q}}$	$= \operatorname{End}(A) \otimes \mathbb{Q}$, 6
Φ_D	$= \varphi_H^{-1} \circ \varphi_D$ mapping $\operatorname{NS}(A)_{\mathbb{Q}} \vert to \operatorname{End}(A)_{\mathbb{Q}}$, 7
$A(K)_{\text{tors}}$	torsion subgroup of $A(K)$, 8
$\operatorname{rank} A(K)$	rank of the group $A(K)$, 8
$(B/k, \varphi)$	the K/k trace of an abelian variety, 9
$\operatorname{Twist}(A/K)$	the set of twists of an abelian variety, 10
$\boldsymbol{\mu}_m$	the group of mth roots of unity, 10
ξ_D	the twisting isomorphism $\xi_D : A_D \to A$, 10
K	a global field, i.e., a number field or the function field of a curve., 12
M_K	a complete set of normalized absolute values on K, 12
M_K^∞	the archimedean absolute values in M_K, 12
M_K°	the non-archimedean absolute values in M_K, 12
K_v	the completion of K at the absolute value $v \in M_K$, 12
H_K, h_K	height on $\mathbb{P}^N(K)$ relative to K, 12
H, h	absolute height on $\mathbb{P}^N(\bar{K})$, 12
M_K	the set of standard absolute values on K, 13
$\lvert\alpha\rvert_\infty$	the standard archimedean absolute value on $\mathbb{Q}$, 13
$\lvert\alpha\rvert_p$	the standard p-adic absolute value on $\mathbb{Q}$, 13
M_K	the set of standard absolute values on K, 13
h_K	the logarithmic Weil height on $\mathbb{P}^N$, 13
h	the absolute logarithmic Weil height on $\mathbb{P}^N$, 13

f_D	projective embedding $f_D : D \to \mathbb{P}^{\ell(D)-1}$, 15
$h_{X,D}$	the absolute logarithmic Weil height on X relative to D, 15
C/k	a smooth projective curve, 17
$k(C)$	the function field of C, 17
h	the absolute logarithmic Weil height on $\mathbb{P}^N$ over a function field, 17
$h(X)$	height of a variety, 18
$\mathcal{A}_g$	moduli space of principally polarized abelian of dimension g, 20
$h_{\mathcal{A}_g}$	Weil height function on the moduli space $\mathcal{A}_g$, 20
$j(A)$	image of A in the moduli space $\mathcal{A}_g$, 20
$h(A/K)$	the height of the principally polarized abelian variety A/K, 20
$h_K^{(m)}$	m-power free height, 22
$\asymp$	asymptotic functions, 22
ρ	2 if D is symmetric, 1 if D is anti-symmetric, 23
$\hat{h}_{A,D}(P)$	the canonical height function on an abelian variety, 23
$\hat{h}_{A,D}(P)$	the canonical height function on an abelian variety, 25
$\langle\,\cdot\,,\,\cdot\,\rangle_{A,D}$	the Neron–Tate canonical height pairing, 30
$\operatorname{Reg}_D(P_1,\ldots,P_r)$	the Neron–Tate regulator of $P_1,\ldots,P_r$, 30
$\operatorname{Reg}_D(A/K)$	the Neron–Tate regulator of A/K, 30
$\hat{h}_{X,\varphi,D}$	canonical height associated to dynamical system $\varphi\, X \to X$, 31
$A(K)_{\mathbb{R}}$	the $\mathbb{R}$-vector space $A(K)\otimes\mathbb{R}\cong\mathbb{R}^r$, 33
$A(K)_{\mathbb{Z}}$	the image of $A(K)$ in $A(K)\otimes\mathbb{R}$, 34
$\lVert\cdot\rVert_{A,D}$	canonical height norm on $A(K)_{\mathbb{R}}$, 34
$\mathcal{N}\bigl(A(K),h_{A,D},T\bigr)$	point counting function, 34
$\operatorname{dist}_v(P,D)$	the v-adic distance from P to D, 37
$\lambda_{\mathbb{P}^n,D,v}$	local height on $\mathbb{P}^n$, 37
$\hat{\lambda}_{A,D,v}$	local canonical height on $A(\bar{K}_v)$, 38
$\mathbb{B}_{D,v}$	"Bernoulli polynomial" in local height formula, 40
$\mathfrak{H}_g$	Siegel upper half space, 40
$\theta(\mathbf{z},\tau)$	theta function an an abelian variety, 40
$\Theta(\mathbf{z},\tau)$	divisor an an abelian variety associated to a theta function, 40
K^{ab}	the maximal abelian extension of K, 47
$\mathsf{D}_E^{\mathrm{qm}}$	the minimal discriminant of E, 49
$\mathsf{N}_E^{\mathrm{qm}}$	the conductor of E, 49
$\mathsf{S}_E^{\mathrm{qm}}$	the Szpiro-ratio of E, 49
K_P	the field of definition of the point $P\in\mathbb{P}^N(\bar{K})$, 51
$\hat{h}_{\varphi}$	canonical height associated to dynamical system $\varphi\,\mathbb{P}^N\to\mathbb{P}^N$, 53
$A/K(C)$	an abelian variety defined over $K(C)$, 55
$(\mathcal{A},\pi)$	a family $\pi:\mathcal{A}\to C$ of abelian varieties, 55
S_t	the specialization map $A\bigl(\bar{K}(C)\bigr)\to\mathcal{A}_t(\bar{K})$, 61
σ_Q	the section $C\to\mathcal{A}$ associated to a point $Q\in A\bigl(\bar{K}(C)\bigr)$, 61
$C^{\circ}_{\mathrm{noninj}}(\bar{K})$	set of t where specialization is not injective, 61

Index

Appendix A. Projects

This section describes projects that we might work on during the AWS.

Project I: Trace Relations on Abelian Varieties.
Let A/K be an abelian variety defined over a field K, and let L/K be finite a Galois extension with Galois group $G(L/K)$. Then we can define a trace map from $A(L)$ to $A(K)$ by the formula

$$\mathrm{Trace}_{A,L/K} : A(L) \longrightarrow A(K), \quad \mathrm{Trace}_{A,L/K}(P) = \sum_{\sigma\in G(L/K)} \sigma(P).$$

This project asks you to investigate this trace map. Here are some possible questions, but you are encouraged to formulate questions of your own.

(a) Is there a good criterion for when $\mathrm{Trace}_{A,L/K}$ is surjective? You might consider this question for K a number field, a function field, a local field such as $\mathbb{Q}_p$, or a finite field.

(b) Suppose that

$$\mathrm{Trace}_{A,L_v/K_v} : A(L_v) \longrightarrow A(K_v)$$

is surjective for every completion of L/K. Does this imply that $\mathrm{Trace}_{A,L/K}$ is surjective? If not, can you find an obstruction, for example an obstruction that lives in a Galois cohomology group?

(c) Same questions as (a) and (b), but restrict the field. For example, take $[L : K] = 2$, so $A(L)$ is (almost) isomorphic to $A(K) \oplus A^\chi(K)$, where A^χ is the L/K quadratic twist of A/K. Or if A admits CM by dth roots of unity, you could take $G(L/K)$ to be cyclic of order d, so $A(L)$ again (more-or-less) decomposes as a sum $\oplus_\chi A^\chi(K)$, where the A^χ are the L/K twists of A by the characters of $G(L/K)$; see Section 3.3.

(d) Let K be a number field or function field. What, if anything, can one say about the function

$$A(L) \longrightarrow \mathbb{R}, \quad P \longmapsto \hat{h}_{A,D}\bigl(\mathrm{Trace}_{A,L/K}(P)\bigr)? \qquad \text{(A.1)}$$

Note that $\hat{h}_{A,D}\bigl(\sigma(P)\bigr) = \hat{h}_{A,D}(P)$ for all $P \in A(L)$ and all $\sigma \in G(L/K)$. So for example, if D is anti-symmetric, then (A.1) is just the map $P \to [L : K]\hat{h}_{A,D}(P)$. But (A.1) seems more interested if D is ample and symmetric. It might also be interesting to consider the map

$$A(L) \longrightarrow \mathbb{R}, \quad P \longmapsto \det\Bigl(\bigl\langle \sigma(P), \tau(P)\bigr\rangle_{A,D}\Bigr)_{\sigma,\tau\in G(L/K)}. \qquad \text{(A.2)}$$

In particular, the map (A.2) sends P to 0 if and only if the set of Galois conjugates $\{\sigma(P) : \sigma \in G(L/K)\}$ is $\mathbb{Z}$-linearly dependent.

References. There has been work done on local/global trace problems for elliptic curves. In particular, Çiperiani–Ozman [12] includes a detailed analysis of the local/global trace problem for the case $g = 1$ and $[L : K] = 2$, i.e., for elliptic curves and quadratic extensions.

Project II: Northcott and Bogomolov Fields for Abelian Varieties.

For this project, we set the following notation (unless we say otherwise):

$K/\mathbb{Q}$ a number field (although you might consider function fields, too).

A/K an abelian variety of dimension $g \geq 1$.

D $\in \mathrm{Div}_K(A)$, an ample symmetric divisor.

$\hat{h}_{A,D}$ the absolute logarithmic canonical height on $A(\bar{K})$.

Northcott's theorem (Theorem 4.8(d), see also Remark (4.14)) says that $\mathbb{P}^N(K)$ has only finitely many points of bounded height. This implies that there is a constant $c(K) > 0$ so that for all $\alpha \in K$, either $h(\alpha) = 0$ or $h(\alpha) \geq c(K)$. These properties clearly do not hold if we replace K with $\bar{\mathbb{Q}}$, which raises an interesting question: For which infinite algebraic extensions $L/\mathbb{Q}$ are these properties true. This prompts the following two definitions, which we formulate for abelian varieties, but which have been even more widely studied for the Weil height on a field.

Definition A.1. Let $K/\mathbb{Q}$ be a number field, let A/K be an abelian variety, and let $D \in \mathrm{Div}_K(A)$ be an ample symmetric divisor. Let L/K be an algebraic extension (generally of infinite degree).

We say that L has the *Northcott Property for A/K* if for all B, the set

$$\{P \in A(L) : \hat{h}_{A,D}(P) \leq B\} \quad \text{is finite.}$$

We say that L has the *Bogomolov Property for A/K* if there exists a constant $c(A, D, L)$ such that every $P \in A(L)$ satisfies either $\hat{h}_{A,D}(P) = 0$ or $\hat{h}_{A,D}(P) \geq c(A, D, L)$. Equivalently, if

$$\inf\{\hat{h}_{A,D}(P) : P \in A(L) \smallsetminus A_{\mathrm{tors}}(L)\} > 0,$$

where we conventionally set the infimum of the empty set equal to 1.

In this project you'll investigate fields that have the Northcott and Bogomolov Properties. (As a warm-up, prove that the Northcott property implies the Bogomolov property.) In particular, we will look at abelian variety analogues of some of the results that have been proven about fields having the Northcott and Bogomolov Properties; see below for a number of references. But as always, you are encouraged to formulate questions of your own.

(a) Habegger [34] proves that if $E/\mathbb{Q}$ is an elliptic curve, then the field $\mathbb{Q}(E_{\text{tors}})$ generated by the torsion points of E has the Bogomolov property for E, i.e., there exists $\epsilon > 0$ such that then $\hat{h}_{E,(0)}(P) \ge \epsilon$ for all non-torsion points $P \in E\bigl(\mathbb{Q}(E_{\text{tors}})\bigr)$. [Here $\hat{h}_{E,(0)}$ is the height relative to the divisor $(0) \in \operatorname{Div}(E)$.]
   (i) Let $E/\mathbb{Q}$ be an elliptic curve. Does $\mathbb{Q}(E_{\text{tors}})$ have the Bogomolov property for the Weil height h?
   (ii) Let $E/\mathbb{Q}$ and $E'/\mathbb{Q}$ be non-isogenous elliptic curves. Does $\mathbb{Q}(E_{\text{tors}})$ have the Bogomolov property for E'?
   (iii) Consider analogous questions for higher dimensional abelian varieties.

(b) Investigate the following conjecture:

Conjecture A.2 (Remond [77]). *For a subgroup $\Gamma \subseteq A(\bar{K})$, let*

$$\operatorname{rank}(\Gamma) = \dim_{\mathbb{Q}} \Gamma \otimes \mathbb{Q},$$
$$\Gamma_{\text{div}} = \bigl\{P \in A(\bar{K}) : nP \in \Gamma \text{ for some } n \ge 1\bigr\}.$$

Then for all $\Gamma \subseteq A(\bar{K})$ of finite rank, there exists a constant $c(\Gamma) > 0$ such that

$$\hat{h}_{A,D}(P) \ge c(\Gamma) \quad \text{for all } P \in A\bigl(K(\Gamma)\bigr) \smallsetminus \Gamma_{\text{div}}.$$

Note that if we take $\Gamma = \{0\}$, then $\Gamma_{\text{div}} = A(\bar{K})_{\text{tors}}$, so if we also take $g = 1$, this is proven in [34].

(c) There are many results in the literature that give Lehmer-like lower bounds for $\hat{h}_{A,D}(P)$ under the assumption that the field of definition $K(P)$ of P has some special property. For example, there is the very strong bound (8.9) under the assumption that $K(P)/K$ is an abelian extension.

Here's another example. It is a (not yet proven?) abelian variety analogue of a theorem that Amoroso–Masser [2, 3] proved for $\bar{\mathbb{Q}}^*$. Let $K/\mathbb{Q}$ be a number field, let A/K be an abelian variety, let $D \in \operatorname{Div}_K(A)$ be an ample symmetric divisor, and let $\epsilon > 0$. Then there is a constant $c(A/K, D, \epsilon)$ such that if $P \in A(\bar{K})$ is a non-torsion point such that $K(P)/K$ is Galois extension, then

$$\hat{h}_{A,D}(P) \ge \frac{c(A/K, D, \epsilon)}{\bigl[K(P) : K\bigr]^{\epsilon}}. \tag{A.3}$$

Notice that the lower bound (A.3) is stronger than in Lehmer's conjecture (Conjecture 8.7), but weaker than (8.9).

There are also many results in the literature, especially for $\bar{\mathbb{Q}}^*$, but also for abelian varieties, that give Lehmer-type lower bounds under a ramification condition on $K(P)/\mathbb{Q}$. For example, there are papers that assume that $K(P)$ is totally real, or that $K(P)$ is totally p-adic for some prime p; see for example [7, 14, 27, 40]

This project would involve looking at some of these known results to learn about methods used, and also to think about whether there are other sorts of fields for which one might be able to prove a Lehmer-type result.

(d) In the context of Definition A.1, we might say that L/K has the *Universal Northcott Property* if it has the Northcott Property for every abelian variety A/K, and similarly for the *Universal Bogomolov Property*. For example, [4] says that K^{ab} is universally Bogomolov for abelian varieties. Are there other universally Bogomolov fields satisfying $[L:K]=\infty$? What can one say about them? Are there any universally Northcott fields satisfying $[L:K]=\infty$?

References. The following papers have Northcott and Bogomolov results for fields and various sorts of group varieties. It is not meant to be comprehensive, but looking at any of these articles will likely suggest abelian variety problems to work on!

- [1] Francesco Amoroso, Sinnou David, and Umberto Zannier, On fields with Property (B), Proc. Amer. Math. Soc. 142 (2014), no. 6, 1893–1910.
- [7] Bombieri, Enrico; Zannier, Umberto, A note on heights in certain infinite extensions of $\mathbb{Q}$, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl.12 (2001), 5–14.
- [13] Checcoli, S.; Fehm, A. On the Northcott property and local degrees. Proc. Amer. Math. Soc. 149 (2021), no. 6, 2403-–2414.
- [15] Checcoli, Sara; Widmer, Martin. On the Northcott property and other properties related to polynomial mappings. Math. Proc. Cambridge Philos. Soc. 155 (2013), no. 1, 1—12.
- [28] Roberto Dvornicich and Umberto Zannier, On the properties of Northcott and of Narkiewicz for fields of algebraic numbers. part 1, Funct. Approx. Comment. Math. 39 (2008), no. part 1, 163–173.
- [31] Arno Fehm, Three counterexamples concerning the Northcott property of fields, Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl. 29 (2018), no. 2, 309–314.
- [34] P. Habegger, Small height and infinite nonabelian extensions, Duke Math. J. 162 (2013), no. 11, 2027–2076.

- [62] Mello, Jorge; Sha, Min. On the properties of Northcott and Narkiewicz for elliptic curves. Int. J. Number Theory 18 (2022), no. 10, 2129—2144.
- [91] Martin Widmer, On certain infinite extensions of the rationals with Northcott property, Monatsh. Math. 162 (2011), no. 3, 341–353.

Project III: Experimental Investigation of Lehmer's Conjecture for Elliptic Curves.

There is a lot of data available for the classical Lehmer conjecture for $\bar{\mathbb{Q}}^*$; see for example [64]. And there is a conjectural value for the non-root of unity $\alpha \in \bar{\mathbb{Q}}^*$ that minimizes $d(\alpha)h(\alpha)$; see Remark 8.3. As far as I am aware, there are no comparable computations for any elliptic curve. In this project we'll choose a convenient elliptic curve $E/\mathbb{Q}$ and search for non-torsion points $P \in E(\bar{\mathbb{Q}})$ of small degree such that $d(P)\hat{h}_E(P)$ is small.

(1) Choose an elliptic curve, for example, one of

$$\begin{array}{lll} E_1 : y^2 = x^3 - x & \Delta(E) = 2^6 & j(E) = 2^6 \cdot 3^3 \\ E_2 : y^2 + y = x^3 & \Delta(E) = -3^3 & j(E) = 0 \\ E_3 : y^2 + y = x^3 - x^2 & \Delta(E) = -11 & j(E) = -2^{12} \cdot 11^{-1} \\ E_4 : y^2 + y = x^3 - x & \Delta(E) = 37 & j(E) = 2^{12} \cdot 3^3 \cdot 37^{-1} \end{array}$$

For E_1 and E_2, it may be best to go to a field where they have everywhere good reduction and change the Weierstrass equation. For E_3 and E_4, they're already semi-stable, so their equations are already minimal.

(2) Choose a number field, e.g., $K = \mathbb{Q}(\sqrt{7})$ with $R_K = \mathbb{Z}[\sqrt{7}]$, or $K = \mathbb{Q}(\sqrt{-3})$ with $R_K = \mathbb{Z}[(1 + \sqrt{-3})/2]$. Choose some $x \in R_K$ and solve for y in a quadratic extension L of K to get a point $P \in E(L)$. (Or search for a point in $E(K)$.)

(3) If P reduces to the singular point at some prime of bad reduction, replace it by nP for some small n so that this doesn't happen.

(4) For each real and complex embedding of L, compute the corresponding local height of nP. (Note that these local heights may be computed analytically to high accuracy without using information about the ring of integers of L.) Add these to $\frac{1}{2}$ the norm of the denominator of $x(nP)$ to get the canonical height of nP. Then divide by n^2 to get $\hat{h}_E(P)$.

(5) Repeat for a large number of x values and try to find points with small canonical height.

Side Project: It might be useful to develop an explicit estimate of the following form: For $E/\mathbb{Q}$ and $P \in E(\bar{\mathbb{Q}})$, let

$$K = \mathbb{Q}(P) \quad \text{and} \quad d = [K : \mathbb{Q}].$$

Then there is a lower bound

$$\hat{h}_E(P) \ge C_1(d) \cdot \log\bigl|\mathrm{Disc}(K/\mathbb{Q})\bigr| - C_2(d, E).$$

This can be proven by combining an estimate for $\hat{h}_E(P) - h\bigl(x(P)\bigr)$ (there are many articles giving such bounds) and a lower bound for the Weil height of a point in terms of the discriminant of the field generated by its coordinates (see for example [82, Theorem 2]). If we find explicit values for $C_1(d)$ and $C_2(d, E)$ for our curve E, then we might be able to determine, for example, the absolutely smallest possible height over all (say) quadratic or cubic fields.

Mathematics Department, Box 1917, 151 Thayer Street, Brown University, Providence, RI 02912 USA
Email address: joseph_silverman@brown.edu