A LEHMER-TYPE HEIGHT LOWER BOUND FOR
ABELIAN SURFACES OVER FUNCTION FIELDS

NICOLE R. LOOPER AND JOSEPH H. SILVERMAN

Abstract. Let $K$ be a 1-dimensional function field over an algebraically closed field of characteristic 0, and let $A/K$ be an abelian surface. Under mild assumptions, we prove a Lehmer-type lower bound for points in $A(\bar{K})$. More precisely, we prove that there are constants $C_1, C_2 > 0$ such that the normalized Bernoulli-part of the canonical height is bounded below by

$$\hat{h}_A^B(P) \geq C_1 [K(P) : K]^{-2}$$

for all points $P \in A(\bar{K})$ whose height satisfies $0 < \hat{h}_A(P) \leq C_2$.

Contents

1. Introduction 2
2. Survey of Previous Results and Methods 6
3. An Overview of Canonical Local and Global Heights 9
4. Local Heights for Completely Split Multiplicative Reduction 12
5. A Hexagonal Fourier Calculation 15
6. Averaging (Periodic) Functions over (Torsion) Points 24
7. Two Lower Bounds for the Local Height 28
  7.1. A Local Height Lower Bound via Fourier Averaging 28
  7.2. A Local Height Lower Bound via the Pigeonhole Principle 31
8. A Bound for Small Differences Lying on $\Theta$ 33
9. A Lehmer-Type Height Bound for Abelian Surfaces 35
References 45
Appendix A. Verification of some basic formulas 46

Date: August 24, 2021.
2010 Mathematics Subject Classification. Primary: 11G10, 11G50, 14K15; Secondary: 14K25, 37P30.
Key words and phrases. abelian variety, canonical height, Lehmer conjecture.
The first author was supported by NSF grant DMS-1803021. The second author was partially supported by Simons Collaboration Grant #712332.
1. Introduction

The classical Lehmer conjecture says that there is an absolute constant \( C > 0 \) such that if \( \alpha \in \mathbb{Q}^* \) is not a root of unity, then its absolute logarithmic height satisfies
\[
 h(\alpha) \geq \frac{C}{[\mathbb{Q}(\alpha) : \mathbb{Q}]}.
\]

Various authors have extended this conjecture to elliptic curves and to higher dimensional abelian varieties. We review these conjectures and some of the progress made in proving them in Section 2.

The main results of the present paper are: (1) an explicit Fourier expansion of the “Bernoulli-part” of the canonical height on abelian surfaces defined over non-archimedean local fields (Theorem 5.2); (2) a lower bound for a torsion-and-difference average of the Bernoulli part of the height on abelian surfaces defined over function fields (Theorem 9.1). This is an analogue of the key lemma in [16], which dealt with elliptic curves. We also prove: (3) a Lehmer-type lower bound with exponent 2 for the canonical height of non-torsion points on abelian surfaces defined over function fields that is conditional on the assumption that the torsion-and-difference average of the “intersection part” of the canonical height is at least as large as a certain local-global constant (Corollary 9.2).

Before explaining the statements of our results in more detail, we briefly recall the local decomposition of the canonical height on abelian varieties. See Section 3 for further details and references.

Let \( k \) be an algebraically closed field of characteristic 0, let \( K/k \) be a 1-dimensional function field, let \( A/K \) be an abelian variety, and let \( \Delta \in \text{Div}(A) \) be an ample symmetric divisor on \( A \). The associated canonical height
\[
 \hat{h}_{A,\Delta} : A(\bar{K}) \to \mathbb{R}_{\geq 0}
\]
may be decomposed as a sum of normalized local canonical heights
\[
 \hat{\lambda}_{A,\Delta,v} : (A \setminus |\Delta|)(\bar{K}_v) \to \mathbb{R},
\]
one for each absolute value on \( \bar{K} \), where the normalization condition
\[
 \lim_{N \to \infty} \frac{1}{N^{2g}} \sum_{P \in A[N] \setminus |\Delta|} \hat{\lambda}_{A,\Delta,v}(P) = 0 \tag{1}
\]
serves to uniquely determine \( \hat{\lambda}_{A,\Delta,v} \). The local height further decomposes into an “intersection-part” and a “Bernoulli-part,” which we denote respectively by \( \hat{\lambda}^I_{A,\Delta,v} \) and \( \hat{\lambda}^B_{A,\Delta,v} \). The intersection part is given
by
\[ \hat{\lambda}_{A,\Delta,v}^I(P) = \left( \text{intersection index of } P \text{ and } \Delta \text{ on the } v\text{-fiber of the Néron model of } A \right) - \kappa_{I,A,\Delta,v}, \]
where the constant \( \kappa_{I,A,\Delta,v} \) is chosen so that \( \hat{\lambda}_{I,A,\Delta,v} \) itself satisfies the normalization condition (1), and then the Bernoulli part is what’s left over, i.e.,
\[ \hat{\lambda}_{A,\Delta,v}(P) = \hat{\lambda}_{A,\Delta,v}^I(P) + \hat{\lambda}_{A,\Delta,v}^B(P) \]
We also note that
\[ A \text{ has potential good reduction at } v \implies \hat{\lambda}_{A,\Delta,v}^B = \kappa_{I,A,\Delta,v} = 0. \]
(See Section 3 for the further details.)

For any finite extension \( L/K \), we write \( M_L \) for an appropriate normalized set of absolute values on \( L \), and we let
\[ M_{L}^{\text{bad}}(A) = \{ v \in M_L : A \text{ has bad reduction at } v \}. \]
Then for \( P \in A(L) \setminus |\Delta| \), we define an “intersection-part” and a “Bernoulli-part” of the global canonical height via
\[ \hat{h}_{A,\Delta}^B(P) = \frac{1}{[L : K]} \sum_{v \in M_{L}^{\text{bad}}(A)} \hat{\lambda}_{A,\Delta,v}^B(P), \]
\[ \hat{h}_{A,\Delta}^I(P) = \frac{1}{[L : K]} \sum_{v \in M_{L}} \hat{\lambda}_{A,\Delta,v}^I(P). \]
With this notation, there exists a local-global height constant \( \kappa_{A,\Delta} \) so that
\[ \hat{h}_{A,\Delta}(P) = \hat{h}_{A,\Delta}^B(P) + \hat{h}_{A,\Delta}^I(P) - \kappa_{A,\Delta} \text{ for all } P \in A(K) \setminus |\Delta|. \] (2)
We note that if \( \dim(A) = 1 \), i.e., if \( A \) is an elliptic curve, then \( \kappa_{A,\Delta} = 0. \) However, if \( \dim(A) \geq 2 \), then \( \kappa_{A,\Delta} \) is generally positive.

Our main results are a Fourier series calculation and the following lower bound for the Bernoulli part of the canonical height in the case that \( A \) is an abelian surface defined over a function field. This theorem, and the Fourier averaging lemmas that we prove along the way, are analogues of the key lemmas and results in [16], where similar results are proven for elliptic curves. However, we note that the main theorem in [16] is an unconditional Lehmer-type lower bound for the canonical height on elliptic curves (with non-integral \( j \)-invariant), while our result for abelian surfaces only gives a lower bound for a suitable average of the Bernoulli part of the height.

**Theorem 1.1** (Theorem 9.1 and Corollary 9.2). Fix the following quantities:
Let $k$ be an algebraically closed field of characteristic 0. Let $K/k$ be a 1-dimensional function field. Let $(A, \Theta)/K$ be an abelian variety $A$ defined over $K$ with an effective symmetric principal polarization $\Theta \in \text{Div}_K(A)$.

Let $\hat{h}_{A, \Theta}$ be the canonical height on $A$ for the divisor $\Theta$. Let $\hat{h}^B_{A, \Theta}, \hat{h}^I_{A, \Theta}$ be the Bernoulli and intersection parts of the canonical height on $A$ for the divisor $\Theta$.

Assume that for every place $v$ of $K$, the abelian variety $A$ has either potential good reduction at $v$ or totally multiplicative reduction at $v$, and that $A$ has at least one place of multiplicative reduction.

(a) There are constants $C_1, C_2, C_3, C_4 > 0$ and an integer $d \geq 1$ so that for all finite extensions $L/K$ and all sets of points $\Sigma \subset A(L)$ of $\hat{h}_{A, \Theta}$-height at most $C_1$, there is a subset $\Sigma_0 \subseteq \Sigma$ with $\#\Sigma_0 \geq C_2 \#\Sigma$ so that the following double average$^1$ of the Bernoulli part of the heights of the points in $\Sigma_0$ satisfies

$$\text{Avg}_{P, Q \in \Sigma_0} \text{Avg}_{T \in A[d]} \hat{h}^B_{A, \Theta}(P - Q + T) \geq \frac{C_3}{[L : K]^{2/3}} - \frac{C_4}{\#\Sigma}.$$  

(b) Suppose that the subset $\Sigma_0$ in (a) can always be chosen so that it satisfies the further estimate

$$\text{Avg}_{P, Q \in \Sigma_0} \text{Avg}_{T \in A[d]} \hat{h}^B_{A, \Theta}(P - Q + T) \geq \kappa_{\Theta},$$

where $\kappa_{\Theta}$ is the constant appearing in (2). Then there is a constant $C_5 > 0$ so that every non-torsion $P \in A(\overline{K})$ satisfies

$$\hat{h}_{A, \Theta}(P) \geq \frac{C_5}{[K(P) : K]^2}.$$  

We make several remarks, but again we refer the reader to Section 2 for more details of the history and known results surrounding Lehmer’s conjecture. To ease notation, we write

$$D = [K(P) : K]$$

for the degree of the field of definition of $P$ (although we note that later we assign a different meaning to $D$).

Remark 1.2. Our proof uses the Fourier averaging technique that has previously been used for the classical Lehmer conjecture $^5$ Blanksby–Montgomery (1971) and for Lehmer’s conjecture on elliptic curves $^6$.

---

$^1$The averaging notation $\mathcal{A}_\alpha$ is fairly self-explanatory, but see Section 6 for the precise definition.
Hindry–Silverman (1990)]. A crucial ingredient in the one-dimensional cases is that the Fourier series associated to the local heights has non-negative coefficients, a fact that is no long true in the higher dimensional case. Thus our proof has two key components. First we compute the relevant two-dimensional Fourier series attached to a periodic two-variable quadratic form with its associated hexagonal fundamental domain. Second we deal with the issue that the Fourier series has both positive and negative Fourier coefficients via a subsidiary averaging process over a suitable collection of torsion points.

**Remark 1.3.** Currently the best known result for general abelian surfaces over number fields\(^2\) is due to Masser. More generally it is proven in [19, Masser 1984] that on an abelian variety of dimension \(g\), there is a Lehmer estimate

\[
\hat{h}_A(P) \geq \frac{C_\varepsilon(A/K)}{D^{2g+6+2/g+\varepsilon}}.
\]

Thus for abelian surfaces, i.e., for \(g = 2\), Masser’s lower bound is \(O(D^{-11})\), which may be compared to our conditional lower bound of \(O(D^{-2})\) and with the conjectural lower bound of \(O(D^{-1/2})\). Masser’s proof uses auxiliary polynomials and methods from Diophantine approximation, a technique that has been long used in studying Lehmer’s conjecture.

**Remark 1.4.** One would of course like to prove a result for number fields that is analogous to Theorem 1.1, much as was done in [16] for elliptic curves. However, the Fourier expansion for the archimedean local height is likely to include negative Fourier coefficients, just as in the non-archimedean case. And these negative Fourier coefficients would vitiate the argument used in the present paper, since we rely on the fact that our absolute values are discrete, and thus that the component groups on the Néron model are finite and are well-behaved under finite extension. For archimedean absolute values, the “fiber” on the “Néron model” should be viewed as having “bad reduction” with “component group” equal to a real torus. We thus have no consistent way to calculate which multiples of a point lie on (or near) the “identity component” in an archimedean topology.

**Remark 1.5.** Fourier averaging techniques have also been used successfully for studying Lang’s height lower bound conjecture, in which one fixes a field \(K\) and varies the abelian variety. Lang’s conjecture asserts (roughly) that for all abelian varieties of \(A/K\) of dimension \(g\)

\(2\)Presumably Masser’s proof carries over to the function field setting, where the \(\varepsilon\) might well be superfluous.
and all points \( P \in A(K) \) whose multiples are Zariski dense in \( A \), we have
\[
\hat{h}^*_A(P) \geq c_1(K, g)h(A/K) - c_2(K, g),
\]
where \( h(A/K) \) is an appropriate height of the abelian variety. For \( g = 1 \), this was proven for function fields, and conditionally for number fields assuming on Szpiro’s conjecture, using Fourier averaging [16, Hindry–Silverman]. The difficulty in directly extending these proofs to abelian surfaces, even conditionally on a Szpiro-type conjecture, is again the two-fold problem of negative Fourier coefficients and that pesky \( \kappa_{A, A} \) constant. However, see [7, David (1993)] and [21, Pazuki (2013)] for Lang-style lower bounds for abelian varieties in which the lower bound has a correction term that measures the distance to the boundary of moduli space.

**Remark 1.6.** We also hope that it may be possible to extend our results to abelian varieties of dimension three or greater. However, it seems a challenging problem to write down an explicit formula for the Fourier series of the Bernoulli part of the local height in higher dimensions, since our two-dimensional hexagonal fundamental domain (see Figure 1) would be replaced by a \( g \)-dimensional parallelepiped. However, if this could be done, we would not be surprised if it could be used to prove a function field Lehmer-type bound with exponent 2 for the Bernoulli-part of the height.

## 2. Survey of Previous Results and Methods

In this section we give a brief overview of the study of Lehmer-type height lower bounds. We continue with the notation
\[
D = [K(P) : K].
\]
Lehmer’s original conjecture [18, Lehmer (1933)], actually phrased as a question, says that
\[
h(\alpha) \geq CD^{-1}.
\]
General Lehmer-type estimates were proven in the classical case in [5, Blanksby–Montgomery 1971] using Fourier series methods and in [20,}

---

3 We assume throughout this historical survey (Section 2) that “trivial counterexamples” are excluded. Thus for Lehmer’s original conjecture, we assume that \( \alpha \) is non-zero and not a root of unity, for elliptic curves \( P \) is a non-torsion point, and for abelian varieties we assume that the iterates of \( P \) are Zariski dense.

4 We use the phrase “Lehmer-type estimate” to mean a height lower bound that decays at worst polynomially in the degree \( D \). We note that it is relatively easy to obtain exponentially decaying bounds.
Stewart 1978] using auxiliary polynomials. Both proofs give bounds of the form

\[ h(\alpha) \geq CD^{-2}(\log D)^{-1}. \]

Stewart’s methods were applied in [12, Dobrowolski 1979] to achieve the following bound, which is ever-so-close to Lehmer’s conjecture:

\[ h(\alpha) \geq CD^{-1} \left( \frac{\log \log D}{\log D} \right)^3. \]

Dobrowolski’s innovation was to use the Frobenius \( p \)-power map for suitably many \( p \) to greatly increase the power of the vanishing lemma.

The first general Lehmer-type estimate for elliptic curves was given in [2, Anderson–Masser (1980)], where a lower bound of roughly \( D^{-10} \) was given. This was subsequently improved in [20, Masser (1989)] to

\[ \hat{h}_E(P) \geq CD^{-3}(\log D)^{-2}. \]

Masser’s proof uses auxiliary polynomials. A Fourier series proof of the same precision was given in [33, Zhang (1989)].

Stronger results are known for restricted collections of elliptic curves. Notable is the result [17, Laurent (1983)], who proves a Dobrowolski-type bound

\[ \hat{h}_E(P) \geq CD^{-1} \left( \frac{\log \log D}{\log D} \right)^3 \text{ if } E \text{ has complex multiplication.} \]

And building on Zhang’s ideas, Masser’s result was improved in [16, Hindry–Silverman (1990)] to

\[ \hat{h}_E(P) \geq CD^{-2}(\log D)^{-2} \text{ if } j(E) \text{ is non-integral.} \]

There are also many results proving Lehmer-type estimates for points defined over restricted types of fields. One of the earliest such results is the proof [25, Smyth (1971)] that Lehmer’s conjecture is true for all \( \alpha \in \bar{\mathbb{Q}}^* \) such that \( \alpha^{-1} \) is not a \( \bar{\mathbb{Q}}/\mathbb{Q} \)-Galois conjugate of \( \alpha \). (One says that \( \alpha \) is non-reciprocal.) Even stronger results are known for points defined over abelian extensions of the ground field \( K \). It is shown in [11, Amoroso–Dvornicich(2000)] that

\[ h(\alpha) \geq C(K) > 0 \text{ for all non-zero non-roots of unity } \alpha \in K^{ab}, \]

and analogous estimates for points defined over \( K^{ab} \) were proven for elliptic curves in [3, Baker (2003)] and [24, Silverman (2004)], and then for abelian varieties in [4, Baker–Silverman (2004)]. Note that in these abelian extension results, the lower bounds are independent of \( D \). Under the weaker assumption that \( K(P)/K \) is a Galois extension,
it is shown in [13, Galateau–Mahé (2017)] that the elliptic Lehmer conjecture is true.

We next consider higher dimensional abelian varieties. For a simple abelian variety \( A/K \) of dimension \( g \) and appropriate choice of canonical height, the current conjecture \([8, 19]\) appears to be
\[
\hat{h}_A(P) \geq CD^{-1/g},
\]
although no one has yet even managed to get even \( D^{-1} \). It is shown in \([19, \text{Masser (1984)}]\) that
\[
\hat{h}_A(P) \geq C\epsilon D^{-2g-6-2/g-\epsilon},
\]
and if \( A \) has complex multiplication, then Dobrowolski-type bounds have been proven in \([8, \text{David–Hindry (2000)}]\) and \([22, \text{Ratazzi (2008)}]\).

For the \( g \)-fold product \( E^g \) of an elliptic curve, the estimate
\[
\hat{h}_{E^g}(P) \geq CD^{-1-1/2g} (\log D)^{-2/g}
\]
is proven in \([13, \text{Galateau–Mahé (2017)}]\).

A Lehmer-type conjecture involves fixing one geometric object such as \( G_m, E, \) or \( A \) defined over a field \( K \), and finding height lower bounds for points defined over extensions of \( K \). Dem’janenko and Lang conjectured a different sort of height lower bound for elliptic curves by fixing a field \( K \) and allowing the elliptic curve to vary. The original conjecture had the form
\[
\hat{h}_E(P) \geq c_1(K) \log N_{D_{E/K}} - c_2(K)
\]
for all \( E/K \) and all non-torsion \( P \in E(K) \), and this has been generalized to abelian varieties with the log-discriminant replaced by an appropriate height of the abelian variety, e.g., the height \( h(A/K) \) used by Faltings in his proof of the Mordell conjecture. A Fourier series argument was used in \([15, \text{Hindry–Silverman (1988)}]\) to prove that Lang’s conjecture for elliptic curves is a consequence of Szpiro’s conjectured inequality relating the discriminant and the conductor of an elliptic curve, so in particular Lang’s conjecture is a theorem over one-dimensional characteristic 0 function fields. However, for higher dimensional abelian varieties, the best known estimates include an error term that grows as the moduli point of the abelian variety approaches the boundary of the associated moduli space; see for example \([7, \text{David (1993)}]\).

The definition of the canonical height of points on abelian varieties can be extended to assign a canonical height to subvarieties of higher dimension, and one can formulate a Lehmer conjecture and
prove Lehmer-type lower bounds for these higher dimensional heights. See for example the series of papers by David and Philippon [9, 10, 11].

In this brief section, we have only touched on some of the work done on Lehmer’s conjecture. For additional information, the reader might consult the lengthy (unpublished) survey article [28, Verger-Gaugry (2019)] that includes an extensive bibliography of articles related to the conjectures of Lehmer and Schinzel-Zassenhaus.

3. An Overview of Canonical Local and Global Heights

We follow the exposition in Hindry’s notes [14]; see also the articles [29, 30, 31] by Werner (especially [30]). We set the following notation:

- $k$: an algebraically closed field $k$ of characteristic 0.
- $K/k$: a 1-dimensional function field over $k$.
- $M_L$: for finite extensions $L/K$, a complete set of absolute values on $L$, normalized so that $w(L^*) = \mathbb{Z}$ for all $w \in M_L$.
- $A/K$: an abelian variety of dimension $g$ defined over $K$.
- $A$: the Néron model of $A/K$.
- $A^0$: the identity component of the Néron model of $A/K$.

For $P = [x_0, \ldots, x_N] \in \mathbb{P}^N(\bar{K})$, the Weil height of $P$ is defined by choosing a finite extension $L/K$ with $P \in \mathbb{P}^N(L)$ and setting

$$h(P) = \frac{1}{[L : K]} \sum_{w \in M_L} \max \{-w(x_i)\}.$$  

The value is independent of the choice of $L$.

**Theorem 3.1.** (Néron) Let $A/K$ be an abelian variety. There exists is a unique collection of functions

$$\hat{\lambda}_{\Delta,v} : A(K_v) \setminus |\Delta| \to \mathbb{R},$$  

where $\Delta \in \text{Div}_K(A)$ and $v \in M_K$, so that the following are true:

(a) The map $\hat{\lambda}_{\Delta,v}$ is continuous, where we give $A(K_v)$ the $v$-adic topology.

---

5For comparison with [14], we note that Hindry’s $i_v(D, P)$ is our $(D \cdot P)_{A,v}$, and we have adopted his $B_{D,v}(j_v(P))$ notation; cf. [14] (3.10) and (3.11). We also point the reader to the brief discussion of the function field setting in [14, Section 5].

6Those who are familiar with the theory of Weil heights may wonder where the local factor $[L_w : K_w]$ has gone. The answer is that there is no residue degree, since our scalar field $k$ is algebraically closed, and the ramification degree is already absorbed in the way that we have normalized the absolute values in $M_K$ and $M_L$, i.e., if $\alpha \in K^*$ and $w \in M_L$ lies over $v \in M_K$, then $w(\alpha) = e(w/v)v(\alpha)$ already includes the ramification degree.
(b) For all $\Delta, \Delta' \in \text{Div}_K(A)$ and all $v \in M_K$,
$$\hat{\lambda}_{\Delta + \Delta', v} = \hat{\lambda}_{\Delta, v} + \hat{\lambda}_{\Delta', v} \quad \text{on } A(\overline{K_v}) \setminus (|\Delta| \cup |\Delta'|).$$

(c) For all morphisms $\varphi : A \to B$ of abelian varieties over $K$ and all $\Delta \in \text{Div}_K(B)$,
$$\hat{\lambda}_{A, \varphi^* \Delta, v} = \hat{\lambda}_{B, \Delta, v} \circ \varphi \quad \text{on } A(\overline{K_v}) \setminus |\varphi^* \Delta|.$$ 

(d) For all rational functions $f \in K(A)$,
$$\hat{\lambda}_{\text{div}(f), v} = v \circ f \quad \text{on } A(\overline{K_v}) \setminus |\text{div}(f)|.$$ 

(e) (Normalization) For all $\Delta \in \text{Div}_K(A)$ and all $v \in M_K$, we have
$$\lim_{N \to \infty} N^{-2g} \sum_{P \in A[N] \setminus |\Delta|} \hat{\lambda}_{\Delta, v}(P) = 0. \tag{5}$$

(f) (Good Reduction) If $A$ has potential good reduction at $v$, then
$$\hat{\lambda}_{\Delta, v}(P) = \langle \Delta \cdot P \rangle_{A, v}$$
is the intersection index over $v$ of the closures of $\Delta$ and $P$ in $A$. In the case of potential good reduction we have
$$\hat{\lambda}_{\Delta, v}(P) \geq 0 \quad \text{for all } P \in A(\overline{K_v}) \setminus |\Delta|.$$ 

(g) (Bad Reduction) Let
$$j_v : A(K) \longrightarrow (A/A^0)_v(k)$$
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
$$B_{\Delta, v} : (A/A^0)_v(k) \longrightarrow \mathbb{R}$$
so that
$$\hat{\lambda}_{\Delta, v}(P) = \langle \Delta \cdot P \rangle_{A, v} + B_{\Delta, v}(j_v(P)) - \kappa_{\Delta, v}, \tag{6}$$
where again $\kappa_{\Delta, v}$ is chosen so that (5) holds.

---

7Without this normalization, which Néron did not impose in his original formulation, the function $\lambda_{\Delta, v}$ is only well-defined up to an $M_K$-constant. We also mention that if the absolute value on $K$ is archimedean, then the normalization condition is equivalent to \( \int_{A(\overline{K_v})} \hat{\lambda}_{\Delta, v}(P) \, d\mu(P) = 0 \), where $\mu$ is Haar measure on $A(\overline{K_v}) \cong A(\mathbb{C})$.

8Néron proved that this formula is true up to a constant. See [6] for a proof that the average of the intersection multiplicities over torsion points goes to 0, which implies that the constant vanishes.

9Néron further proved that the values of $B_{\Delta, v}$ are rational numbers with denominators dividing $2\#(A/A^0)_v(k)$.
(h) (Local-Global Decomposition) There is a constant $\kappa_\Delta$ so that for all finite extensions $L/K$ and all $P \in A(L) \setminus |\Delta|$, \[
\hat{h}_\Delta(P) = \frac{1}{[L : K]} \sum_{w \in M_K} \hat{\lambda}_{\Delta,w}(P) - \kappa_\Delta.\]

**Definition 3.2.** With notation as in Theorem 3.1, we define the normalized intersection local height and the normalized Bernoulli local height to be, respectively,
\[
\hat{\lambda}_{\Delta,v}^I(P) = \langle \Delta : P \rangle_{A,v} - \kappa_{\Delta,v}^I, \quad \text{and} \quad \hat{\lambda}_{\Delta,v}^B(P) = \mathcal{B}_{\Delta,v}(j_v(P)) - \kappa_{\Delta,v}^B, \tag{7}
\]
where the constants $\kappa_{\Delta,v}^I$ and $\kappa_{\Delta,v}^B$ are chosen to ensure the normalization formulas
\[
\frac{1}{N^{2g}} \sum_{T \in A[N]} \hat{\lambda}_{\Delta,v}^I(T) \xrightarrow{N \to \infty} 0 \quad \text{and} \quad \frac{1}{N^{2g}} \sum_{T \in A[N]} \hat{\lambda}_{\Delta,v}^B(T) \xrightarrow{N \to \infty} 0. \tag{8}
\]

We note that Theorem 3.1(f) says that if $A$ has potential good reduction at $v$, then
\[
\hat{\lambda}_{\Delta,v} = \hat{\lambda}_{\Delta,v}^I, \quad \kappa_{\Delta,v}^I = 0, \quad \text{and} \quad \hat{\lambda}_{\Delta,v}^B = 0,
\]
so the added complication of (7) and (8) are only needed if $A$ does not have potential good reduction.

**Definition 3.3.** We define the global intersection height and the global Bernoulli height as follows: For $P \in A(\bar{K}) \setminus |\Delta|$,
\[
\hat{h}_{A,\Delta}(P) = \frac{1}{[K(P) : K]} \sum_{w \in M_K(P)} \hat{\lambda}_{A,\Delta,w}(P),
\]
\[
\hat{h}_{A,\Delta}^B(P) = \frac{1}{[K(P) : K]} \sum_{w \in M_K(P)} \hat{\lambda}_{A,\Delta,w}^B(P).
\]

The next result summarizes how our various normalizations and normalizing constants fit together.

**Proposition 3.4.** With notation as in Theorem 3.1 and Definitions 3.2 and 3.3, we have
\[
\hat{\lambda}_{\Delta,v}(P) = \hat{\lambda}_{\Delta,v}^I(P) + \hat{\lambda}_{\Delta,v}^B(P). \tag{9}
\]
\[
\hat{h}_\Delta(P) = \hat{h}^I_\Delta(P) + \hat{h}_\Delta^B(P) - \kappa_{A,\Delta}. \tag{10}
\]

**Important Note:** When the local heights are normalized via (5), then their weighted sum will generally differ from the glocal height by a non-zero constant that we have denoted $\kappa_{A,\Delta}$; see [14] Appendix for an example. However, if $\dim(A) = 1$, then $\kappa_{A,\Delta} = 0$, which is why this issue does not arise when working with elliptic curves.
Proof. Using (5), (6), (7), and (8), we see that

\[
0 = \lim_{N \to \infty} \frac{1}{N^{2g}} \sum_{T \in \mathcal{A}[N]} \left( \hat{\lambda}_{\Delta,v}(T) - \hat{\lambda}_{\Delta,v}^I(T) - \hat{\lambda}_{\Delta,v}^B(T) \right)
\]

\[
= \lim_{N \to \infty} \frac{1}{N^{2g}} \sum_{T \in \mathcal{A}[N]} \left( -\kappa_{\Delta,v} + \kappa_{\Delta,v}^I + \kappa_{\Delta,v}^B \right)
\]

Thus \( \kappa_{\Delta,v} = \kappa_{\Delta,v}^I + \kappa_{\Delta,v}^B \), which gives (9). Then

\[
\hat{h}_\Delta(P) + \kappa_{A,\Delta} = \frac{1}{[L : K]} \sum_{w \in M_K} \hat{\lambda}_{\Delta,w}(P) \quad \text{from Theorem 3.1(h),}
\]

\[
= \frac{1}{[L : K]} \sum_{w \in M_K} \left( \hat{\lambda}_{\Delta,w}^I(P) + \hat{\lambda}_{\Delta,w}^B(P) \right) \quad \text{from (9),}
\]

\[
= \hat{h}_\Delta^I(P) + \hat{h}_\Delta^B(P) \quad \text{from Definition 3.3,}
\]

which proves (10). \( \square \)

Remark 3.5. We note that \( j_v \) and Néron’s Bernoulli function \( B_{\Delta,v} \) are defined at all points, so the Bernoulli-part of the local height is well-defined everywhere,

\[
\hat{\lambda}_{\Delta,v}^B : A(\bar{K}) \longrightarrow \mathbb{R}.
\]

This is in contrast to the intersection-part \( \hat{\lambda}_{\Delta,v}^I \) of the local height, which is only defined off of the support of the divisor \( \Delta \), since if \( P \in |\Delta| \), then the local intersection index \( \langle \Delta \cdot P \rangle_{A,v} \) is not defined.

4. Local Heights for Completely Split Multiplicative Reduction

We continue with our discussion of (local) heights based on the material in [14]. For this section, we fix a non-archimedean place \( v \in M_K \) such that \( A_v^o \cong \mathbb{G}_m^g \) is a split torus. There is then a \( v \)-adic uniformization

\[
\mathbb{G}_m^g(K_v)/\Omega \sim A(K_v),
\]

where \( \Omega \) is a (multiplicative) lattice, say spanned by the columns of

\[
\Omega = \text{Multiplicative-Span} \left( \begin{array}{cccc}
q_{11}^2 & q_{12}^2 & \cdots & q_{1g}^2 \\
q_{21}^2 & q_{22}^2 & \cdots & q_{2g}^2 \\
\vdots & \vdots & \ddots & \vdots \\
q_{g1}^2 & q_{g2}^2 & \cdots & q_{gg}^2
\end{array} \right),
\]

\( ^11 \)The \( q_{ij} \) may live in a multi-quadratic extension of \( K \).
We define matrices
\[
q = \begin{pmatrix}
q_{11} & q_{12} & \cdots & q_{1g} \\
q_{21} & q_{22} & \cdots & q_{2g} \\
\vdots & \vdots & \ddots & \vdots \\
q_{g1} & q_{g2} & \cdots & q_{gg}
\end{pmatrix}
\quad \text{and} \quad
Q = \begin{pmatrix}
v(q_{11}) & v(q_{12}) & \cdots & v(q_{1g}) \\
v(q_{21}) & v(q_{22}) & \cdots & v(q_{2g}) \\
\vdots & \vdots & \ddots & \vdots \\
v(q_{g1}) & v(q_{g2}) & \cdots & v(q_{gg})
\end{pmatrix},
\]
where \(q\) and \(Q\) are symmetric, and \(Q\) is positive-definite. In general, when we apply \(v\) to vectors and matrices with entries in \(K_v\), we mean the associated vector or matrix obtained by applying \(v\) to the entries. So for example, we have \(Q = v(q)\), and for \(u \in \mathbb{G}_m(K_v)\), we have \(v(u) = (v(u_1), \ldots, v(u_g))\).

We introduce notation that will make it easier to work with linear algebra on multiplicative spaces. For (column) vectors \(u = (u_1, \ldots, u_g) \in \mathbb{G}_m(K_v)\) and \(m = (m_1, \ldots, m_g) \in \mathbb{Z}^m\), we define\(^\text{12}\)
\[
^t m \ast u = \prod_{i=1}^g u_i^{m_i}.
\]
Similarly, for the multiplicative period matrix \(q\) and integer vectors \(m, n \in \mathbb{Z}^g\), we define
\[
^t m \ast q \ast n = \prod_{i,j=1}^g q_{ij}^{m_i n_j}.
\]
In particular, we note that
\[
v(^t m \ast q \ast n) = ^t m Q n = \sum_{i,j=1}^g m_i n_j v(q_{ij})
\]
is the value of the bilinear form associated to the positive-definite matrix \(Q\).

Just as in the classical case over \(\mathbb{C}\), the change-of-basis formula for the multiplicative period matrix \(q\) may be described using the symplectic group
\[
\text{Sp}_{2g}(\mathbb{Z}) = \left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \text{Mat}_{2g \times 2g}(\mathbb{Z}) : \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \right\}.
\]
\(^\text{12}\)The intuition is that \(^t m \ast u\) is \(\exp(^t m \log u)\).
For our purposes, it suffices to describe the action of $\text{Sp}_{2g}(\mathbb{Z})$ on the period valuation matrix $Q$. It is given by the formula

$$
\begin{pmatrix} A & B \\ C & D \end{pmatrix} \star Q = (AQ + B)(CQ + D)^{-1}.
$$

(11)

The following normalization lemma for the 2-dimensional case will be used later.

**Lemma 4.1.** Let $Q$ be a positive definite symmetric 2-by-2 matrix. Then the $\text{Sp}_4(\mathbb{Z})$ equivalence class of $Q$ via the action (11) contains a matrix

$$
\begin{pmatrix} a & b \\ b & c \end{pmatrix} \in \text{Sp}_4(\mathbb{Z}) \star Q
$$

satisfying

$$
ac > b^2 \quad \text{and} \quad 0 \leq 2b \leq a \leq c.
$$

(12)

We will say that a matrix $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ satisfying (12) is normalized.

**Proof.** Standard reduction theory of positive definite binary quadratic forms (Gaussian reduction) tells us that there is a matrix $A \in \text{SL}_2(\mathbb{Z})$ such that

$$
AQ^tA = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \quad \text{with} \quad 0 \leq |2b| \leq a \leq c.
$$

We note that

$$
AQ^tA = Q \star \begin{pmatrix} A & 0 \\ 0 & t_A^{-1} \end{pmatrix}, \quad \text{where} \quad \begin{pmatrix} A & 0 \\ 0 & t_A^{-1} \end{pmatrix} \in \text{Sp}_4(\mathbb{Z}).
$$

This completes the proof if $b \geq 0$. And if $b < 0$, then we can change the sign of $b$ using the following element of $\text{Sp}_4(\mathbb{Z})$:

$$
\begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \star \begin{pmatrix} a & b \\ b & c \end{pmatrix} = \begin{pmatrix} a & -b \\ -b & c \end{pmatrix}.
$$

This completes the proof of Lemma 4.1. □

**Definition 4.2.** The **theta function** associated to the half-period matrix $q$ is the function

$$
\theta(\cdot, q) : \mathbb{G}_m^g(K_v) \longrightarrow K_v,
$$

$$
\theta(u, q) = \sum_{m \in \mathbb{Z}^g} (t^m \star q \star m)(t^m \star u).
$$
Written out in full, we have

\[ \theta(u, q) = \sum_{m \in \mathbb{Z}} \prod_{i,j=1}^{g} q_{ij}^{m_{ij}} \cdot \prod_{k=1}^{g} u_{k}^{m_{k}}. \]

The positive-definiteness of \( Q = v(q) \) ensures that the sum converges for all \( u \in \tilde{K}_{v}^{*} \).

We next compute the transformation formula for \( \theta \) when \( u \) is translated by an element of \( \Omega \). We observe that an element of \( \Omega \) is a product of powers of the columns of the matrix whose entries are \( q_{ij}^{2} \), so they are elements of \( G_{g}^{m}(K_{v}) \) of the form \( q^{2}n \) with \( n \in \mathbb{Z} \).

**Proposition 4.3.** Let \( n \in \mathbb{Z} \) and \( u \in G_{g}^{m}(K_{v}) \).

(a) \( \theta(u \cdot (q \cdot 2n), q) = (t^{n}q \cdot n)Q^{-1}(t^{n}u) - \theta(u, q) \).

(b) \( v(\theta(u \cdot (q \cdot 2n), q)) = v(\theta(u, q)) - t^{n}Qn - t^{n}v(u) \).

**Proof.** We give the elementary verification in Appendix A see Proposition A.1.

**Proposition 4.4.** The function

\[ \Lambda(\cdot, q) : G_{g}^{m}(K_{v}) \to \mathbb{R}, \]

\[ \Lambda(u, q) = v(\theta(u, q)) + \frac{1}{4} t^{n}v(u)Q^{-1}v(u), \]

is \( \Omega \)-invariant, and hence descends to a function

\[ \Lambda(\cdot, q) : A(K_{v}) \cong G_{g}^{m}(K_{v})/\Omega \to \mathbb{R}. \]

**Proof.** We give the elementary verification in Appendix A see Proposition A.2.

**Theorem 4.5.** Let \( (A, \Theta)/K_{v} \) be a principally polarized abelian surface having totally split multiplicative reduction, where \( \Theta \in \text{Div}_{K}(A) \) is an effective symmetric principal polarization, and let \( q \subset G_{g}^{m}(K_{v}) \) be an associated multiplicative period matrix.

(a) There is a 2-torsion point \( T_{0} \in A[2] \) so that

\[ \Theta = \Theta_{0} + T_{0} \quad \text{with} \quad \Theta_{0} = \text{div}(\theta(\cdot, q)). \]

(b) Let

\[ P : G_{g}^{m}(\tilde{K}_{v}) \to A(\tilde{K}_{v}) \]

denote the \( v \)-adic analytic uniformization of \( A \). Then there is a \( \kappa'_{v} \in \mathbb{Q} \) so that for all \( u \in G_{g}^{m}(\tilde{K}_{v}) \),

\[ \hat{\lambda}_{\Theta_{0}, v}(P(u)) = v(\theta(u, q)) + \frac{1}{4} t^{n}v(u)Q^{-1}v(u) - \kappa'_{v}. \]
(c) Write
\[
\hat{\lambda}_{\Theta,v}(P) = \langle \Theta_0 \cdot \overline{P} \rangle_{A,v} + \mathcal{B}_{\Theta,v}(j_v(P)) - \kappa_v
\]
\[
= \hat{\lambda}_{\Theta,v}^{I}(P) + \hat{\lambda}_{\Theta,v}^{\mathcal{B}}(P)
\]
as in (6) and (9). Then
\[
\hat{\lambda}_{\Theta,v}^{I}(P + T_0) = \max_{u \in \mathcal{G}_m(\mathcal{K}_v)} \frac{1}{4} v(\theta(u,q)) + v(\mathcal{P}(u)) = \mathcal{P}(u = P)
\]
\[
\hat{\lambda}_{\Theta,v}^{\mathcal{B}}(P + T_0) = \min_{u \in \mathcal{G}_m(\mathcal{K}_v)} \frac{1}{4} v(\theta(u,q)) + v(\mathcal{P}(u)) = \mathcal{P}(u = P)
\]
(13)

In the next section we are going to give an explicit formula for the Fourier series of the Bernoulli local height $\hat{\lambda}_{\Theta,v}^{\mathcal{B}}$ when $\dim(A) = 2$. For notational reasons, it is easier to renormalize the lattice and work with the standard torus $\mathbb{R}^g/\mathbb{Z}^g$. Roughly speaking, we want to write $u \in \mathcal{G}_m(\mathcal{K}_v)$ as a (multiplicative) linear combination of the lattice vectors. But since we only require the valuations, we define a function
\[
x : \mathcal{G}_m(\mathcal{K}_v) \rightarrow \mathbb{R}^g, \quad x(u) = Q^{-1} v(u).
\]
(14)
For $P \in \mathcal{A}(\mathcal{K}_v)$, we write $P = \mathcal{P}(u_P)$ for some choice of $u_P \in \mathcal{G}_m(\mathcal{K}_v)$, and then we set
\[
x_P = x(u_P) \in \mathbb{R}^g.
\]
We note that $x_P$ is well defined in $\mathbb{R}^g/\mathbb{Z}^g$.

We associate to the period valuation matrix $Q$ the “periodic quadratic form”
\[
L_Q : \mathbb{R}^g \rightarrow \mathbb{R}, \quad L_Q(x) = \min_{x \equiv x_0 \pmod{\mathbb{Z}^g}}^t x Q x,
\]
(15)
and we write its associated Fourier series as
\[
L_Q(x) = \sum_{n \in \mathbb{Z}^g} \hat{L}_Q(n) e^{2\pi i n x}.
\]

Then
\[
\frac{1}{4} \hat{L}_Q(0) = \int_{\mathbb{R}^g/\mathbb{Z}^g} \frac{1}{4} L_Q(x) dx
\]
\[
= \lim_{N \to \infty, N \text{ even}} \sum_{t \in N^{-1} \mathbb{Z}^g} \frac{1}{4} L_Q(t)
\]
\[
= \lim_{N \to \infty, N \text{ even}} \sum_{T \in \mathcal{A}[N]} \left( \hat{\lambda}_{\Theta,v}^{\mathcal{B}}(T) + \kappa_v^{\mathcal{B}} \right) \quad \text{from (13)},
\]
\[
= \kappa_v^{\mathcal{B}} \quad \text{from (8)}.
\]
(16)
Then the Bernoulli local height is given by the formula
\[
\lambda_{\Theta,v}^B(P + T_0) = \min_{v(u) = P} \frac{1}{4} t v(u) Q^{-1} v(u) - \kappa_v^B \quad \text{from (13)},
\]
\[
= \min_{v(u) = P} \frac{1}{4} t x(u) Q x(u) - \kappa_v^B \quad \text{from (14) and} \quad t Q = Q,
\]
\[
= \frac{1}{4} L_Q(x_P) - \kappa_v^B \quad \text{from (15)},
\]
\[
= \frac{1}{4} L_Q(x_P) - \frac{1}{4} \hat{L}_Q(0) \quad \text{from (16)}.
\]

We record this result as a proposition.

**Proposition 4.6.** With notation as in this and the previous section,
\[
\lambda_{\Theta,v}^B(P + T_0) = \frac{1}{4} L_Q(x_P) - \frac{1}{4} \hat{L}_Q(0) \quad \text{for all} \quad P \in A(\bar{K}_v).
\]

5. **A Hexagonal Fourier Calculation**

**Definition 5.1.** Figure 2 gives a list of notation and conventions that will remain fixed throughout the remainder of this article.

We note that the following formal identities are true in the polynomial ring \(\mathbb{Z}[a, b, c, m, n]:\)

\[
\begin{align*}
F_0 - aF_1 &= Dn, \\
F_0 - \gamma F_2 &= Dm, \\
F_0 + bF_3 &= D(m + n), \\
aF_1 + bF_2 &= Dm, & bF_1 + cF_2 &= Dn, \\
aF_1 + bF_3 &= Dm, & -\gamma F_1 + cF_3 &= Dn, \\
-\alpha F_2 + aF_3 &= Dm, & \gamma F_2 + bF_3 &= Dn.
\end{align*}
\]

Our next result gives the Fourier expansion of \(L(x, y),\) which is the \(\mathbb{Z}^2\)-periodic version of the quadratic form \(F.\)

**Theorem 5.2.** With notation as in Figure 2 and in particular with \(L : (\mathbb{R}/\mathbb{Z})^2 \to \mathbb{R}\) being the periodic function
\[
L(x, y) = \min_{\xi \in \mathbb{Z}, \eta \in \mathbb{Z}} \alpha \xi^2 + 2b \xi \eta + c \eta^2,
\]
the Fourier expansion
\[
L(x, y) = \sum_{m, n \in \mathbb{Z}} \hat{L}(m, n) e(mx + ny)
\]
Figure 1. The hexagon where $F = L$, and the associated decomposition of the unit square as an octagon and four triangles

$$Q_{12} = \left( \frac{a(c-b)}{2(ac-b^2)}, \frac{a(c-b)}{2(ac-b^2)} \right)$$

$$Q_{34} = \left( -\frac{a(c-b)}{2(ac-b^2)}, \frac{c(a-b)}{2(ac-b^2)} \right)$$
If we need to specify $a, b, c$ in the notation, we write

$$F(x, y) = F_{a,b,c}(x, y) = F(a, b, c; x, y),$$
$$L(x, y) = L_{a,b,c}(x, y) = L(a, b, c; x, y),$$

and similarly for $F_0, \ldots, F_3$. We say that $F$, $L$, and $(a, b, c)$ are normalized if they satisfy (cf. Lemma 4.1)

$$c \geq a \geq 2b \geq 0.$$

If we are working over the $v$-adic completion of a field, all of the associated quantities will have a subscript $v$, e.g., $(a_v, b_v, c_v)$ and $L_v$ and $F_{i,v}$.

**Figure 2.** Notation and Conventions and Formulas

of $L(x, y)$ has Fourier coefficients given by the following formulas$^{13}$

$$L(m, n) = \begin{cases} \frac{a^2c + ac^2 - 2b^5}{12D} & \text{if } (m, n) = (0, 0), \\ \frac{(-1)^n \alpha c^2}{2\pi^2 Dn^2} & \text{if } F_1 = cm - bn = 0, (m, n) \neq (0, 0), \\ \frac{(-1)^m \gamma a^2}{2\pi^2 Dm^2} & \text{if } F_2 = an - bm = 0, (m, n) \neq (0, 0), \\ \frac{(-1)^{m+n+1} \alpha \gamma b}{2\pi^2 Dmn} & \text{if } F_3 = \gamma m + \alpha n = 0, (m, n) \neq (0, 0), \\ \frac{D^2 \sin \left( \frac{\cos m + a \gamma n}{2D} \right)}{4\pi^3 (cm - bn)(an - bm)(\gamma m + \alpha n)} & \text{otherwise.} \end{cases}$$

$^{13}$We note that since $ac \neq 0$, the assumptions $F_1 = 0$ and $(m, n) \neq (0, 0)$ imply that $n \neq 0$, and similarly $F_2 = 0$ and $(m, n) \neq (0, 0)$ imply that $m \neq 0$. Further, the fact that $D = ac - b^2 \neq 0$ tells us that at least one of $\alpha = a - b$ and $\gamma = c - b$ is non-zero, so $F_3 = 0$ and $(m, n) \neq (0, 0)$ implies that at least one of $m$ and $n$ is non-zero. And if we normalize $a, b, c$, then $\alpha \gamma \neq 0$, so $F_3 = 0$ and $(m, n) \neq (0, 0)$ implies that both $m$ and $n$ are non-zero.
Proof. The region
\[ \mathcal{H} = \{(x, y) \in \mathbb{R}^2 : F(x, y) = L(x, y)\} \]
where \( F \) and \( L \) are equal is a hexagon, as shown in the bottom illustration in Figure 1. The intersection of \( \mathcal{H} \) with the unit square \( S = \{(x, y) \in \mathbb{R}^2 : |x| \leq \frac{1}{2}, |y| \leq \frac{1}{2}\} \) is the central octagon in both illustrations in Figure 1. The set difference \( S \setminus \mathcal{H} \) consists of four triangles, which are shifted versions of the set difference \( \mathcal{H} \setminus S \) again as shown in Figure 1. We label the four triangular regions as follows:

| Region  | Equation |
|---------|----------|
| I       | \( L(x, y) = F(x, y - 1) \) |
| II      | \( L(x, y) = F(x - 1, y) \) |
| III     | \( L(x, y) = F(x, y + 1) \) |
| IV      | \( L(x, y) = F(x + 1, y) \) |

We start with some observations that we use in the computation of the Fourier coefficients of \( L \). The functions \( F \) and \( L \) have the following symmetries:

\[
F_{a,b,c}(x, y) = F_{a,b,c}(-x, -y) = F_{c,b,a}(y, x),
L_{a,b,c}(x, y) = L_{a,b,c}(-x, -y) = L_{c,b,a}(y, x).
\]

The sign change symmetry identifies Regions I and III, leading to the equality

\[
\int_{\text{III}} \left\{ F_{a,b,c}(x, y - 1) - F_{a,b,c}(x, y) \right\} e(mx + ny)
= \int_{\text{I}} \left\{ F_{a,b,c}(x, y + 1) - F_{a,b,c}(x, y) \right\} e(-mx - ny),
\]
and similarly for Regions II and IV. The reflection symmetry together with the parameter swap \( a \leftrightarrow c \) identifies Regions I and II, leading to the equality

\[
\int_{\text{II}} \left\{ F_{a,b,c}(x, y - 1) - F_{a,b,c}(x, y) \right\} e(mx + ny)
= \int_{\text{I}} \left\{ F_{c,b,a}(x, y - 1) - F_{c,b,a}(x, y) \right\} e(mx + ny),
\]
and similarly for Regions III and IV.

Using these observations, we find that

\[
\hat{L}(m, n) = \iint_{S} L(x, y) e(mx + ny)
\]
\[
\begin{align*}
&= \iint_S F(x, y)e(mx + ny) \\
&\quad + \left( \iint_{I} + \iint_{II} + \iint_{III} + \iint_{IV} \right) \{ L(x, y) - F(x, y) \} e(mx + ny) \\
&= \iint_S F(x, y)e(mx + ny) \\
&\quad + \iint_{I} \{ F(x, y - 1) - F(x, y) \} e(mx + ny) \\
&\quad + \iint_{II} \{ F(x - 1, y) - F(x, y) \} e(mx + ny) \\
&\quad + \iint_{III} \{ F(x, y + 1) - F(x, y) \} e(mx + ny) \\
&\quad + \iint_{IV} \{ F(x + 1, y) - F(x, y) \} e(mx + ny) \\
&\quad + 2 \iint_{II} \{ F(x - 1, y) - F(x, y) \} \cos(mx + ny) \\
&\quad + 2 \iint_{III} \{ F(x - 1, y) - F(x, y) \} \cos(mx + ny) \\
&= \iint_S F(x, y) \cos(mx + ny) \\
&\quad + 2 \int \int_{I} \{ F(x, y - 1) - F(x, y) \} \cos(mx + ny) \\
&\quad + 2 \int \int_{II} \{ F(x - 1, y) - F(x, y) \} \cos(mx + ny) \\
&\quad + 2 \int \int_{III} \{ F(x, y + 1) - F(x, y) \} \cos(mx + ny) \\
&\quad + 2 \int \int_{IV} \{ F(x + 1, y) - F(x, y) \} \cos(mx + ny) \\
&\quad + \int \int_{I} \{ F(x, y - 1) - F(x, y) \} \cos(mx + ny) \\
&\quad + \int \int_{II} \{ F(x - 1, y) - F(x, y) \} \cos(mx + ny) \\
&\quad + \int \int_{III} \{ F(x, y + 1) - F(x, y) \} \cos(mx + ny) \\
&\quad + \int \int_{IV} \{ F(x + 1, y) - F(x, y) \} \cos(mx + ny) \\
&= \int \int_{\frac{1}{2}} \frac{1}{2} (ax^2 + 2bxy + cy^2) \cos(mx + ny) \, dx \, dy \\
&\quad + 2 \int \int_{\frac{1}{2}} \frac{1}{2} \left( \frac{c-b}{a-b} \right) \left( \frac{1}{2} - \frac{c-b}{a-b} \right) \cos(mx + ny) \, dx \, dy \\
&\quad + \int \int_{\frac{1}{2}} \frac{1}{2} \left( \frac{a-b}{c-b} \right) \left( \frac{1}{2} - \frac{a-b}{c-b} \right) \cos(mx + ny) \, dx \, dy. \\
\end{align*}
\]

It is now an easy task\[14\] to compute these integrals. We start with the case that \(m\) and \(n\) are non-zero integers, and we assume for the moment that 

\[ F_1(m, n)F_2(m, n)F_3(m, n) \neq 0. \]

Then the integral over Regions I and III is given explicitly by 

\[ \left( \iint_{I} + \iint_{III} \right) \{ L(x, y) - F(x, y) \} e(mx + ny) \]

\[14\] Easy using a computer algebra system such as Mathematica, otherwise it is a feasible, but tedious, task.
\[
\int_{1}^{\frac{a+c-b}{2(ac-b^2)}} \int_{\frac{c}{b} \left(\frac{1}{2}-y\right)}^{\frac{1}{2}-\frac{c-b}{2(a-b)}} (c-2bx - 2cy) \cdot \cos(mx + ny) \, dx \, dy \\
= \frac{b(a-b)}{4\pi^2 m((c-b)m + (a-b)n)} \cdot (-1)^{m+n} \\
+ \frac{1}{4\pi^3 (cm-bn)} \cdot \left(\frac{ac-b^2}{(c-b)m + (a-b)n}\right)^2 \\
\times \sin\left(\frac{\left(c(a-b)m + a(c-b)n\right)}{2(ac-b^2)}\right) \\
= \frac{(-1)^{m+n}b\alpha}{4\pi^2 mF_3} + \frac{D^2}{4\pi^3 F_1^2F_3^2} \sin\left(\frac{F_0}{2D}\right). 
\]

Further, as noted earlier, the integral over Regions II and IV is the same with the swaps \(a \leftrightarrow c\) and \(m \leftrightarrow n\). Hence

\[
\left(\int_{I} + \int_{\text{II}} + \int_{\text{III}} + \int_{IV}\right) \{L(x,y) - F(x,y)\} \, e^{mx + ny} \, dx \, dy \\
= \left\{ \frac{(-1)^{m+n}b\alpha}{4\pi^2 mF_3} + \frac{D^2}{4\pi^3 F_1^2F_3^2} \sin\left(\frac{F_0}{2D}\right) \right\} \\
+ \left\{ \frac{(-1)^{m+n}b\gamma}{4\pi^2 nF_3} + \frac{D^2}{4\pi^3 F_2^2F_3^2} \sin\left(\frac{F_0}{2D}\right) \right\} \\
= \frac{(-1)^{m+n}b}{4\pi^2 mn} + \frac{D^2 \sin\left(F_0/2D\right)}{4\pi^3 F_1^2F_2F_3^3}. 
\]

On the other hand, the integral of \(F\) over the square is simply

\[
\int_{S} F(x,y) \cos(mx + ny) = \frac{(-1)^{m+n+1}b}{4\pi^2 mn}, 
\]

which cancels the first term in the sum of the four-triangle integrals. Hence

\[
\hat{L}(m,n) = \frac{D^2 \sin\left(F_0/2D\right)}{4\pi^3 F_1^2F_2F_3^3}. \tag{18} 
\]

One can check by a direct computation that the formula (18) for \(\hat{L}(m,n)\) is valid if one, but not both, of \(m\) and \(n\) is 0. (This despite the fact that \(m\) and \(n\) seem to appear in the denominators of some of the intermediate calculations.)

\[\text{We note that these swaps leave } F_0 \text{ and } F_3 \text{ invariant, while swapping } \alpha \leftrightarrow \gamma \text{ and } F_1 \leftrightarrow F_2.\]

\[\text{Presumably this cancellation is not a coincidence!}\]
We note that we can use (17) to rewrite the formula for \( \hat{L}(m, n) \) so that the argument of the sine function is instead related to one of the other \( F_i \). Thus

\[
\begin{align*}
\sin \left( \frac{F_0}{2D} \right) &= \sin \left( \frac{\alpha F_1 + Dn}{2D} \right) = (-1)^n \sin \left( \frac{\alpha F_1}{2D} \right), \\
\sin \left( \frac{F_0}{2D} \right) &= \sin \left( \frac{\gamma F_2 + Dm}{2D} \right) = (-1)^m \sin \left( \frac{\gamma F_2}{2D} \right), \\
\sin \left( \frac{F_0}{2D} \right) &= \sin \left( \frac{-b F_3 + D(m + n)}{2D} \right) = (-1)^{m+n+1} \sin \left( \frac{b F_3}{2D} \right).
\end{align*}
\]

Substituting these into (18) gives three additional formulas for \( \hat{L}(m, n) \),

\[
\begin{align*}
\hat{L}(m, n) &= \frac{(-1)^n D^2 \sin(\alpha F_1/2D)}{4 \pi^3 F_1 F_2 F_3}, \\
\hat{L}(m, n) &= \frac{(-1)^m D^2 \sin(\gamma F_2/2D)}{4 \pi^3 F_1 F_2 F_3}, \\
\hat{L}(m, n) &= \frac{(-1)^{m+n+1} D^2 \sin(b F_3/2D)}{4 \pi^3 F_1 F_2 F_3}.
\end{align*}
\]

Continuing with our assumption that \((m, n) \neq (0, 0)\), we consider the case that one of \( F_1, F_2, F_3 \) vanishes. An important observation is that (17) and the fact that \( D \neq 0 \) tells us that at most one of these \( F_i(m, n) \) can vanish.

We fix an integer pair \((m, n) \neq (0, 0)\), we take a sequence of values of \((a, b, c)\) that cause one of the \( F_i \) to vanish. The integrals that occur in the computation of \( \hat{L}(a, b, c; m, n) \) are integrals of a continuous function \( L \) over a compact set, where \( L \) depends continuously on \((a, b, c)\), so we can move the limit as \( F_i(a, b, c; m, n) \to 0 \) across the integral.

Thus (19) yields

\[
\lim_{F_1 \to 0} \hat{L}(m, n) = \lim_{F_1 \to 0} \frac{(-1)^n D^2 \sin(\alpha F_1/2D)}{4 \pi^3 F_1 F_2 F_3} = \frac{(-1)^n D^2}{4 \pi^2} \lim_{F_1 \to 0} \frac{1}{F_2 F_3}.
\]

As \( F_1 = cm - bn \to 0 \), we have (note that \( c \neq 0 \))

\[
\lim_{F_1 \to 0} F_2 = \lim_{cm \to bn} an - bm = \lim_{cm \to bn} an - \frac{b}{c} \cdot cm = an - \frac{b}{c} \cdot bn = \frac{Dn}{c},
\]
\[
\lim_{F_1 \to 0} F_3 = \lim_{cm \to bn} \gamma m + \alpha n = \lim_{cm \to bn} \frac{\gamma}{c} \cdot cm + \alpha n = \frac{\gamma}{c} \cdot bn + \alpha n = \frac{Dn}{c}.
\]
Substituting these two limits into (22) yields
\[
\lim_{F_1 \to 0} \hat{L}(m, n) = \left(-1\right)^n \alpha c^2 4\pi^2 Dn^2.
\]
Similar computations using (20) and (21) give the analogous formulas for the values of \( \hat{L}(m, n) \) as \( F_2 \to 0 \) and as \( F_3 \to 0 \). We leave the details to the reader.

It remains to compute \( \hat{L}(0, 0) \), for which the relevant integrals are easy and left as an exercise. This concludes the proof of Theorem 5.2.

\[\square\]

6. Averaging (Periodic) Functions over (Torsion) Points

We introduce a convenient notation for the expected value (average) of a function over a set, and in particular over the \( d \)-torsion points of an abelian group.

**Definition 6.1.** Let \( S \) be a finite set, and let \( f : S \to \mathbb{R} \) be a real-valued function. We write
\[
\mathcal{Avg}_S f(x) = \frac{1}{\# S} \sum_{x \in S} f(x).
\]
Similarly,
\[
\mathcal{Avg}_{x, y \in S \atop x \neq y} f(x - y) = \frac{1}{\# S^2 - \# S_{x, y \in S \atop x \neq y}} \sum_{x, y \in S \atop x \neq y} f(x - y).
\]
If \( S = A \) is an abelian group and \( d \geq 1 \), by a slight abuse of notation we write
\[
(\mathcal{Avg}_d f)(x) = \mathcal{Avg}_S f(x + t) = \frac{1}{\# A[d]} \sum_{t \in A[d]} f(x + t)
\]
for the average of \( f \) at the \( d \)-torsion translates of \( x \), and we call \( \mathcal{Avg}_d f \) the \( d \)-average of \( f \).

**Example 6.2.** We illustrate Definition 6.1 with three examples:

1. For a function \( L : (\mathbb{R}/\mathbb{Z})^2 \to \mathbb{R} \) such as the one defined in Theorem 5.2, we have
\[
(\mathcal{Avg}_d L)(x, y) = \frac{1}{d^2} \sum_{i=0}^{d-1} \sum_{j=0}^{d-1} L \left( x + \frac{i}{d}, y + \frac{j}{d} \right).
\]
(2) For an abelian variety $A$ of dimension $g$ and a function $\lambda : A \to \mathbb{R}$, we have
\[
(\mathcal{A}_g \lambda)(P) = \frac{1}{d^{2g}} \sum_{T \in A[d]} \lambda(P + T).
\]

(3) For any integer $m$ and the function $e_m(x) = e^{2\pi imx}$, we have
\[
(\mathcal{A}_g e_m)(x) = \begin{cases} 
  e_m(x) & \text{if } d \mid m, \\
  0 & \text{if } d \nmid m.
\end{cases}
\]

**Definition 6.3.** The 2nd periodic Bernoulli polynomial is the function defined by
\[
B_2(x) = x^2 - x + \frac{1}{6} \text{ for } 0 \leq x \leq 1, \text{ and } B_2(x + n) = B_2(x) \text{ for } n \in \mathbb{Z}.
\]
The well-known Fourier expansion of $B_2(x)$ is
\[
B_2(x) = \frac{1}{2\pi^2} \sum_{k \in \mathbb{Z}} e(kx) \frac{k^2}{k^2}, \tag{23}
\]
from which we immediately obtain the distribution relation
\[
(\mathcal{A}_N B_2)(x) = \frac{1}{N^2} B_2(Nx). \tag{24}
\]

We recall a Fejér kernel type estimate for $B_2$.

**Lemma 6.4.** Let $R \geq 1$ be an integer, and let $T \subset \frac{1}{R} \mathbb{Z}$ with $N = \#T$ be a set of $N$ distinct rational numbers whose denominators divide $R$. Then
\[
\mathcal{A}_{s,t \in T, s \neq t} B_2(s - t) \geq \frac{1}{6R^2} - \frac{1}{6(N - 1)}.
\]

**Proof.** Let $T = \{t_1, \ldots, t_N\}$. We compute
\[
\mathcal{A}_{s,t \in T, s \neq t} B_2(s - t)
\]
\[
= \frac{1}{N^2 - N} \sum_{i,j=1}^{N} B_2(t_i - t_j)
\]
\[
= \frac{1}{N^2 - N} \sum_{i,j=1}^{N} \frac{1}{2\pi^2} \sum_{k \in \mathbb{Z}} e(k(t_i - t_j)) \frac{k^2}{k^2} \text{ from (23),}
\]
\[ = \frac{1}{2\pi^2(N^2 - N)} \sum_{k \in \mathbb{Z}}' \frac{1}{k^2} \sum_{i,j=1}^{N} e(k(t_i - t_j)) \]

\[ = \frac{1}{2\pi^2(N^2 - N)} \sum_{k \in \mathbb{Z}}' \frac{1}{k^2} \left\{ \left\lfloor \sum_{i=1}^{N} e(kt_i) \right\rfloor^2 - N \right\} \]

this quantity is always \(\geq 0\), and if \(R \mid k\), then it equals \(N^2\)

\[ \geq \frac{1}{2\pi^2(N^2 - N)} \sum_{k \in \mathbb{Z}}' \frac{N^2 - N}{(Rk)^2} - \frac{1}{2\pi^2(N^2 - N)} \sum_{k \in \mathbb{Z}}' \frac{N}{k^2} \]

\[ = \frac{1}{2\pi^2 R^2} \cdot 2\zeta(2) - \frac{1}{2\pi^2(N - 1)} \cdot 2\zeta(2) \]

\[ = \frac{1}{6R^2} - \frac{1}{6(N - 1)} \cdot 2\zeta(2). \]

This completes the proof of Lemma 6.4. \(\Box\)

We next express certain \(d\)-averages of the function \(L(x, y)\) in Theorem 5.2 in terms of \(d\)-averages of the second Bernoulli polynomial.

**Corollary 6.5.** Let \(a, b, c \in \mathbb{Z}\) with \(D = ac - b^2 > 0\), let \(\alpha = a - b\) and \(\gamma = c - b\), and let \(d\) be an integer satisfying

\[ d \equiv 0 \pmod{2D \gcd(a, b, c)^2}. \tag{25} \]

Then the \(d\)-average of the \(\mathbb{Z}^2\)-periodic function

\[ L(x, y) = \min_{\xi \in x + \mathbb{Z}} a\xi^2 + 2b\xi + c\eta^2 \]

is given by the formula

\[ \mathcal{A}_{d} L(x, y) = \widehat{L}(0, 0) + \frac{\alpha(c, b)^2}{Dd^2} B_2 \left( \frac{d(bx + cy)}{\gcd(c, b)} \right) + \frac{\gamma(a, b)^2}{Dd^2} B_2 \left( \frac{d(ax + by)}{\gcd(a, b)} \right) + \frac{b(\alpha, \gamma)^2}{Dd^2} B_2 \left( \frac{d(ax - \gamma y)}{\gcd(\alpha, \gamma)} \right). \tag{26} \]

**Proof.** The congruence condition (25) says that \(d\) satisfies \(d \gcd(a, b, c)^2 \in \mathbb{Z}\).
which in turn implies that
\[
\sin \left( \frac{\alpha \gamma \beta}{2D} \right) = \sin \left( \pi \cdot \frac{\gcd(a, b, c)^2}{D} \cdot \frac{\alpha m + \beta n}{\gcd(a, b, c)^2} \right) = 0,
\]
since \(a, c, \alpha, \gamma\) are all divisible by \(\gcd(a, b, c)\). Then Theorem 5.2 says that the associated Fourier coefficient satisfies
\[
\hat{L}(dm, dn) = 0 \quad \text{unless} \quad F_1(m, n)F_2(m, n)F_3(m, n) = 0.
\]

We note that if \(D \neq 0\) and \((m, n) \neq (0, 0)\), then at most one of the linear forms \(F_1, F_2, F_3\) may vanish, so aside from \(\hat{L}(0, 0)\), the Fourier series splits into three sums. We compute (note \(d\) is even, so the powers with \((-1)^d\) may be omitted)
\[
\mathcal{M}_{\phi_d} L(x, y) - \hat{L}(0, 0) = \sum'_{m,n \in \mathbb{Z}} \frac{(-1)^{dn} \alpha c^2}{2\pi^2 D(dn)^2} e(dx + dy)
\]
\[
+ \sum'_{m,n \in \mathbb{Z}} \frac{(-1)^{dm} \gamma a^2}{2\pi^2 D(dm)^2} e(dx + dy)
\]
\[
+ \sum'_{m,n \in \mathbb{Z}} \frac{(-1)^{dm+dn+1} \alpha \gamma b}{2\pi^2 D(dm)(dn)} e(dx + dy)
\]
using the formulas for \(\hat{L}(m, n)\) from Theorem 5.2
\[
= \frac{\alpha c^2}{2\pi^2 D d^2} \sum'_{m,n \in \mathbb{Z}} \frac{1}{n^2} e(dx + dy)
\]
\[
+ \frac{\gamma a^2}{2\pi^2 D d^2} \sum'_{m,n \in \mathbb{Z}} \frac{1}{m^2} e(dx + dy)
\]
\[
+ \frac{\alpha \gamma b}{2\pi^2 D d^2} \sum'_{m,n \in \mathbb{Z}} \frac{1}{mn} e(dx + dy).
\]

We rewrite the last three sums using
\[
\{(m, n) \in \mathbb{Z}^2 : F_1(m, n) = 0\} = \left\{ \left( \frac{bk}{(c, b)}, \frac{ck}{(c, b)} \right) : k \in \mathbb{Z} \right\},
\]
This yields
\[
\mathcal{A}_{\mathbb{Z}} L(x, y) - \hat{L}(0, 0) = \frac{\alpha(c, b)^2}{2\pi^2 Dd^2} \sum_{k \in \mathbb{Z}}' \frac{1}{k^2} e \left( \frac{d(bx + cy)}{(c, b)}k \right) + \frac{\alpha(c, b)^2}{2\pi^2 Dd^2} \sum_{k \in \mathbb{Z}}' \frac{1}{k^2} e \left( \frac{d(ax + by)}{(a, b)}k \right) + \frac{b(\alpha, \gamma)^2}{2\pi^2 Dd^2} \sum_{k \in \mathbb{Z}}' \frac{1}{k^2} e \left( \frac{d(\alpha x - \gamma y)}{(\alpha, \gamma)}k \right).
\] (27)

Using the Fourier series (23) for \( B_2 \) for the three sums in (27) gives the desired result. \( \square \)

7. TWO LOWER BOUNDS FOR THE LOCAL HEIGHT

In this section we prove two lower bounds for averages of the Bernoulli part of the local height, one via Fourier averaging and one via the pigeonhole principle. Both estimates will be used in the proof of our Lehmer-type lower bound for the global height. The notation in Figure 3 is used in the statement of both lemmas.

7.1. A Local Height Lower Bound via Fourier Averaging. The main result of this section is an abelian surface analogue of the elliptic curve result [16, Proposition 1.2]. In order to handle the fact that for abelian surfaces, many of the Fourier coefficients of the Bernoulli part of the local height are negative, the proof includes an average over \( d \)-torsion points that eliminates the negative coefficients. For our eventual application to Lehmer-type height bounds, it is crucial that the value of \( d \) does not change when the base field is replaced by a (ramified) extension.

Lemma 7.1. With notation as in Figure 3, we have\(^{17}\)
\[
\mathcal{A}_{\mathbb{Z}} \mathcal{V} \mathcal{A}_{\mathbb{Z}} \lambda_{\Theta, v}^{\mathbb{A}}(P - Q + T) \geq \frac{1}{24d^2} \left( \frac{\alpha + \gamma + b}{D} - \frac{\alpha(c, b)^2 + \gamma(a, b)^2 + b(\alpha, \gamma)^2}{D(N - 1)} \right).
\]

\(^{17}\)We recall that although \( \lambda_{\Theta, v} \) is only defined on the complement of the support of its associated divisor, we can extend \( \lambda_{\Theta, v}^{\mathbb{A}} \) to all of \( A(K_v) \). See Remark 3.5.
$K_v$ a field that is complete with respect to a non-archimedean absolute value $v$.

$(A, \Theta)/K_v$ an abelian variety $A$ defined over $K_v$ with an effective symmetric principal polarization $\Theta$, and such that $A$ has totally split multiplicative reduction.

$(a, b, c)$ a normalized period valuation triple for $A/K_v$, i.e., if the period matrix is $q$, then

$$a = v(q_{11}), \quad b = v(q_{12}) = v(q_{21}), \quad c = v(q_{22}).$$

$D = ac - b^2$.

$d$ a positive integer satisfying

$$d \equiv 0 \pmod{\gcd(a, b, c)^2}.$$

$\Sigma$ a finite subset of $A(K_v)$.

$N = \#\Sigma$.

**Figure 3.** Notation and Setup for Lemmas 7.1 and 7.2.

**Proof.** We first note that since we are averaging over $d$-torsion points and $d$ is even, we may as well replace the principal polarization $\Theta$ with the divisor of $\theta(u, q)$, since they differ by a 2-torsion point that will disappear when we take the average; cf. Theorem 4.5.

An important observation is that for any point $P$, the vector $(x_P, y_P)$ is given by the coordinates of $P$ in the group $\mathbb{Z}^2/A\mathbb{Z}^2$ relative to the basis given by the columns of the matrix $A = (a \ b \ c)$. Thus

$$\begin{pmatrix} x_P \\ y_P \end{pmatrix} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}^{-1} \begin{pmatrix} u_P \\ v_P \end{pmatrix} = \frac{1}{D} \begin{pmatrix} c & -b \\ -b & a \end{pmatrix} \begin{pmatrix} u_P \\ v_P \end{pmatrix}$$

for some $u_p, v_p \in \mathbb{Z}$. (28)

This yields the useful formulas

$$b x_P + cy_P = v_P, \quad ax_P + by_P = u_p, \quad ax_P - \gamma y_P = u_p - v_p. \quad (29)$$

We also note that for any points $P$ and $Q$, we have

$$x_{P-Q} \equiv x_P - x_Q \pmod{\mathbb{Z}} \quad \text{and} \quad y_{P-Q} \equiv y_P - y_Q \pmod{\mathbb{Z}}. \quad (30)$$
To ease notation, we drop the gcd from the notation \( \gcd(a, b) \). We compute

\[
\mathbb{E}_{P, \Theta, \nu} \mathbb{E}_{Q, T \in A[d]} (P - Q + T)
\]

\[
= \mathbb{E}_{P, \Theta, \nu} \mathbb{E}_{Q, T \in A[d]} \left( L(x_{P - Q}, y_{P - Q}) - \hat{L}(0, 0) \right)
\]

\[
= \frac{1}{N^2 - N} \sum_{P, Q \in \Sigma \atop P \neq Q} \left\{ \frac{\alpha(c, b)^2}{Dd^2} B_2 \left( \frac{d(bx_{P - Q} + cy_{P - Q})}{(c, b)} \right) + \frac{\gamma(a, b)^2}{Dd^2} B_2 \left( \frac{d(ax_{P - Q} + by_{P - Q})}{(a, b)} \right) + \frac{b(\alpha, \gamma)^2}{Dd^2} B_2 \left( \frac{d(\alpha x_{P - Q} - \gamma y_{P - Q})}{(\alpha, \gamma)} \right) \right\}
\]

from Proposition 4.6,

\[
= \frac{\alpha(c, b)^2}{Dd^2} \frac{1}{N^2 - N} \sum_{P, Q \in \Sigma \atop P \neq Q} B_2 \left( \frac{d(bx_P + cy_P) - d(bx_Q + cy_Q)}{(c, b)} \right)
\]

from Corollary 6.5,

\[
+ \frac{\gamma(a, b)^2}{Dd^2} \frac{1}{N^2 - N} \sum_{P, Q \in \Sigma \atop P \neq Q} B_2 \left( \frac{d(ax_P + by_P) - d(ax_Q + by_Q)}{(a, b)} \right)
\]

\[
+ \frac{b(\alpha, \gamma)^2}{Dd^2} \frac{1}{N^2 - N} \sum_{P, Q \in \Sigma \atop P \neq Q} B_2 \left( \frac{d(\alpha x_P - \gamma y_P) - d(\alpha x_Q - \gamma y_Q)}{(\alpha, \gamma)} \right)
\]

from (30),

\[
= \frac{\alpha(c, b)^2}{Dd^2} \frac{1}{N^2 - N} \sum_{P, Q \in \Sigma \atop P \neq Q} B_2 \left( \frac{d(v_P - v_Q)}{(c, b)} \right)
\]

\[
+ \frac{\gamma(a, b)^2}{Dd^2} \frac{1}{N^2 - N} \sum_{P, Q \in \Sigma \atop P \neq Q} B_2 \left( \frac{d(u_P - u_Q)}{(a, b)} \right)
\]

\[
+ \frac{b(\alpha, \gamma)^2}{Dd^2} \frac{1}{N^2 - N} \sum_{P, Q \in \Sigma \atop P \neq Q} B_2 \left( \frac{d((u_P - v_P) - (u_Q - v_Q))}{(\alpha, \gamma)} \right)
\]

from (29),
\[
\begin{align*}
\geq & \frac{\alpha(c, b)^2}{Dd^2} \cdot \frac{1}{6} \left( \frac{1}{(c, b)^2} - \frac{1}{N - 1} \right) \\
+ & \frac{\gamma(a, b)^2}{Dd^2} \cdot \frac{1}{6} \left( \frac{1}{(a, b)^2} - \frac{1}{N - 1} \right) \\
+ & \frac{b(\alpha, \gamma)^2}{Dd^2} \cdot \frac{1}{6} \left( \frac{1}{(\alpha, \gamma)^2} - \frac{1}{N - 1} \right)
\end{align*}
\]
from Lemma 6.4, since \( d, u_P, v_P \in \mathbb{Z} \).

A little bit of algebra yields the desired result, which concludes the proof of Lemma 7.1. \( \square \)

7.2. A Local Height Lower Bound via the Pigeonhole Principle. The main result of this section is an analogue for abelian surfaces of [23, Lemma 4] and [16, Proposition 1.3]. However, the proof is intrinsically more complicated than in the case of elliptic curves, since it relies on a lower bound for the average of the local height over a carefully chosen set of torsion points, and that lower bound ultimately relies on the explicit Fourier expansion of the periodic quadratic form given in Theorem 5.2.

**Lemma 7.2.** With notation as in Figure 3 there exists a subset \( \Sigma' \subseteq \Sigma \) containing
\[
\#\Sigma' \geq 6^{-3} \#\Sigma
\]
elements such that for all distinct \( P, Q \in \Sigma' \) we have
\[
\mathcal{N}_{\Theta}^\ast \lambda_{\Theta, v}(P - Q) \geq \frac{\alpha(c, b)^2 + \gamma(a, b)^2 + b(\alpha, \gamma)^2}{144 Dd^2}.
\]

**Proof of Lemma 7.2.** As in the proof of Lemma 7.1 the fact that we’re taking the \( d \)-average with \( d \) even means that we may replace the principal polarization \( \Theta \) with the divisor of \( \theta(u, q) \).

We start with the formula
\[
\mathcal{N}_{\Theta}^\ast 4 \lambda_{\Theta, v}(R) = \mathcal{N}_{\Theta} L(x_R, y_R) - \tilde{L}(0, 0)
\]
from Proposition 4.6,
\[
= \frac{\alpha(c, b)^2}{Dd^2} B_2 \left( \frac{d(bx_R + cy_R)}{\gcd(c, b)} \right)
\]
from Corollary 6.5,
\[
+ \frac{\gamma(a, b)^2}{Dd^2} B_2 \left( \frac{d(ax_R + by_R)}{\gcd(a, b)} \right)
\]
\[
+ \frac{b(\alpha, \gamma)^2}{Dd^2} B_2 \left( \frac{d(ax_R - \gamma y_R)}{\gcd(\alpha, \gamma)} \right).
\]
(31)
To ease notation, we momentarily define
\[
\| \cdot \|_Z : \mathbb{R} \rightarrow \left[ 0, \frac{1}{2} \right], \quad \| t \|_Z = \min_{n \in \mathbb{Z}} |t + n|,
\]
i.e., \(\| t \|_Z\) is the distance from \(t\) to the closest integer to \(t\). It is easy to check that for all \(t \in \mathbb{R}\), the periodic Bernoulli polynomial satisfies
\[
\| t \|_Z \leq \frac{1}{6} \Rightarrow \mathbb{B}_2(t) \geq \frac{1}{36},
\]
since by periodicity and symmetry \(\mathbb{B}_2(-t) = \mathbb{B}_2(t)\), it suffices to check for \(0 \leq t \leq \frac{1}{6}\). Hence if \(R\) satisfies the three inequalities
\[
\begin{align*}
\| d(\alpha x_R + c y_R) \|_{(c, b)} &\leq \frac{1}{6}, \\
\| d(ax_R + by_R) \|_{(a, b)} &\leq \frac{1}{6}, \\
\| d(\alpha x_R - \gamma y_R) \|_{(\alpha, \gamma)} &\leq \frac{1}{6};
\end{align*}
\]
then each of the three Bernoulli polynomial values appearing in (31) is greater \(1/36\). This proves that
\[
\mathcal{N}_{\psi, \Theta, v}(R) \geq \frac{\alpha(c, b)^2 + \gamma(a, b)^2 + b(\alpha, \gamma)^2}{36Dd^2} \quad \text{if } R \text{ satisfies (32)}.
\]
We consider the map
\[
\Sigma \rightarrow (\mathbb{R}/\mathbb{Z})^3,
\]
\[
R \mapsto \left( \frac{d(bx_R + cy_R)}{(c, b)}, \frac{d(ax_R + by_R)}{(a, b)}, \frac{d(\alpha x_R - \gamma y_R)}{(\alpha, \gamma)} \right).
\]
We divide the centered fundamental domain for \((\mathbb{R}/\mathbb{Z})^3\) into \(6^3\) equally sized cubes whose sides have length \(6^{-1}\). Then the pigeon-hole principle ensures that we can find a subset
\[
\Sigma' \subseteq \Sigma \quad \text{with} \quad \#\Sigma' \geq B^{-3}\#\Sigma
\]
such that the points in \(\Sigma'\) all lie in the same small cube. It follows that for all pairs \(P, Q \in \Sigma'\) we have
\[
\begin{align*}
\| d(\alpha x - P + cy - Q) \|_{(c, b)} &\leq \frac{1}{6}, \\
\| d(ax - P + by - Q) \|_{(a, b)} &\leq \frac{1}{6}, \\
\| d(\alpha x - P + \gamma y - Q) \|_{(\alpha, \gamma)} &\leq \frac{1}{6};
\end{align*}
\]
We note that for the three equalities in (33), (34) and (35), we are using the fact that the quantities \( x_{P-Q} \) and \( y_{P-Q} \) are multiplied by integers. This combined with the fact that we are using the norm on \( \mathbb{R}/\mathbb{Z} \) justifies the equalities. Thus all differences of points in \( \Sigma' \) satisfy (32), which completes the proof of Lemma 7.2.

8. A Bound for Small Differences Lying on \( \Theta \)

As noted earlier, the Bernoulli part of the local height \( \hat{\lambda}_{\Delta^0}^\mathbb{R} \) is defined at every point, but the intersection part \( \hat{\lambda}_{\Delta^0}^\mathbb{I} \) is defined only away from the support of the associated divisor \( \Delta \). That means that if we want to use the local-global decomposition of the global height \( \hat{h}_{\Delta} \) described in Theorem 3.1(h), we must restrict to points lying in the complement \( A(\bar{K}) \setminus |\Delta| \) of the support of \( \Delta \). However, since ultimately we want to study points of small height, it will suffice to use the following lemma, whose proof relies on Ullmo and Zhang’s proof of the Bogomolov conjecture.

**Lemma 8.1.** Let \( \bar{K} \) be an algebraically closed field of characteristic 0, let \( A/\bar{K} \) be an abelian surface, let \( \Theta \subset A \) be an irreducible curve of genus at least 2, and let \( \hat{h}_A \) be a canonical height on \( A \) relative to some ample symmetric divisor. There are constants \( C_6, C_7 > 0 \) that depend only on \( A/\bar{K}, \Theta, \hat{h}_A \) so that for all finite subsets

\[
\Sigma \subset \Theta \cap \{ P \in A(\bar{K}) : \hat{h}_A(P) \leq C_6 \}
\]

there exists a subset \( \Sigma' \subset \Sigma \) satisfying

\[
\#\Sigma' \geq C_7 \cdot \#\Sigma \quad \text{and} \quad (P - Q + A_{\text{tors}}) \cap \Theta = \emptyset \text{ for all distinct } P, Q \in \Sigma'.
\]

**Proof.** The Bogomolov conjecture for (curves on) abelian varieties, which was proven by Ullmo [27] and Zhang [32], says that there is a constant \( C_8 > 0 \), depending only on \( A, \Theta, \hat{h}_A \), such that the set

\[
\Xi = \Xi(A, \Theta, \hat{h}_A) := \left( \Theta \cap \{ P \in A(\bar{K}) : \hat{h}_A(P) \leq C_6 \} \right)
\]

is finite. In other words, there are a bounded number of points of \( A \) that lie on \( \Theta \) and have small height.

We set \( C_6 = \frac{1}{4} C_8 \). Then

\[
P, Q \in \Sigma \quad \text{and} \quad T \in A_{\text{tors}} \quad \text{and} \quad P - Q + T \in \Theta
\]

\[
\Rightarrow \quad \hat{h}_A(P - Q + T) = \hat{h}_A(P - Q) \leq 2\hat{h}_A(P) + 2\hat{h}_A(Q)
\]

parallelogram formula,
\[
\Rightarrow \hat{h}_A(P - Q + T) \leq 4C_6
\]
from (36), since \(P, Q \in \Sigma\),

\[
\Rightarrow \hat{h}_A(P - Q + T) \leq C_8, \quad \text{since} \quad C_6 = \frac{1}{4}C_8,
\]

\[
\Rightarrow P - Q + T \in \Xi.
\]

To ease notation, we let

\[
N = \#\Sigma \quad \text{and} \quad \nu = \nu(A, \Theta, \hat{h}_A) := \max\{\#\Xi, 2\},
\]

and we let

\[
\Sigma = \{P_1, P_2, \ldots, P_N\}.
\]

We build the set \(\Sigma'\) one step at a time. We first consider the differences of \(P_1\) with the other elements of \(\Sigma\), translated by torsion points, i.e., we consider the sets

\[
P_1 - P_2 + A_{\text{tors}}, \ P_1 - P_3 + A_{\text{tors}}, \ldots, \ P_1 - P_N + A_{\text{tors}}.
\]

The implication proven earlier implies that at most \(\nu = \#\Xi\) of these sets may contain a point lying on \(\Theta\), so relabeling the elements of \(\Sigma\), we have shown that

\[
(P_1 - P_2 + A_{\text{tors}}) \cap \Theta = \emptyset,
\]

\[
(P_1 - P_3 + A_{\text{tors}}) \cap \Theta = \emptyset,
\]

\[
\vdots
\]

\[
(P_1 - P_{N-\nu} + A_{\text{tors}}) \cap \Theta = \emptyset.
\]

We next consider the differences of \(P_2\) with the higher-indexed elements of \(\Sigma\), again translated by torsion points,

\[
P_2 - P_3 + A_{\text{tors}}, \ P_2 - P_4 + A_{\text{tors}}, \ldots, \ P_2 - P_{N-\nu} + A_{\text{tors}}.
\]

As in the previous step, at most \(\nu\) of these sets contains a point lying on \(\Theta\), so relabeling again, we have shown that

\[
(P_2 - P_3 + A_{\text{tors}}) \cap \Theta = \emptyset, \ldots, \ (P_2 - P_{N-2\nu} + A_{\text{tors}}) \cap \Theta = \emptyset.
\]

Continuing in this fashion, at the \(k\)th step (until we run out of points in \(\Sigma\)), we will have shown that

\[
(P_k - P_{k+1} + A_{\text{tors}}) \cap \Theta = \emptyset, \ldots, \ (P_k - P_{N-k\nu} + A_{\text{tors}}) \cap \Theta = \emptyset.
\]

This works as long as

\[
N - k\nu > k, \quad \text{and thus as long as} \quad k < \frac{N}{\nu + 1}.
\]
Since $\nu \geq 2$ by assumption, we may certainly run the above algorithm until $k = \lceil N/2\nu \rceil$. Then by construction the set

$$\Sigma' = \{P_1, P_2, \ldots, P_k\}$$

has the property that

$$(P_i - P_j + A_{\text{tors}}) \cap \Theta = \emptyset \quad \text{for all } 1 \leq i < j \leq k,$$

and the size of the set $\Sigma'$ satisfies

$$\#\Sigma' \geq \left\lceil \frac{N}{2\nu} \right\rceil \geq \frac{1}{2\nu}\#\Sigma.$$ 

This completes the proof of Lemma 8.1 with $C_7 = 1/2\nu$. □

9. A Lehmer-Type Height Bound for Abelian Surfaces

In this section we prove an unconditional, albeit somewhat technical, lower bound for average values of the Bernoulli part of the canonical height. We also prove a corollary giving an exponent 2 Lehmer-type lower bound for the canonical height that is conditional on the assumption that the average of the intersection part of the canonical height is at least as large as the local-global constant $\kappa_\Theta$ appearing in Theorem 3.1(h).

**Theorem 9.1.** We set the following notation:

- $k$ an algebraically closed field of characteristic 0.
- $K/k$ a 1-dimensional function field.
- $(A, \Theta)/K$ an abelian variety $A$ defined over $K$ with an irreducible effective symmetric principal polarization $\Theta \in \text{Div}_K(A)$.
- $h_{A, \Theta}$ the canonical height on $A$ for the divisor $\Theta$.
- $h_{A, \Theta}^B$ the Bernoulli part of the canonical height on $A$ for the divisor $\Theta$; see Definition 3.3.

Assume that for every place $v$ of $K$, the abelian variety $A$ has either potential good reduction at $v$ or totally multiplicative reduction at $v$, and that $A$ has at least one place of multiplicative reduction.\(^{18}\) There are constants $C_9, C_{10}, C_{11}, C_{12} > 0$ and an integer $d \geq 1$ that depend only on $A/K$ so that the following holds:

\(^{18}\)For ease of exposition, we have excluded abelian surfaces having partial multiplicative reduction (surface with fibers $A_v = \mathcal{E} \times \mathbb{G}_m$ where $\mathcal{E}$ is an elliptic curve), although we expect that these cases could be handled similarly. We also note that although the assumption that $A$ have at least one place of potential multiplicative reduction is required for our proof, it is a relatively weak assumption. For example, if $A/K$ has everywhere good reduction and is not isotrivial, then it necessarily has a non-simple fiber $A_v$, i.e., a fiber that is isogenous to a product of elliptic curves.
For all finite extensions $L/K$ and all sets of points $\Sigma \subseteq \{ P \in A(L) : \hat{h}_{A,\Theta}(P) \leq C_9 \}$, there is a subset $\Sigma_0 \subseteq \Sigma$ having the following three properties:

- $\#\Sigma_0 \geq C_{10} \cdot \#\Sigma$ (38)
- $P - Q + T \notin |\Theta|$ for all distinct $P, Q \in \Sigma_0$ and all $T \in A_{\text{tors}}$. (39)
- $\mathcal{A}_{\Sigma_0} \mathcal{A}_{\Sigma_0} \hat{h}^g_{A,\Theta}(P - Q + T) \geq \frac{C_{11}}{[L : K]^{2/3}} - \frac{C_{12}}{\#\Sigma}$. (40)

**Corollary 9.2.** With notation as in Theorem 9.1, suppose that for every finite $L/K$ and every set of points $\Sigma$ satisfying (37), there is a subset $\Sigma_0 \subseteq \Sigma$ satisfying (38), (39), (40), and also

- $\mathcal{A}_{\Sigma_0} \mathcal{A}_{\Sigma_0} \hat{h}^g_{A,\Theta}(P - Q + T) \geq \kappa_{\Theta}$, (41)

where $\kappa_{\Theta}$ is the constant appearing in Theorem 3.1(h). Then every non-torsion $P \in A(K)$ satisfies

$$\hat{h}_{A,\Theta}(P) \geq \frac{C_{13}}{[K(P) : K]^2}.$$

**Remark 9.3.** The assumption (41) in Corollary 9.2 says roughly that (on average) the intersection part of the local heights, by itself, is sufficient to compensate for the difference between the canonical height and the sum of the local heights. It is unclear to the authors whether this is likely to be true, but we have included it in order to explain how the somewhat technical estimate in Theorem 9.1 can be incorporated into the proof of a Lehmer-type estimate, as was done unconditionally for elliptic curves in [16].

**Proof of Theorem 9.1.** We first replace $K$ by a finite extension over which $A$ has everywhere good or totally multiplicative reduction, which may require some adjustment in the constants. We let

$$n = [L : K].$$

As in the statement of the theorem, all of the constants may depend on $A/K$, but they are independent of $L$, $n$ and $P \in A(L)$. We let

$$S = \{ v \in M_K : A \text{ has bad reduction at } v \}.$$

\textsuperscript{19}We note that (39) ensures that $\hat{h}^g_{A,\Theta}$ is well-defined at all of the $P - Q - T$ points under consideration.
For each \( v \in S \) we fix a uniformization

\[
\mathbb{G}_m^2(\bar{K}_v) \to A(\bar{K}_v)
\]

with kernel spanned (multiplicatively) by the columns of the matrix

\[
q_v = \begin{pmatrix} q_{v,11} & q_{v,12} \\ q_{v,21} & q_{v,22} \end{pmatrix}
\]

whose associated \( \theta \)-function has divisor that is a translation of \( \Theta \) be a 2-torsion point. The valuation matrix

\[
Q_v = v(q_v) = \begin{pmatrix} v(q_{v,11}) & v(q_{v,12}) \\ v(q_{v,21}) & v(q_{v,22}) \end{pmatrix} = \begin{pmatrix} a_v & b_v \\ b_v & c_v \end{pmatrix}
\]

is symmetric and positive-definite. As usual, we let

\[
\alpha_v = a_v - b_v \quad \text{and} \quad \gamma_v = c_v - b_v.
\]

After a change of basis as described in Lemma 4.1, we may assume that the triple is normalized, and thus that

\[
D_v = a_v c_v - b_v^2 > 0 \quad \text{and} \quad 0 \leq 2b_v \leq a_v \leq c_v.
\]

To ease notation, we define two functions on \( \mathbb{Z}^3 \), where we note that the expressions \( \xi(a, b, c) \) and \( \Delta(a, b, c) \) are the quantities appearing in both Lemma 7.1 and Lemma 7.2:

\[
\Delta(a, b, c) = \frac{D}{\gcd(a, b, c)^2},
\]

\[
(42)
\]

\[
\xi(a, b, c) = \frac{\alpha \gcd(c, b)^2 + \gamma \gcd(a, b)^2 + b \gcd(\alpha, \gamma)^2}{D}.
\]

(43)

For the proof of Theorem 9.1 it is crucial to observe that these functions satisfy the homogeneity formulas

\[
\xi(ea, eb, ec) = e \xi(a, b, c) \quad \text{and} \quad \Delta(ea, eb, ec) = \Delta(a, b, c),
\]

since these homogeneity properties allow us to control the height bounds as for ramified extensions \( L_w/K_v \).

For \( w \in M_L \) with \( w \mid v \), we denote the ramification index of \( w/v \) by \( e_w \), so \( w|K = e_w v \). In particular, the valuations of the multiplicative periods of \( A \) are multiplied by \( e_w \) when we move from \( K \) to \( L \). Thus for places \( v \) of bad reduction, we have

\[
\begin{align*}
 a_w &= e_w a_v, & b_w &= e_w b_v, & c_w &= e_w c_v, \\
 \alpha_w &= e_w \alpha_v, & \gamma_w &= e_w \gamma_v, \\
 D_w &= a_w c_w - b_w^2 = e_w^2 D_v, \\
 \Delta(a_w, b_w, c_w) &= \Delta(a_v, b_v, c_v), \\
 \xi(a_w, b_w, c_w) &= e_w \xi(a_v, b_v, c_v).
\end{align*}
\]

(44)
We define the integer \(d\) by the formula
\[
d = 2 \text{LCM}\{\Delta(a_v, b_v, c_v) : v \in S\}.
\]
We note that \(d\) depends only on \(A/K\), i.e., it is independent of the extension field \(L/K\). We may thus replace \(L\) with the compositum of \(L\) and \(K(A[d])\), at the potential cost of multiplying \(n = [L : K]\) be up to \(d^4\). Since \(d\) depends only on \(A/K\), this requires only an adjustment of various constants. We henceforth assume that
\[
A[d] \subset A(L).
\]
We choose a place \(v_0 \in M_K\) such that the fiber of the Néron model of \(A\) is a torus, i.e., \(A_v(k) \cong \mathbb{G}_m^2(k)\). (By assumption, there is at least one such place.) Then among the \(w \in M_L\) lying over \(v_0\), we choose \(w_0\) to have largest ramification index, i.e.,
\[
e_{w_0} = \max\{e_w : w \in M_L, w \mid v\}.
\]
We also let
\[
M_{A/K}^{\text{bad}} = \{v \in M_K : A\text{ has bad reduction at } v\},
\]
and similarly for \(M_{A/L}^{\text{bad}}\).

Let \(\Sigma\) be a set satisfying (37). We start by applying Lemma 7.2 to \(\Sigma \subset A(L) \subset A(L_{w_0})\) to find a subset \(\Sigma' \subseteq \Sigma\) satisfying
\[
\#\Sigma' \geq 6^{-3}\#\Sigma \quad \text{(45)}
\]
and such that for all distinct \(P, Q \in \Sigma'\) we have
\[
_{\Theta, w_0} \hat{\lambda}_d(P - Q) \geq \frac{\xi(a_{w_0}, b_{w_0}, c_{w_0})}{144d^2} = \frac{e_{w_0} \xi(a_{w_0}, b_{w_0}, c_{w_0})}{144d^2}. \quad \text{(46)}
\]
We next apply Lemma 8.1 to the set \(\Sigma'\) to find a subset \(\Sigma_0 \subseteq \Sigma'\) satisfying
\[
N := \#\Sigma_0 \geq C_7 \cdot \#\Sigma' \quad \text{(47)}
\]
and
\[
P - Q + T \notin |\Theta| \quad \text{for all distinct } P, Q \in \Sigma_0 \text{ and all } T \in A_{\text{tors}}. \quad \text{(48)}
\]
We now estimate the double average (40) for the set \(\Sigma_0\) and the integer \(d\). We note that (48) ensures that the points \(P - Q + T\) appearing

\[\text{[20] If we only want the lower bound on the Bernoulli part of the height, it is not necessary to use Lemma 8.1, since the Bernoulli part of the height is defined on all of } A. \text{ However, any application to the global height will need to also include the intersection part of the height, which is not defined on the support of } \Theta.\]
in this calculation do not lie on the divisor $\Theta$, and thus the local heights are well-defined at all such points. Thus

$$\mathcal{A}_0 \mathcal{A}_0 \hat{h}_{A,\Theta}^B(P - Q + T)$$

$$= \mathcal{A}_0 \mathcal{A}_0 \sum_{P \neq Q} \frac{1}{n} \lambda_{\Theta, w}^B(P - Q + T)$$

$$= \sum_{w \in M_{A/L} \backslash v_0} \frac{1}{n} \mathcal{A}_0 \mathcal{A}_0 \lambda_{\Theta, w}^B(P - Q + T). \quad (49)$$

We split the sum in (49) into three pieces:

1. For the absolute value $w_0$, we use the lower bound from Lemma 7.2.
2. For the absolute values $w$ dividing $v_0$ that are not equal to $w_0$, we use the lower bound provided by the full strength of Lemma 7.1.
3. For the absolute values $w$ with $w \in M_{A/L} \backslash v_0$ that do not divide $v_0$, we again use Lemma 7.1, but we discard the positive contribution coming from the $1/D^2$ terms.

Carrying out these three estimates yields the following:

1. $\frac{1}{n} \mathcal{A}_0 \mathcal{A}_0 \lambda_{\Theta, w_0}^B(P - Q + T)$

$$\geq \frac{1}{n} \frac{e_{w_0} \xi(a_{w_0}, b_{w_0}, c_{w_0})}{144d^2} \quad \text{from (46),}$$

$$= C_{14} \cdot \frac{e_{w_0}}{n}. \quad (50)$$

2. $\sum_{w \in M_{A/L} \backslash v_0} \frac{1}{n} \mathcal{A}_0 \mathcal{A}_0 \lambda_{\Theta, w}^B(P - Q + T)$

$$\geq \sum_{w \in M_{A/L} \backslash v_0} \frac{1}{n} \frac{1}{24d^2} \left( \frac{\alpha_w + \gamma_w + b_w}{D_w} - \frac{\xi(a_w, b_w, c_w)}{N - 1} \right)$$

applying Lemma 7.1 to $\Sigma_0$ and $w$,

$$= \frac{1}{24nd^2} \sum_{w \in M_{A/L} \backslash v_0} \frac{1}{e_w} \left( \frac{e_w(\alpha_{w_0} + \gamma_{w_0} + b_{w_0})}{e_w^2D_{w_0}} - \frac{e_w v(a_{w_0}, b_{w_0}, c_{w_0})}{N - 1} \right)$$

using the homogeneity formulas (44),

$$= \frac{\alpha_{w_0} + \gamma_{w_0} + b_{w_0}}{24nd^2D_{w_0}} \sum_{w \in M_{A/L} \backslash v_0} \frac{1}{e_w} - \frac{\xi(a_{w_0}, b_{w_0}, c_{w_0})}{24nd^2(N - 1)} \sum_{w \in M_{A/L} \backslash v_0} e_w.$$
\[
\frac{\alpha v_0 + \gamma v_0 + b v_0}{24nd^2D v_0} \left( \sum_{w \in M_{\text{bad}}^{A/L}} \frac{1}{e_w} \right) - \frac{\zeta(a v_0, b v_0, c v_0)(n - e_{w_0})}{24nd^2(N - 1)}
\]

since in \(\sum_{w \mid v} e_w = n\) for all \(v\),

\[
\geq \frac{C_{15}}{n} \sum_{w \in M_{\text{bad}}^{A/L}} \frac{1}{e_w} - \frac{C_{16}}{(N - 1)}. \tag{51}
\]

\[
\sum_{w \in M_{\text{bad}}^{A/L}} n_PQ \in \Sigma_0 T \in A[d] \frac{1}{n} \frac{1}{e_w} (P - Q + T)
\]

\[
\geq \frac{1}{n} \sum_{w \in M_{\text{bad}}^{A/L}} \frac{1}{24d^2} \left( -\frac{\xi(a_w, b_w, c_w)}{N - 1} \right)
\]

applying Lemma 7.1 to \(\Sigma_0\) and \(w\),

\[
= \frac{1}{n} \sum_{w \in M_{\text{bad}}^{A/L}} \frac{1}{24d^2} \left( -\frac{e_w \xi(a_w, b_w, c_w)}{N - 1} \right)
\]

using the homogeneity formulas (44),

\[
= -\frac{1}{24d^2n} \sum_{v \in M_{\text{bad}}^{A/K}} \left( \frac{\xi(a_v, b_v, c_v)}{N - 1} \right) \sum_{w \in M_L} e_w
\]

\[
= -\frac{1}{24d^2(N - 1)} \sum_{v \in S(A/K)} \xi(a_v, b_v, c_v)
\]

\[
= -\frac{C_{17}}{N - 1}. \tag{52}
\]

Substituting the sum of the three estimates (50), (51), (52) into (49), we find that

\[
\sum_{P \neq Q} \frac{1}{n} \left[ C_{14} e_{w_0} + C_{15} \sum_{w \in M_L} \frac{1}{e_w} \right] - \frac{C_{16} + C_{17}}{N - 1}. \tag{53}
\]

Since \(e_{w_0} = \max \{e_w : w \mid v_0\\}\) and \(\sum_{w \mid v_0} e_w = n\),

we can apply Lemma 9.4 to the quantity in braces in (53) to obtain the following lower bound, with newly relabeled constants depending...
on $A/K$ and where we have used (45) and (47) to estimate $N = \#\Sigma_0$ in terms of $\#\Sigma$.

$$
\mathcal{M}_p \mathcal{M}_q \mathcal{M}_{P \neq Q, T \in A[\mathfrak{d}]} \hat{h}_B^\Sigma(P - Q + T) \geq \frac{1}{n} \cdot C_{21} n^{1/3} - \frac{C_{22}}{N - 1}
$$

$$
\geq \frac{C_{11}}{n^{2/3}} - \frac{C_{12}}{\#\Sigma}.
$$

This completes the proof of Theorem 9.1. □

**Proof of Corollary 9.2.** Let $P_0 \in A(\bar{K})$ be a non-torsion point, and to ease notation, let

$$
L = K(P_0) \quad \text{and} \quad n = [L : K].
$$

We take $M$ to be the largest integer satisfying

$$
M^2 \leq \frac{C_9}{\hat{h}_{A,\Theta}(P_0)},
$$

where $C_9$ is the constant appearing in (37). We consider the set of points

$$
\Sigma = \{mP_0 : 0 \leq m \leq M - 1\} \subset \{P \in A(L) : \hat{h}_{A,\Theta}(P) \leq C_9\},
$$

where the inclusion follows from $\hat{h}_{A,D}(mP_0) = m^2 \hat{h}_{A,D}(P_0)$ and our choice of $M$.

Then, according to (38), (40), and (41), we can find a subset $\Sigma_0 \subseteq \Sigma$ with $\#\Sigma_0 \geq C_{10} \#\Sigma = C_{10} M$ that satisfies

$$
\mathcal{M}_p \mathcal{M}_q \mathcal{M}_{P \neq Q, T \in A[\mathfrak{d}]} \hat{h}_{A,\Theta}^\Sigma(P - Q + T) \geq \frac{C_{11}}{n^{2/3}} - \frac{C_{12}}{M}.
$$

(55)

$$
\mathcal{M}_p \mathcal{M}_q \mathcal{M}_{P \neq Q, T \in A[\mathfrak{d}]} \hat{h}_I^\Sigma(P - Q + T) \geq \kappa_{\Theta}.
$$

(56)

Proposition 10 says that

$$
\hat{h}_{A,\Theta} = \hat{h}_{A,\Theta}^I + \hat{h}_{A,\Theta}^R - \kappa_{\Theta}.
$$

We remark that in order to apply Lemma 9.4, the integer $n$ must satisfy $n^2 \geq C_{15}/C_{14}$. There is no harm in our making this assumption, since these constants are given explicitly by

$$
C_{15} = \frac{\alpha_{\tau \omega} + 24D_{\tau \omega}}{24D_{\tau \omega}^2} \quad \text{and} \quad C_{14} = \frac{\xi(a_{\tau \omega}, b_{\tau \omega}, c_{\tau \omega})}{144d^2} \geq \frac{\alpha_{\tau \omega} + \gamma_{\tau \omega} + 24D_{\tau \omega}}{144D_{\tau \omega}},
$$

and thus $C_{15}/C_{14} \leq 6$. Hence it suffices to assume that $n \geq 3$. 21
so adding (55) to (56) yields

\[
\text{Avg}_{P,Q \in \Sigma_0} \text{Avg}_{P \neq Q} \hat{h}_{A,\Theta}(P - Q) \geq \frac{C_{11}}{n^{2/3}} - \frac{C_{12}}{M}.
\]

But for any points \(P, Q \in \Sigma\) and for any torsion point \(T \in A_{\text{tors}}\), we have

\[
\hat{h}_{A,\Theta}(P - Q + T) = \hat{h}_{A,\Theta}(P - Q)
\leq 2\hat{h}_{A,\Theta}(P) + 2\hat{h}_{A,\Theta}(Q)
\leq 4 \max_{P \in \Sigma} \hat{h}_{A,\Theta}(P)
\leq 4 \max_{0 \leq m < M} \hat{h}_{A,\Theta}(mP_0)
\leq M^2 \hat{h}_{A,\Theta}(P_0).
\]

Hence

\[
\text{Avg}_{P,Q \in \Sigma_0} \text{Avg}_{P \neq Q} \hat{h}_{A,\Theta}(P - Q + T) \leq M^2 \hat{h}_{A,\Theta}(P_0).
\]

Combining (57) and (58) yields

\[
M^2 \hat{h}_{A,\Theta}(P_0) \geq \frac{C_{11}}{n^{2/3}} - \frac{C_{12}}{M}.
\]

Setting \(M\) to be the smallest integer satisfying

\[
M \geq \frac{2C_{12}n^{2/3}}{C_{11}}
\]

yields (after adjusting constants)

\[
n^{4/3} \hat{h}_{A,\Theta}(P_0) \geq \frac{C_{13}}{n^{2/3}}.
\]

This completes the proof of Corollary 9.2 provided that we can justify choosing \(M\) to satisfy (59), since we earlier in (54) assumed that \(M\) satisfies an upper bound. In other words, we need to check that there is an integer \(M\) in the interval

\[
\frac{2C_{12}n^{2/3}}{C_{11}} \leq M \leq \sqrt[3]{\frac{C_9}{\hat{h}_{A,\Theta}(P_0)}}.
\]

But if there is no such \(M\), then we find that

\[
\sqrt[3]{\frac{C_9}{\hat{h}_{A,\Theta}(P_0)}} \leq \frac{2C_{12}n^{2/3}}{C_{11}} + 1,
\]
and squaring both sides and adjusting constants, we see that
\[ \hat{h}_{A, \Theta}(P_0) \geq \frac{C_{23}}{n^{4/3}}, \]
which is an even stronger inequality than the one that we are trying to prove. \qed

The following is a more precise and fully explicated version of [16, Lemma 3.1].

**Lemma 9.4.** Let \( \alpha, \beta, n > 0 \) be positive real numbers satisfying
\[ n^2 \geq \beta/\alpha, \]
and let \( e_0, \ldots, e_r > 0 \) be positive real numbers satisfying
\[ e_0 = \max\{e_0, \ldots, e_r\} \quad \text{and} \quad n = e_0 + \cdots + e_r. \]
Then
\[ \alpha e_0 + \beta \sum_{i=1}^{r} \frac{1}{e_i} \geq (\alpha^2 \beta n)^{\frac{1}{3}}. \]

**Proof.** Since \( e_0 \) is the largest of the \( e_i \) and \( n \) is the sum of the \( e_i \), we can estimate
\[ e_0 \geq \frac{e_0 + \cdots + e_r}{r + 1} = \frac{n}{r + 1}. \]

We compute
\[ r^2 = \left( \sum_{i=1}^{r} e_i^{1/2} \cdot e_i^{-1/2} \right)^2 \]
\[ \leq \left( \sum_{i=1}^{r} e_i \right) \left( \sum_{i=1}^{r} e_i^{-1} \right) \quad \text{Cauchy-Schwartz inequality}, \]
\[ = (n - e_0) \left( \sum_{i=1}^{r} e_i^{-1} \right) \quad \text{since} \ e_0 + \cdots + e_r = n, \]
\[ \leq \frac{rn}{r + 1} \left( \sum_{i=1}^{r} e_i^{-1} \right) \quad \text{using (62).} \]

We use this estimate to bound the left-hand side of (61) as
\[ \alpha e_0 + \beta \sum_{i=1}^{r} \frac{1}{e_i} \geq \frac{\alpha n}{r + 1} + \frac{\beta r^2 + r}{n} \quad \text{using (62) and (63)}, \]
\[ \geq \inf_{t > 0} \left\{ \frac{\alpha n}{t + 1} + \frac{\beta (t^2 + t)}{n} \right\} \]
\[
\begin{align*}
&= \inf_{x > 1} \left\{ \frac{\alpha n}{x} + \frac{\beta}{n} (x^2 - x) \right\} \quad \text{setting } x = t + 1, \\
&= (\alpha^2 \beta n)^{1/3} \inf_{u > \gamma} \left\{ \frac{1}{u} + u^2 - \gamma u \right\} \\
&\quad \text{setting } \gamma = \left( \frac{\beta}{\alpha n^2} \right)^{1/3} \text{ and } u = \gamma x.
\end{align*}
\]

To ease notation, we let
\[
f(\gamma, u) = u^{-1} + u^2 - \gamma u.
\]

The fact that
\[
\frac{d^2}{du^2} (u^{-1} + u^2 - \gamma u) = 2u^{-3} + 2 > 0 \quad \text{for all } u > 0
\]
shows that \( f(\gamma, u) \) has at most one minimum on the half-line \( u > 0 \), and then the fact that \( f(\gamma, u) \to \infty \) as \( u \to 0^+ \) and as \( u \to \infty \) shows that it has a unique minimum. We thus get a well-defined function
\[
F(w) = \inf_{u > 0} f(w, u) = \inf_{u > 0} \{ u^{-1} + u^2 - wu \} \quad \text{for } w \in \mathbb{R}.
\]

We claim that \( F(w) \) is a strictly decreasing function. To see why, we note that our earlier discussion shows that
\[
F(w) = f(w, U(w)) = U(w)^{-1} + U(w)^2 - wU(w),
\]
where \( u = U(w) \) is the unique real solution to the equation
\[
\frac{\partial f}{\partial u}(w, u) = -u^{-2} + 2u - w = 0.
\]
Hence
\[
\frac{dF}{dw} = \frac{d}{dw} f(w, U(w))
\]
\[
= \frac{\partial f}{\partial w}(w, U(w)) + \frac{\partial f}{\partial u}(w, U(w)) \cdot \frac{dU}{dw}(w)
\]
\[
= -U(w) < 0.
\]

Returning to our earlier calculation and using the assumption (60) that \( \gamma \leq 1 \), we find that
\[
\alpha e_0 + \beta \sum_{i=1}^{r} \frac{1}{e_i} \geq (\alpha^2 \beta n)^{1/3} \inf_{u > \gamma} \{ u^{-1} + u^2 - \gamma u \} \quad \text{from (64)},
\]
\[
\geq (\alpha^2 \beta n)^{1/3} F(\gamma) \quad \text{by definition of } F(w),
\]
\[
\geq (\alpha^2 \beta n)^{1/3} F(1) \quad \text{for all } 0 \leq \gamma \leq 1, \text{ since } F(w)
\]
\text{is a decreasing function,}
\[
(\alpha^2 \beta n)^{1/3} \quad \text{since it is easy to compute } F(1) = 1.
\]

This completes the proof of Lemma 9.4. \qed

Acknowledgements. The authors would like to thank Dan Abramovich, Matt Baker, and David Grant for their helpful advice.

References

[1] Francesco Amoroso and Roberto Dvornicich. A lower bound for the height in abelian extensions. *J. Number Theory*, 80(2):260–272, 2000.
[2] M. Anderson and David W. Masser. Lower bounds for heights on elliptic curves. *Math. Z.*, 174(1):23–34, 1980.
[3] Matthew H. Baker. Lower bounds for the canonical height on elliptic curves over abelian extensions. *Int. Math. Res. Not.*, (29):1571–1589, 2003.
[4] Matthew H. Baker and Joseph 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] P. E. Blanksby and H. L. Montgomery. Algebraic integers near the unit circle. *Acta Arith.*, 18:355–369, 1971.
[6] John L. Boxall. Une propriété des hauteurs locales de Néron-tate sur les variété abéliennes. *J. Théor. Nombres Bordeaux*, 7(1):111–119, 1995.
[7] Sinnou David. Minimations de hauteurs sur les variétés abéliennes. *Bull. Soc. Math. France*, 121(4):509–544, 1993.
[8] Sinnou David and Marc Hindry. Minoramtion de la hauteur de Néron-Tate sur les variétés abéliennes de type C. M. *J. Reine Angew. Math.*, 529:1–74, 2000.
[9] Sinnou David and Patrice Philippon. Minorations des hauteurs normalisées des sous-variétés de variétés abéliennes. In *Number theory (Tiruchirapalli, 1996)*, volume 210 of *Contemp. Math.*, pages 333–364. Amer. Math. Soc., Providence, RI, 1998.
[10] Sinnou David and Patrice Philippon. Minorations des hautesurs normalisées des sous-variétés de variétés abéliennes. II. *Comment. Math. Helv.*, 77(4):639–700, 2002.
[11] Sinnou David and Patrice Philippon. Minorations des hauteurs normalisées des sous-variétés de puissances des courbes elliptiques. *Int. Math. Res. Pap. IMRP*, (3):Art. ID rpm006, 113, 2007.
[12] E. Dobrowolski. On a question of Lehmer and the number of irreducible factors of a polynomial. *Acta Arith.*, 34:391–401, 1979.
[13] Aurélien Galateau and Valéry Mahé. Some consequences of Masser’s counting theorem on elliptic curves. *Math. Z.*, 285(1-2):613–629, 2017.
[14] Marc 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.
[15] Marc Hindry and Joseph H. Silverman. The canonical height and integral points on elliptic curves. *Invent. Math.*, 93(2):419–450, 1988.
[16] Marc Hindry and Joseph 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.
Appendix A. Verification of some basic formulas

Proposition A.1. Let $n \in \mathbb{Z}^g$ and $u \in \mathbb{C}_m^g(K_v)$.

(a) $\Theta(u \cdot (q \ast 2n), q) = (\check{t}n \ast q \ast n)^{-1}(\check{t}n \ast u)^{-1}\Theta(u, q)$.

(b) $v\left(\Theta(u \cdot (q \ast 2n), q)\right) = v\left(\Theta(u, q)\right) - \check{t}nQn - \check{t}nv(u)$. 

[23] Joseph H. Silverman. Lower bounds for height functions. Duke Math. J., 51(2):395–403, 1984.
[24] Joseph H. Silverman. A lower bound for the canonical height on elliptic curves over abelian extensions. J. Number Theory, 104(2):353–372, 2004.
[25] C. J. Smyth. On the product of the conjugates outside the unit circle of an algebraic integer. Bull. London Math. Soc., 3:169–175, 1971.
[26] Cameron L. Stewart. Algebraic integers whose conjugates lie near the unit circle. Bull. Soc. Math. France, 106(2):169–176, 1978.
[27] Emmanuel Ullmo. Positivité et discrétion des points algébriques des courbes. Ann. of Math. (2), 147(1):167–179, 1998.
[28] Jean-Louis Verger-Gaugry. A survey on the conjecture of Lehmer and the conjecture of Schinzel–Zassenhaus, 2019. hal.archives-ouvertes.fr/hal-02315014
[29] Annette Werner. Local heights on Mumford curves. Math. Ann., 306(4):819–831, 1996.
[30] Annette Werner. Local heights on abelian varieties with split multiplicative reduction. Compositio Math., 107(3):289–317, 1997.
[31] Annette Werner. Local heights on abelian varieties and rigid analytic uniformization. Doc. Math., 3:301–319, 1998.
[32] Shou-Wu Zhang. Equidistribution of small points on abelian varieties. Ann. of Math. (2), 147(1):159–165, 1998.
[33] Shouwu Zhang. Lower bounds for heights on elliptic curves, June 1989. unpublished
Proof. (a) We compute

\[
\Theta(u \cdot (q \star 2n), q) = \sum_{m \in \mathbb{Z}^g} (t^m \star q \star m)(t^m \star (u \cdot (q \star 2n)))
\]
\[
= \sum_{m \in \mathbb{Z}^g} (t^m \star q \star m)(t^m \star u)(t^m \star q \star 2n)
\]
\[
= \sum_{m \in \mathbb{Z}^g} (t^m \star q \star (m + 2n))(t^m \star u)
\]
\[
= \sum_{m \in \mathbb{Z}^g} (t^m \star q \star (m + n))(t^m \star u)
\]
\[
= (t^n \star q \star n)^{-1}(t^n \star u)^{-1}\Theta(u, q).
\]

(b) We have elementary formulas

\[
v(q \star n) = Qn, \quad v(t^n \star q \star n) = t^n Qn, \quad v(t^n \star u) = t^n v(u). \quad (65)
\]

We verify the first of these and leave the others to the reader.

\[
v(q \star n) = \begin{pmatrix} v(q_{11}^n \cdots q_{1g}^n) \\ \vdots \\ v(q_{g1}^n \cdots q_{gg}^n) \end{pmatrix} = \sum_{j=1}^g n_j \begin{pmatrix} v(q_{1j}) \\ \vdots \\ v(q_{gj}) \end{pmatrix} = Qn.
\]

Then applying \(v\) to the formula in (a) gives the stated result. \(\square\)

**Proposition A.2.** The function

\[
\Lambda(\cdot, q) : \mathbb{G}_m^g(K_v) \rightarrow \mathbb{R},
\]

\[
\Lambda(u, q) = v(\Theta(u, q)) + \frac{1}{4}t^v(u)Q^{-1}v(u),
\]

is \(\Omega\)-invariant, and hence descends to a function

\[
\Lambda(\cdot, q) : A(K_v) \cong \mathbb{G}_m^g(K_v)/\Omega \rightarrow \mathbb{R}.
\]

**Proof.** We use the elementary formulas (65) to compute what happens when we translate \(u\) by an element of the lattice.

\[
\Lambda(u \cdot (q \star 2n), q) - \Lambda(u, q)
\]
\[
= \left\{ v(\Theta(u \cdot (q \star 2n), q)) + \frac{1}{4}t^v(u \cdot (q \star 2n))Q^{-1}v(u \cdot (q \star 2n)) \right\}
\]
\[
- \left\{ v(\Theta(u, q)) + \frac{1}{4}t^v(u)Q^{-1}v(u) \right\}
\]
\[
\begin{align*}
&= \left\{ v(\Theta(u \cdot (q \star 2n), q)) - v(\Theta(u, q)) \right\} \\
&+ \left\{ \frac{1}{4} t(v(u \cdot (q \star 2n))Q^{-1}v(u \cdot (q \star 2n)) - \frac{1}{4} t(v(u)Q^{-1}v(u)) \right\} \\
&= v \left( (n \star q \star n)^{-1}(n \star u)^{-1} \right) \\
&\quad + \frac{1}{4} \left\{ t(v(q \star 2n)Q^{-1}v(u) + t(v(u))Q^{-1}v(q \star 2n) \\
&\quad + t(v(q \star 2n))Q^{-1}v(q \star 2n) \right\} \\
&= \left\{ -(nQn - nQn) \right\} \\
&\quad + \frac{1}{4} \left\{ t(2Qn)Q^{-1}v(u) + t(v(u))Q^{-1}2Qn + t(2Qn))Q^{-1}(2Qn) \right\} \\
&= 0 \quad \text{since } tQ = Q.
\end{align*}
\]