Algebra & Number Theory

Volume 15
2021
No. 8

Roth's theorem over arithmetic function fields

Paul Vojta

msp

ALGEBRA AND NUMBER THEORY 15:8 (2021)
https://doi.org/10.2140/ant.2021.15.1943

Roth's theorem over arithmetic function fields

Paul Vojta

Roth's theorem is extended to finitely generated field extensions of $\mathbb{Q}$, using Moriwaki's theory of heights.

In his work dating back at least to the 1970s, Serge Lang observed that many results in diophantine geometry that were true over number fields were also true for fields finitely generated over $\mathbb{Q}$. Following Moriwaki, the latter will be called *arithmetic function fields* in this paper.

Lang felt that such fields were a more natural setting for diophantine geometry; see [Lang 1974; 1986]. For example, the Mordell–Weil theorem and Faltings' theorem on the Mordell conjecture are true over arithmetic function fields — see Corollaries 4.3 and 2.2, respectively, in Chapter I of [Lang 1991].[1] Both are proved using induction on the transcendence degree, using the cases of the theorems over (classical) function fields in the inductive step. Correspondingly, Lang phrased his early conjectures on "Mordellicity" in terms of rational points over subfields of $\mathbb{C}$ finitely generated over $\mathbb{Q}$ (i.e., arithmetic function fields).

As for integral points, Siegel's theorem on integral points holds also for points integral over entire rings of finite type over $\mathbb{Z}$ [Lang 1960, Theorem 4]; see also [Lang 1991, Chapter IX, Theorem 3.1] and Corollary 4.11, below. In that spirit, Lang conjectured that results on integral points over number rings should extend to integral points over entire rings finitely generated over $\mathbb{Z}$; see [Lang 1974].

A weaker version of the Dirichlet unit theorem is also true (it gives only an inequality for the rank, since finiteness of the class group does not hold for arithmetic function fields). One can then extend the Mordell–Weil theorem to include integral points on semiabelian varieties over arithmetic function fields. This is done in the usual way.

More recently, Moriwaki [2000] formulated a theory of heights over arithmetic function fields, and showed that they have a many of the standard properties, including independence up to $O(1)$ of the choices made, Northcott's theorem, and canonical heights on abelian varieties.

Moriwaki's work opened the door for theorems on diophantine approximation to be extended to arithmetic function fields.

This paper takes a first step in this direction, by extending Roth's theorem to arithmetic function fields. This uses Moriwaki's theory of height functions and an obvious extension of his work to Weil functions (local heights). As a consequence, it follows that arithmetic function fields are quite close to number

MSC2010: primary 11J68; secondary 11J97, 14G40.
Keywords: Diophantine approximation, arithmetic function field, Roth's theorem, Thue–Siegel method.

[1] Lang apparently forgot to state the necessary assumption that X has genus ≥ 2.

fields, in the sense that Roth's theorem can be proved using extensions of the standard proof over number fields, as opposed to arguments that reduce to the number field case.

This paper was suggested by a paper of Rastegar [2015] (as it turns out, though, his theorem can be proved more easily without using results of this paper).

Schmidt's subspace theorem should extend to arithmetic function fields using the same methods. That will be the subject of future work. I thank one of the referees for pointing me in this direction.

The Masser–Oesterlé abc conjecture should also extend to arithmetic function fields. (A proof of the abc conjecture has been proposed by Mochizuki, but it has not attained wide acceptance yet.) Also, I conjecture that Conjectures 15.6, 23.4, 25.1, 25.3, 26.1, and 30.1 of [Vojta 2011] generalize to arithmetic function fields.

Recall that in the classical diophantine theory of function fields, the function field in question is the function field K of a projective variety B over a ground field F. Often F is taken to be algebraically closed; hence, following Moriwaki [2002, Section 1], we refer to such function fields as *geometric function fields.* When $\dim B > 1$, it is necessary to choose a projective embedding of B, in order to determine degrees of the prime divisors on B to be used in the product formula. When the ground field is infinite, there may be infinitely many elements of K whose height is below a fixed bound. It is true, however, that a set of such elements can belong to only a finite number of algebraic families. A similar principle applies also to Northcott's finiteness theorem (for algebraic points on a projective variety, rational over a field of bounded degree over K and of bounded height relative to an ample divisor).

Moriwaki [2002, Section 2] refers to fields finitely generated over $\mathbb{Q}$ as *arithmetic function fields.* They have this name because they have features of both function fields and number fields. An arithmetic function field K arises as the function field of an arithmetic variety; i.e., an integral scheme B, flat and projective over $\operatorname{Spec}\mathbb{Z}$. As is the case of geometric function fields, when $\dim B > 1$ it is necessary to choose metrized line sheaves on B in order to determine weights for the prime divisors on B. Unlike the case of geometric function fields, though, not all places are nonarchimedean; in fact, if $\dim B > 1$ then there are *uncountably many* archimedean places. (When $\dim B = 1$, one recovers the classical case of number fields.) For all values of $\dim B$, though, Northcott's theorem gives actual finiteness (as opposed to finiteness of algebraic families in the geometric function field case).

We recall here the statement of Roth's theorem [1955], as generalized by LeVeque [1956, Theorem 4-15], Ridout [1958], and Lang [1962] (Lang's work also covered geometric function fields).

Theorem 0.1. *Let k be a number field, and let S be a finite set of places of k. For each $v \in S$ let α_v be algebraic over k, and assume that v is extended to $\bar{k}$ in some way. Then, for all $\epsilon > 0$, the set of $\xi \in k$ satisfying the approximation condition*

$$\prod_{v\in S} \min\{1, \|\xi - \alpha_v\|_v\} \le \frac{1}{H_k(\xi)^{2+\epsilon}} \tag{0.1.1}$$

is finite.

If one extends this statement to arithmetic function fields in a straightforward way, then the resulting statement is false — see Examples 4.12 and 4.13. Instead we impose the additional condition that the set of all α_v is finite, as v varies over S (which is now in general uncountable, as described below). See Theorem 4.6. (Theorem 4.6 is actually stronger than the above — see Remark 4.7.) Theorem 4.6 reduces to Theorem 0.1 in the number field case, and is strong enough to imply Siegel's theorem on integral points (Corollary 4.11).

Actually, we give four equivalent formulations of Roth's theorem over arithmetic function fields (Theorems 4.3–4.6), and show in Proposition 4.8 that they are equivalent. Sections 5–10 contain a proof of Theorem 4.5, which then implies the other three variants.

The proof of Roth's theorem in this paper follows the same general outline as the classical proofs of Thue, Siegel, and Roth. In particular, it is ineffective (i.e., it does not give a constructive proof for the upper bound on the heights of exceptions to the main inequality). Fundamentally different proofs of Roth's theorem over geometric function fields (of characteristic 0) have been obtained by Osgood [1985] and Wang [1996], using "Nevanlinna–Kolchin systems" and Steinmetz's method in Nevanlinna theory, respectively; the latter is effective. Roth's theorem can be proved over geometric function fields of characteristic 0 by the Thue–Siegel method; see [Lang 1983]. The current paper does not add anything to this proof.

Unfortunately, Roth's theorem over arithmetic function fields does not yet imply any new applications. However, as noted above it is anticipated that Schmidt's subspace theorem will also extend to arithmetic function fields, and that result has numerous diophantine consequences whose counterparts over arithmetic function fields will be new.

The main difficulty in generalizing Roth's theorem to arithmetic function fields concerns the part of the proof often referred to as "reduction to simultaneous approximation." In that part, it is shown that it suffices to prove the theorem with the approximation condition (0.1.1) replaced by conditions

$$\min\{1, \|\xi - \alpha_v\|_v\} \le H_k(\xi)^{-\lambda_v(2+\epsilon')}$$

for each v, where $0 < \epsilon' < \epsilon$, and for each $v \in S$ a constant $\lambda_v \ge 0$ is given such that

$$\sum_{v \in S} \lambda_v = 1.$$

In addressing this difficulty, a key idea came from a proof of Wirsing [1971]. Wirsing extended Roth's theorem to approximation by rational numbers of bounded degree. In his proof the number of archimedean places was still finite, but it grew exponentially with the number of solutions to (his equivalent of) (0.1.1) under consideration. The idea was to ignore a small proportion of those places, and this is also done here. See the introduction to Section 6 for more details on this.

The paper is organized as follows. Section 1 summarizes the basic results and conventions from number theory, algebraic geometry, and Arakelov theory used in the paper. Section 2 describes the positivity properties of metrized line sheaves that are needed in the paper. Section 3 introduces Moriwaki's theory of heights for arithmetic function fields, and describes how this theory can be extended to give a theory of Weil functions (often called *local heights*). Thus, one can decompose the height into proximity and counting functions, as in Nevanlinna theory [Vojta 2011]. In Section 4, the main theorem of the paper is formally stated, in four different forms, and the four are shown to be equivalent.

Section 5 begins the main line of the proof of the theorem, by showing that it suffices to prove the theorem under some additional assumptions. Sections 6–8 give the proof of reduction to simultaneous approximation for arithmetic function fields; this is the technical core of the paper. Mostly this focuses on the archimedean places, and involves analysis of Green functions. More details are given in the introduction to Section 6. Sections 9 and 10 then conclude the proof by formulating and proving Siegel's lemma for arithmetic function fields, constructing the auxiliary polynomial, and deriving a contradiction to conclude the proof. For the latter, we use a version of Dyson's lemma [1947] due to Esnault and Viehweg [1984] instead of Roth's lemma, since the former is true for arbitrary fields of characteristic zero, and therefore needs no adaptation for arithmetic function fields.

1. Basic notation and conventions

In this paper, $\mathbb{N} = \{0, 1, 2, \dots\}$ and $\mathbb{Z}_{>0} = \{1, 2, 3, \dots\}$. Also,

$$\log^+ x = \log\max\{1, x\} \quad \text{and} \quad \log^- x = \log\min\{1, x\}.$$

Throughout this paper, the notation c_1 always refers to either a Chern form or a Chern class. The letter c with any other subscript refers to a constant — and this includes, for example, c_i when $i = 1$. Higher Chern classes do not occur in this paper.

1A. ***Algebraic geometry.*** A *variety* over a field k is an integral separated scheme of finite type over k, and a *curve* over k is a variety over k of dimension 1. A *line sheaf* is an invertible sheaf. For a point x on a scheme X, $\boldsymbol{\kappa}(x)$ denotes the residue field of x. If X is a variety or integral scheme, then $\boldsymbol{\kappa}(X)$ denotes its function field (this equals the residue field $\boldsymbol{\kappa}(\xi)$ for the generic point ξ of X).

For more details on these conventions, see [Vojta 2011].

Also, following [Moriwaki 2014], if s is a nonzero rational section of a line sheaf on an integral scheme X or a nonzero rational function on X, then $\operatorname{div}(s)$ is the associated Cartier divisor of s.

1B. ***Number theory.*** For a number field k, the subring $\mathscr{O}_k$ is its ring of integers (the integral closure of $\mathbb{Z}$ in k). The set M_k is the set of all places of k; this is the disjoint union of the sets of archimedean and nonarchimedean places of k. These are in canonical bijection with the set of injections $k \hookrightarrow \mathbb{C}$ and with the set of nonzero prime ideals in $\mathscr{O}_k$, respectively.

For each place $v \in M_k$ we define an *absolute value* $\|\cdot\|_v$, as follows:

$$\|x\|_v = \begin{cases} |\sigma(x)| & \text{if } v \text{ is archimedean and corresponds to } \sigma : k \hookrightarrow \mathbb{C}; \\ (\mathscr{O}_k : \mathfrak{p})^{-\operatorname{ord}_{\mathfrak{p}}(x)} & \text{if } v \text{ is nonarchimedean and corresponds to } \mathfrak{p} \subseteq \mathscr{O}_k. \end{cases}$$

(In the nonarchimedean case, the formula assumes $x \neq 0$; of course $\|0\|_v = 0$ for all v.) Note that two nonreal complex conjugate embeddings $\sigma, \bar{\sigma} : k \hookrightarrow \mathbb{C}$ are regarded as different places but give rise to the same absolute value. This is the usual convention in Arakelov theory. These absolute values satisfy the *product formula*

$$\prod_{v \in M_k} \|x\|_v = 1 \quad \text{for all } x \in k^*. \tag{1.1}$$

Heights are always taken to be logarithmic but not absolute. The reason for the latter is that, for a general function field K (either arithmetic or geometric) there is no canonical choice of "base field" to play the role of $\mathbb{Q}$, for which K is a finite extension (other than K itself).

As a specific example, the height of a point $P \in \mathbb{P}^n(k)$ with homogeneous coordinates $[x_0 : x_1 : \cdots : x_n]$ is

$$h_k(P) = \sum_{v \in M_k} \log \max\{\|x_0\|_v, \ldots, \|x_n\|_v\}.$$

For more information on the basic properties of heights, see [Hindry and Silverman 2000, Part B] or [Lang 1983, Ch. 4].

1C. ***Complex analytic spaces.*** A *complex analytic space*, or just *complex space*, is as defined in [Hartshorne 1977, Appendix B]. Examples include X^{an}, where X is a reduced quasiprojective scheme over $\mathbb{C}$ (note that X may be reducible, and may have singularities); and the unit discs

$$\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\} \quad \text{and} \quad \mathbb{D}^d := \{z \in \mathbb{C}^d : |z| < 1\}$$

in $\mathbb{C}$ and $\mathbb{C}^d$ ($d \in \mathbb{Z}_{>0}$), respectively.

In this paper, complex spaces are always assumed to be Hausdorff and reduced.

This paper generally follows the definitions of [Zhang 1995].

For the rest of this subsection, let T be a complex space.

A function $f : T \to \mathbb{R}$ is *smooth* if for any holomorphic map $\phi : \mathbb{D}^d \to T$, the composite function $f \circ \phi$ is smooth (i.e., C^∞). Smoothness of differential forms is defined similarly.

Let $\mathscr{L}$ be a line sheaf on T. Then a *smooth hermitian metric* or *continuous hermitian metric* on $\mathscr{L}$ is defined as usual in Arakelov theory, with the metric varying smoothly or continuously, respectively. A *smoothly metrized line sheaf* or *continuously metrized line sheaf* $\mathscr{L}$ on T is a pair $(\mathscr{L}_{\mathrm{fin}}, \|\cdot\|_{\mathscr{L}})$, consisting of a line sheaf $\mathscr{L}_{\mathrm{fin}}$ on T, together with a smooth or continuous hermitian metric $\|\cdot\|_{\mathscr{L}}$ on $\mathscr{L}_{\mathrm{fin}}$, respectively. Most hermitian metrics in this paper are assumed to be smooth. Here the subscript "fin" means *finite*, and is used to refer to the underlying nonmetrized line sheaf (this terminology will make more sense when we get to Arakelov theory). We do not use bars to denote metrized line sheaves: an object $\mathscr{L}$ is what it was said to be. This is because metrized line sheaves are the most natural objects to consider when working in Arakelov theory. For the remainder of this paper, all line sheaves on complex analytic spaces written using notation not involving a subscript "fin" are metrized.

The subscript $\mathscr{L}$ on $\|\cdot\|_{\mathscr{L}}$ may be omitted if $\mathscr{L}$ should be clear from the context.

Let $\mathscr{L}$ be a continuously metrized line sheaf on T. Then a *section* of $\mathscr{L}$ over an open subset U of T is a section of $\mathscr{L}_{\text{fin}}$ over U. A global section s of $\mathscr{L}$ is *small* (resp. *strictly small*) if $\|s\| \le 1$ (resp. $\|s\| < 1$) everywhere on T.

If $\mathscr{L}$ is a smoothly metrized line sheaf on a complex manifold M, then it has a *Chern form* $c_1(\|\cdot\|_{\mathscr{L}})$ (well-)defined by the condition that $c_1(\|\cdot\|_{\mathscr{L}})|_U = -dd^c \log\|s\|^2$ for all open $U \subseteq M$ and all nowhere-vanishing sections s of $\mathscr{L}$ over U. Note that if $\mathscr{L}$ is a smoothly metrized line sheaf on a reduced complex space T, then it may not be possible to define a Chern form $c_1(\|\cdot\|_{\mathscr{L}})$ at singular points of T.

For $n \in \mathbb{Z}_{>0}$, the line sheaf $\mathscr{O}(1)$ on $\mathbb{P}^n(\mathbb{C})$ can be smoothly metrized by the *standard metric*, also called the *Fubini–Study metric*. It is defined uniquely by the condition that, for all global sections $s = a_0z_0 + \cdots + a_nz_n$, where $z_0, \dots, z_n$ are homogeneous coordinates on $\mathbb{P}^n(\mathbb{C})$,

$$\|s\|(p_0 : \cdots : p_n) = \frac{|a_0p_0 + \cdots + a_np_n|}{\sqrt{|p_0|^2 + \cdots + |p_n|^2}}. \tag{1.2}$$

When $n = 1$, the Chern form of this metric is

$$c_1(\|\cdot\|) = \frac{\sqrt{-1}}{2\pi}\frac{1}{(1+|z|^2)^2}dz \wedge d\bar{z} = \frac{dd^c|z|^2}{(1+|z|^2)^2}.$$

Recall that a form on a complex manifold M is *real* if it can be written as a form with real coefficients when M is regarded as a manifold over $\mathbb{R}$. For a $(1, 1)$-form ω on M, written as

$$\omega = \sqrt{-1}\sum_{i,j} h_{ij}(z)dz_i \wedge d\bar{z}_j, \tag{1.3}$$

this is equivalent to $(h_{ij}(z))$ being a Hermitian matrix for all holomorphic local coordinate systems $(z_1, \dots, z_n)$ and all z. Following [Moriwaki 2014, Section 1.12 and 1.14], this form is *positive* (resp. *semipositive*) if it is real and if $(h_{ij}(z))$ is positive definite (resp. positive semidefinite) for all z. A $(1, 1)$-form on a complex projective variety is *semipositive* if its pull-back to a desingularization is semipositive.

Likewise, an (n, n)-form θ on a complex manifold M, written in local coordinates as

$$\theta = \rho(z)dd^c|z_1|^2 \wedge \cdots \wedge dd^c|z_n|^2, \tag{1.4}$$

is real if and only if $\rho(z) \in \mathbb{R}$ for all z, and is *positive* (resp. *semipositive*) if it is real and $\rho(z) > 0$ (resp. $\rho(z) \ge 0$) for all z.

Proposition 1.5. *Positivity of forms as in (1.3) and (1.4) are related as follows*:

(a) *Let M be a complex manifold of dimension n, and let $\omega_1, \dots, \omega_n$ be positive* (*resp. semipositive*) $(1, 1)$*-forms on M. Then $\omega_1 \wedge \cdots \wedge \omega_n$ is positive* (*resp. semipositive*).

(b) *Let X be a complex projective variety of dimension n, and let $\omega_1, \dots, \omega_n$ be semipositive* $(1, 1)$*-forms on X. Then $\omega_1 \wedge \cdots \wedge \omega_n$ is also semipositive.*

Proof. See [Lang 1987, Ch. IV, Lemma 2.4]. For the convenience of the reader, we provide more details here.

It will suffice to prove part (a), since (b) follows by passing to a desingularization.

First assume that $\omega_1, \dots, \omega_n$ are positive. We will use induction on n. The $n = 1$ case is trivial. Fix a point $p \in M$, and let $(z_1, \dots, z_n)$ be a local coordinate system on M near p. We may assume that p corresponds to $z_1 = \dots = z_n = 0$. Write

$$\omega_1 = \sqrt{-1} \sum_{i,j} h_{ij}(z) dz_i \wedge d\bar{z}_j.$$

By Gram–Schmidt, we may assume that $(h_{ij}(0))$ is a diagonal matrix. Since ω_1 is positive, the diagonal entries $\lambda_1, \dots, \lambda_n$ of this matrix are real and positive. For all $i = 1, \dots, n$ and all $j > 1$, $\omega_j|_{z_i=0}$ is positive, so if we write

$$\omega_2 \wedge \dots \wedge \omega_n|_{z_i=0} = \rho_i(z) dd^c|z_1|^2 \wedge \dots \wedge \widehat{(dd^c|z_i|^2)} \wedge \dots \wedge dd^c|z_n|^2$$

for all i, then by induction $\rho_i(0)$ is real and positive. Let $\theta = \omega_1 \wedge \dots \wedge \omega_n$ and let ρ be as in (1.4). Then $\rho(0) = (2\pi)^{-1} \sum \lambda_i \rho_i(0) > 0$, so θ is positive.

The argument for the semipositive case is similar. □

1D. *Arakelov theory.* An *arithmetic variety* is an integral scheme, flat and projective over $\operatorname{Spec} \mathbb{Z}$.

All arithmetic varieties in this paper will be assumed to be normal.

Let X be an arithmetic variety. Let $K = \boldsymbol{\kappa}(X)$; this is a finitely generated extension field of $\mathbb{Q}$. We also write $X_{\mathbb{Q}} = X \times_{\mathbb{Z}} \mathbb{Q}$. The set $X(\mathbb{C})$ will often be regarded as a complex space (with the classical topology).

We say that X is *generically smooth* if $X_{\mathbb{Q}}$ is smooth over $\mathbb{Q}$. If X is generically smooth, then $X(\mathbb{C})$ is a complex manifold (not necessarily connected). If X is an (arbitrary) arithmetic variety, then a *generic resolution of singularities* of X is a proper birational morphism $\pi : Y \to X$, where Y is a generically smooth arithmetic variety.

A *smoothly metrized line sheaf* $\mathscr{L}$ on X is a pair $(\mathscr{L}_{\text{fin}}, \|\cdot\|_{\mathscr{L}})$ consisting of a line sheaf $\mathscr{L}_{\text{fin}}$ on X and a smooth hermitian metric $\|\cdot\|_{\mathscr{L}}$ on $(\mathscr{L}_{\text{fin}})_{\mathbb{C}}$, where $(\mathscr{L}_{\text{fin}})_{\mathbb{C}}$ is the pull-back of $\mathscr{L}_{\text{fin}}$ to $X(\mathbb{C})$. A *continuously metrized line sheaf* on X is defined similarly. In both cases, we will always assume that the hermitian metric is *of real type*; i.e., that it is invariant under the complex conjugation map F_∞ on $X(\mathbb{C})$; see [Moriwaki 2014, (5.2)].

As discussed earlier, metrized line sheaves are not denoted using bars. In this paper, all line sheaves on arithmetic varieties written using notation not involving a subscript "fin" are metrized.

If $\mathscr{L} = (\mathscr{L}_{\text{fin}}, \|\cdot\|_{\mathscr{L}})$ is a smoothly or continuously metrized line sheaf on X, then $\mathscr{L}_{\mathbb{C}}$ will denote the smoothly or continuously metrized line sheaf $((\mathscr{L}_{\text{fin}})_{\mathbb{C}}, \|\cdot\|_{\mathscr{L}})$ on $X(\mathbb{C})$, respectively. We also let $\mathscr{L}_{\mathbb{Q}}$ denote the (nonmetrized) line sheaf $(\mathscr{L}_{\text{fin}})_{\mathbb{Q}}$ on $X_{\mathbb{Q}}$ obtained by restriction.

A *section* of $\mathscr{L}$ over an open subset U of X is a section of $\mathscr{L}_{\text{fin}}$ over U. A global section s of $\mathscr{L}$ is *small* (resp. *strictly small*) if the corresponding section of $\mathscr{L}_{\mathbb{C}}$ is small (resp. strictly small).

If $\mathscr{L}$ is a smoothly metrized line sheaf on X, then its Chern form is the form $c_1(\|\cdot\|_{\mathscr{L}})$; it is a smooth $(1,1)$-form on $X(\mathbb{C})$ and is again denoted $c_1(\|\cdot\|_{\mathscr{L}})$. If $\mathscr{L}$ is a line sheaf (resp. smoothly metrized line sheaf) on a scheme (resp. arithmetic variety) X, then $c_1(\mathscr{L})$ will denote the first Chern class (resp. arithmetic first Chern class) of $\mathscr{L}$; it is a cycle (resp. Arakelov cycle) of codimension 1 on X. In particular, if $\mathscr{L}$ is a smoothly metrized line sheaf on an arithmetic variety X, then $c_1(\mathscr{L}_{\mathbb{Q}})$ is an (ordinary, i.e., non-Arakelov) cycle of codimension 1 on $X_{\mathbb{Q}}$.

In order to simplify the notation, we will often omit $\deg(\cdot)$, even though the product of a number of Chern classes is technically a 0-cycle, not a number. It will always be the intersection number that is meant.

Finally, if $\mathscr{L}$ is a smoothly metrized line sheaf on X then we recall that the *height function* $h_{\mathscr{L}}\colon X(\overline{\mathbb{Q}}) \to \mathbb{R}$ is defined by

$$h_{\mathscr{L}}(x) = \frac{c_1(\mathscr{L}|_{\bar{x}})}{[\boldsymbol{\kappa}(x) : \mathbb{Q}]}, \tag{1.6}$$

where $x \in X(\overline{\mathbb{Q}})$ and $\bar{x}$ is its closure in X. This is an absolute height.

1E. ***Arithmetic intersection theory of Cartier divisors.*** At the present time, a theory of resolution of singularities on arithmetic varieties is not available, so the theory of arithmetic intersection theory on regular varieties, as in [Gillet and Soulé 1990] or [Soulé 1992] cannot be used. Gillet and Soulé [1990, 4.5] construct an intersection theory on generically smooth arithmetic varieties, at the cost of allowing rational coefficients in the Chow groups.

However, since we only need to work with the subring of the Chow ring generated by arithmetic Cartier divisors, it is simpler to use the theory of [Faltings 1992, Lect. 1]. That is what we will do here. It does not require passing to rational coefficients. Moreover, this theory can be extended to an arbitrary arithmetic variety by pulling back to a generic resolution of singularities.

(For generically smooth arithmetic varieties, however, the results of [Gillet and Soulé 1990, Section 1–2], on Green currents and Green forms, can be applied. In fact, they play a key role in this paper.)

Here we follow [Moriwaki 2014, Section 5.4]. A brief summary of his definitions follows.

Let X be a generically smooth arithmetic variety. For $p \in \mathbb{N}$, an *arithmetic cycle* of codimension p on X is a pair $Z = (Z_{\text{fin}}, T)$, where Z_{fin} is a cycle of codimension p on X and T is a current on $X(\mathbb{C})$ of type $(p-1, p-1)$. These form an abelian group under componentwise addition, which is denoted $\mathrm{Z}^p_{\mathrm{D}}(X)$. Note that $\mathrm{Z}^0_{\mathrm{D}}(X) = \mathbb{Z} \cdot (X, 0)$. (Moriwaki denotes this group $\hat{\mathrm{Z}}^p_{\mathrm{D}}(X)$, but the hat is redundant since the subscript "D" already implies that there is a component at infinity.)

A $(p-1, p-1)$-current T on $X(\mathbb{C})$ is said to be *of real type* if $F_\infty^*(T) = (-1)^p T$. Note that this is different from a current (or form) being real; i.e., $\overline{T} = T$.

If Z is a cycle of codimension p on X, then we say that a *Green current* for Z is a current T on $X(\mathbb{C})$ of type $(p-1, p-1)$ such that

$$dd^c T + \delta_Z = [\omega]$$

for some smooth (p, p)-form ω on X. An arithmetic cycle $Z = (Z_{\text{fin}}, T) \in \mathrm{Z}^p_{\mathrm{D}}(X)$ is said to be of *Green type* if T is a Green current for Z_{fin}. These cycles form a subgroup of $\mathrm{Z}^p_{\mathrm{D}}(X)$.

Moriwaki defines $\widehat{\mathrm{Rat}}^p(X)$ to be the subgroup of $\mathrm{Z}_{\mathrm{D}}^p(X)$ generated by (i) cycles $i_*(\operatorname{div}(f)_{\mathrm{fin}}, -\log\|f\|^2)$, where Y is an integral closed subscheme of X of codimension $p-1$, $i\colon Y \to X$ is the corresponding closed embedding, and f is a nonzero rational function on Y; and (ii) $(0, \partial A)$ and $(0, \bar\partial B)$, where A and B are currents on $X(\mathbb{C})$ of type $(p-2, p-1)$ and $(p-1, p-2)$, respectively. He then defines

$$\mathrm{CH}_{\mathrm{D}}^p(X) = \mathrm{Z}_{\mathrm{D}}^p(X)/\widehat{\mathrm{Rat}}^p(X).$$

By way of comparison, Gillet and Soulé [1992, III 1.1] define $\hat{\mathrm{Z}}^p(X)$ to be the subgroup of $\mathrm{Z}_{\mathrm{D}}^p(X)$ consisting of all pairs (Z, T) of Green type such that T is real and of real type, and they let $\widehat{\mathrm{CH}}^p(X)$ be the image of $\hat{\mathrm{Z}}^p(X)$ in $\mathrm{CH}_{\mathrm{D}}^p(X)$. In this paper, all currents under consideration come from (smoothly) metrized line sheaves, so they are real and of real type, but not all pairs $(Z, T) \in \mathrm{Z}_{\mathrm{D}}^p(X)$ in this paper are of Green type, since it is sometimes useful (e.g., in the proof of Lemma 1.11) to split up $(Z, T) \in \mathrm{Z}_{\mathrm{D}}^p(X)$ into a sum $(Z, 0) + (0, T)$.

At times it will be useful to consider intersections on integral closed subschemes of an arithmetic scheme X, including those that are not flat over $\operatorname{Spec}\mathbb{Z}$. Therefore, consider for now an integral scheme X, projective over $\operatorname{Spec}\mathbb{Z}$, which lies entirely over a single closed point $(p) \in \operatorname{Spec}\mathbb{Z}$. Such schemes X will be said to be *vertical*. Since $X_{\mathbb{Q}} = \varnothing$, this scheme is always generically smooth. Similarly, since $X(\mathbb{C}) = \varnothing$, a metrized line sheaf on X (defined as above) is just a pair $\mathscr{L} = (\mathscr{L}_{\mathrm{fin}}, \varnothing)$, and the same definitions as above give that $\mathrm{Z}_{\mathrm{D}}^p(X)$ is the group of pairs $(Z_{\mathrm{fin}}, 0)$, where Z_{fin} is a cycle of codimension p on X in the classical (non-Arakelov) sense. Similarly, $\mathrm{CH}_{\mathrm{D}}^p(X)$ is canonically isomorphic to the classical Chow group $\mathrm{CH}^p(X)$.

Let X be an integral scheme, projective (but not necessarily flat) over $\operatorname{Spec}\mathbb{Z}$, and generically smooth. Let $\mathscr{L}$ be a smoothly metrized line sheaf on X. By [Moriwaki 2014, Def. 5.16, Thm. 5.20, and Section 5.2], the formula

$$(Z, g) \mapsto (\operatorname{div}(s)_{\mathrm{fin}} \cdot Z,\ i_*[-\log\|s|_Z\|^2] + c_1(\|\cdot\|_{\mathscr{L}}) \wedge g) \tag{1.7}$$

gives a well-defined group homomorphism $\mathrm{CH}_{\mathrm{D}}^p(X) \to \mathrm{CH}_{\mathrm{D}}^{p+1}(X)$, denoted $c_1(\mathscr{L})\,\cdot$, where $(Z, g) \in \mathrm{Z}_{\mathrm{D}}^p(X)$ is such that Z is a closed integral subscheme of X, $i\colon Z \to X$ is the corresponding closed embedding, and s is a rational section of $\mathscr{L}$ whose restriction to Z is nonzero. (If Z is vertical, then $Z(\mathbb{C}) = \varnothing$, and therefore $i_*[-\log\|s|_Z\|^2] = 0$.) It is easy to check that (i) $c_1(\mathscr{L}) \cdot (X, 0) = c_1(\mathscr{L})$, where $c_1(\mathscr{L})$ on the right-hand side is as defined earlier, and (ii) if X is regular and if $\alpha \in \widehat{\mathrm{CH}}^p(X)$, then $c_1(\mathscr{L}) \cdot \alpha$ as defined here coincides with the definition from [Gillet and Soulé 1990, Section 3] (or with classical intersection theory if X is vertical).

Let X be an integral scheme, projective over $\operatorname{Spec}\mathbb{Z}$ and generically smooth; let $n = \dim X$; and let $\mathscr{L}_1, \dots, \mathscr{L}_n$ be smoothly metrized line sheaves on X. Then we have a well-defined element

$$c_1(\mathscr{L}_1) \cdots c_1(\mathscr{L}_n) \in \mathrm{CH}_{\mathrm{D}}^n(X).$$

Since this is a cycle of dimension 0 on X, we can take its degree [Moriwaki 2014, Def. 5.22] to get a real number, which will also (by the usual abuse of notation) be denoted $c_1(\mathscr{L}_1) \cdots c_1(\mathscr{L}_n)$. This degree is always taken in the Arakelov sense, even if X is vertical.

The map $c_1(\mathscr{L})\cdot$ satisfies the following projection formula. Let X and Y be integral schemes, projective over $\operatorname{Spec}\mathbb{Z}$ and generically smooth; let $f\colon X\to Y$ be a morphism; let $\mathscr{L}_1,\dots,\mathscr{L}_n$ be smoothly metrized line sheaves on Y; and let $\alpha\in \mathrm{CH}^p_{\mathrm{D}}(X)$. Then

$$f_*(c_1(f^*\mathscr{L}_1)\cdots c_1(f^*\mathscr{L}_n)\cdot\alpha)=c_1(\mathscr{L}_1)\cdots c_1(\mathscr{L}_n)\cdot f_*\alpha. \tag{1.8}$$

Indeed, when $n=1$ this is [Moriwaki 2014, Thm. 5.20 (2) and Prop. 5.5], and the general case follows by induction. In particular, if f is birational and $n=\dim X$, then (taking degrees) we have

$$\begin{aligned} c_1(f^*\mathscr{L}_1)\cdots c_1(f^*\mathscr{L}_n) &= c_1(f^*\mathscr{L}_1)\cdots c_1(f^*\mathscr{L}_n)\cdot(X,0)\\ &= c_1(\mathscr{L}_1)\cdots c_1(\mathscr{L}_n)\cdot(Y,0)\\ &= c_1(\mathscr{L}_1)\cdots c_1(\mathscr{L}_n). \end{aligned}\tag{1.9}$$

By pulling back to a generic resolution of singularities and using (1.9), one can also define this quantity without assuming that X is generically smooth; see [Moriwaki 2014, Def. 5.24]. With this definition, (1.9) holds without the assumption that X and Y are generically smooth.

We conclude this section with a result which is implicit several places in Moriwaki's work, and obvious to the experts, but which seems not to be explicitly stated or proved anywhere.

Definition 1.10. Let X be an arithmetic variety, and let $\mathscr{M}$ be a continuously metrized line sheaf on X. Then, for any nonzero rational section s of $\mathscr{M}$, we define

$$c_1(\mathscr{M})=(\operatorname{div}(s)_{\mathrm{fin}},-\log\|s\|^2)\in \mathrm{CH}^1_{\mathrm{D}}(X).$$

This definition is independent of the choice of s, and is compatible with the definition of $c_1(\mathscr{M})\in \mathrm{CH}^1(X)$ when the metric on $\mathscr{M}$ is smooth.

Lemma 1.11. *Let X be a generically smooth arithmetic variety of dimension n. Let $\mathscr{L}_1,\dots,\mathscr{L}_{n-1}$ be smoothly metrized line sheaves on X, let $\mathscr{M}$ be a continuously metrized line sheaf on X, and let s be a nonzero rational section of $\mathscr{M}$. Write $\operatorname{div}(s)_{\mathrm{fin}}$ as a finite sum $\sum n_Z Z$, where $n_Z\in\mathbb{Z}$ for all Z and each Z is a prime divisor on X. Then*

$$\begin{aligned} &c_1(\mathscr{L}_1)\cdots c_1(\mathscr{L}_{n-1})\cdot c_1(\mathscr{M})\\ &\quad=\sum n_Z(c_1(\mathscr{L}_1|_Z)\cdots c_1(\mathscr{L}_{n-1}|_Z))-\int_{X(\mathbb{C})}\log\|s\|\,c_1(\|\cdot\|_{\mathscr{L}_1})\wedge\cdots\wedge c_1(\|\cdot\|_{\mathscr{L}_{n-1}}). \end{aligned}\tag{1.11.1}$$

Proof. Since both sides of (1.11.1) are linear in $\mathscr{M}$ (and correspondingly in n_Z and $-\log\|s\|$), we may assume that there is only one prime divisor Z, and that $n_Z=1$. Then $c_1(\mathscr{M})$ is represented by the cycle $(Z,-\log\|s\|^2)$, and we have

$$\begin{aligned} &c_1(\mathscr{L}_1)\cdots c_1(\mathscr{L}_{n-1})\cdot c_1(\mathscr{M})\\ &\quad= c_1(\mathscr{L}_1)\cdots c_1(\mathscr{L}_{n-1})\cdot(Z,0)+c_1(\mathscr{L}_1)\cdots c_1(\mathscr{L}_{n-1})\cdot(0,-\log\|s\|^2). \end{aligned}\tag{1.11.2}$$

We first consider the first term on the right-hand side.

Let $\tilde{Z}$ be a generic resolution of singularities of Z, and let $f\colon \tilde{Z} \to X$ be the corresponding map to X. By (1.8) and (1.9),

$$\begin{aligned} c_1(\mathscr{L}_1)\cdots c_1(\mathscr{L}_{n-1})\cdot(Z,0) &= c_1(\mathscr{L}_1)\cdots c_1(\mathscr{L}_{n-1})\cdot f_*(\tilde{Z},0) \\ &= c_1(f^*(\mathscr{L}_1))\cdots c_1(f^*(\mathscr{L}_{n-1}))\cdot(\tilde{Z},0) \\ &= c_1(f^*(\mathscr{L}_1))\cdots c_1(f^*(\mathscr{L}_{n-1})) \\ &= c_1(\mathscr{L}_1|_Z)\cdots c_1(\mathscr{L}_{n-1}|_Z) \end{aligned} \tag{1.11.3}$$

(where the last formula is computed on Z).

Now consider the second term on the right-hand side of (1.11.2).

If g is a current of type $(n-1-i, n-1-i)$ on $X(\mathbb{C})$, then by (1.7),

$$c_1(\mathscr{L}_i)\cdot(0,g) = (0, c_1(\|\cdot\|_{\mathscr{L}_i})\wedge g), \tag{1.11.4}$$

and therefore (taking the degree)

$$c_1(\mathscr{L}_1)\cdots c_1(\mathscr{L}_{n-1})\cdot(0,-\log\|s\|^2) = -\int_{X(\mathbb{C})}\log\|s\|\, c_1(\|\cdot\|_{\mathscr{L}_1})\wedge\cdots\wedge c_1(\|\cdot\|_{\mathscr{L}_{n-1}}).$$

Combining (1.11.2)–(1.11.4) then gives (1.11.1). □

Proposition 1.12. *Let X be an integral scheme of dimension n, projective over* $\operatorname{Spec}\mathbb{Z}$.

(a) *Let $\mathscr{L}_1,\dots,\mathscr{L}_n$ be nef, smoothly metrized line sheaves on X, or*

(b) *let $\mathscr{L}_1,\dots,\mathscr{L}_{n-1}$ be nef, smoothly metrized line sheaves on X, and let $\mathscr{L}_n$ be a continuously metrized line sheaf on X for which some positive tensor power has a small nonzero global section.*

Then

$$c_1(\mathscr{L}_1)\cdots c_1(\mathscr{L}_n)\ge 0.$$

Proof. If X an arithmetic variety, then part (a) is [Moriwaki 2000, Prop. 2.3 (1)] or [Moriwaki 2014, Thm. 6.15]. Otherwise, it is a standard result in algebraic geometry.

Part (b) is [Moriwaki 2000, Prop. 2.3 (2)]. It follows from part (a) and Lemma 1.11. □

2. Positivity conditions on metrized line sheaves

This section defines the conditions nef, big, and ample for a smoothly metrized line sheaf on an arithmetic variety, and gives some of their main properties.

References for this section are [Zhang 1995; Yuan 2008; 2009; Moriwaki 2014].

Throughout this section, $\mathscr{L}$ is a smoothly metrized line sheaf on an arithmetic variety X.

2A. ***Nef metrized line sheaves.***

Definition 2.1 [Moriwaki 2014, Definition 5.38(3)]. (a) $\mathscr{L}$ is *vertically nef* if $\mathscr{L}_{\text{fin}}$ is nef on all closed fibers of $X \to \operatorname{Spec}\mathbb{Z}$ and if the metric on $\mathscr{L}$ is semipositive, and

(b) $\mathscr{L}$ is *nef* if it is vertically nef and if $h_{\mathscr{L}}(x) \ge 0$ for all $x \in X(\overline{\mathbb{Q}})$.

Proposition 2.2. *Let $f : X' \to X$ be a surjective generically finite morphism of arithmetic varieties. If $\mathscr{L}$ is nef, then so is $f^*\mathscr{L}$.*

Proof. This is clear from the definition. □

2B. ***Big metrized line sheaves.*** The definition of a big metrized line sheaf given here is modeled after the definition of big in the classical case.

Definition 2.3. (a) Let $H^0(X, \mathscr{L})$ denote the set of small sections of $\mathscr{L}$:

$$H^0(X, \mathscr{L}) = \{s \in H^0(X, \mathscr{L}_{\text{fin}}) : \|s\|_{\sup} \le 1\},$$

and let

$$h^0(X, \mathscr{L}) = \log \#H^0(X, \mathscr{L}).$$

(b) Let $n = \dim X$. Then the *volume* of $\mathscr{L}$ is

$$\operatorname{vol}(\mathscr{L}) = \limsup_{m\to\infty} \frac{h^0(X, \mathscr{L}^{\otimes m})}{m^n/n!}.$$

By [Yuan 2009, Section 1.1 and Theorem 2.7] (see also [Chen 2008]), this lim sup converges as a limit.

(c) We say that $\mathscr{L}$ is *big* if $\operatorname{vol}(\mathscr{L}) > 0$.

Remark 2.4. Sometimes $H^0(X, \mathscr{L})$ is defined to be the set of strictly small sections of $\mathscr{L}$, and this definition is used to define bigness. This definition of big (and several others) are equivalent to Definition 2.3(c), by [Yuan 2008, Corollary 2.4] and [Moriwaki 2009, Theorem 4.6].

Proposition 2.5. *Let $f : X' \to X$ be a surjective generically finite morphism of arithmetic varieties. If $\mathscr{L}$ is big, then so is $f^*\mathscr{L}$.*

Proof. This is immediate from the fact that pulling back by f induces an injection $H^0(X, \mathscr{L}) \to H^0(X', f^*\mathscr{L})$. □

2C. ***Ample metrized line sheaves.*** To define ampleness, we follow [Yuan 2008, Secton 2.1].

Definition 2.6. We say that $\mathscr{L}$ is *horizontally positive* if $c_1(\mathscr{L}|_Y)^{\cdot \dim Y} > 0$ for all horizontal integral closed subschemes Y of X. Here an integral subscheme Y of X is *horizontal* if it is flat over $\operatorname{Spec}\mathbb{Z}$.

Definition 2.7. A smoothly metrized line sheaf $\mathscr{L}$ is *ample* if

(i) $\mathscr{L}_{\mathbb{Q}}$ is ample;

(ii) $\mathscr{L}$ is vertically nef; and

(iii) $\mathscr{L}$ is horizontally positive.

Remark 2.8. Moriwaki defines ampleness differently. He defines $\mathscr{L}$ to be ample if (i) $\mathscr{L}_{\text{fin}}$ is ample (on X), (ii) the metric on $\mathscr{L}$ is positive, and (iii) there is some integer $n > 0$ such that $H^0(X, \mathscr{L}_{\text{fin}}^{\otimes n})$ is generated by strictly small sections [Moriwaki 2014, Definition 5.38 (2)]. This definition is stronger than Definition 2.7. Indeed, (i) and (ii) of Definition 2.7 follow from Moriwaki's (i) and (ii), and horizontal positivity follows from [Moriwaki 2014, Proposition 5.39]. The converse implication is false: for example, if $\mathscr{L}$ is ample on X in the sense of Moriwaki, then its pull-back to the blowing-up of X at a closed point is ample in the sense of Definition 2.7, but not in Moriwaki's sense.

Proposition 2.9. *If $\mathscr{L}$ is ample, then it is nef and big.*

Proof. The fact that $\mathscr{L}$ is nef follows immediately by comparing Definitions 2.7 and 2.1. That $\mathscr{L}$ is big follows from [Yuan 2008, Corollary 2.4]. □

2D. *An openness property.* Because a metrized line sheaf is only required to be vertically nef in order to be ample, arithmetical ampleness is not an open condition. However, it is true that arithmetical ampleness is preserved under changing the metric by a constant multiple sufficiently close to 1, provided that the arithmetic variety is generically smooth. This is the conclusion of Proposition 2.12, which is the goal of this subsection.

Note that the definition of ampleness is comparable to the Nakai–Moishezon criterion. This implies something comparable to the more common definition of ampleness in the non-Arakelov setting [Zhang 1995, Corollary 4.8].

We start with a result that may be regarded as a counterpart to the theorem in classical algebraic geometry that says that the Nakai–Moishezon and Kleiman criteria for ampleness are equivalent.

Lemma 2.10. *Assume that X is generically smooth, that $\mathscr{L}_{\mathbb{Q}}$ is ample, and that the metric on $\mathscr{L}$ is semipositive. Then $\mathscr{L}$ is horizontally positive if and only if the height function $h_{\mathscr{L}}$ has a positive lower bound on X.*

Proof. This proof makes use of the condition that a smoothly metrized line sheaf be *relatively semiample*. We will not quote the definition here (see [Zhang 1995, (3.1)]); instead, it is sufficient to know that $\mathscr{L}|_Y$ is relatively semiample for all horizontal integral closed subschemes Y of X [Zhang 1995, Theorem 3.5].

This proof follows fairly easily from the equivalence (ii)$\iff$(iii) of [Zhang 1995, Corollary 5.7]. This says the following. Let $\mathscr{M}$ be a smoothly metrized line sheaf on an arithmetic variety Y. Assume that $\mathscr{M}_{\mathbb{Q}}$ is ample, that $\mathscr{M}$ is relatively semiample, and that $h_{\mathscr{M}}(y) \ge 0$ for all $y \in Y(\overline{\mathbb{Q}})$. Then the following conditions are equivalent: (ii) there is a nonempty Zariski-open subset U of Y such that $h_{\mathscr{M}}$ has a positive lower bound on U (i.e., on $U(\overline{\mathbb{Q}})$), and (iii) $c_1(\mathscr{M})^{\cdot \dim Y} > 0$.

We will apply this result with Y equal to a horizontal integral closed subscheme of X and with $\mathscr{M} = \mathscr{L}|_Y$. In this situation, $\mathscr{M}$ is relatively semiample as noted above, and $\mathscr{M}_{\mathbb{Q}}$ is ample because $\mathscr{L}_{\mathbb{Q}}$ is.

We first prove the converse assertion. Assume that $h_{\mathscr{L}}$ has a positive lower bound on X, let Y be a horizontal integral closed subscheme on X, and let $\mathscr{M} = \mathscr{L}|_Y$. Then condition (ii) in Zhang's lemma holds for Y and $\mathscr{M}$ with $U = Y$, and also the hypothesis $h_{\mathscr{M}}(y) \ge 0$ holds. Therefore, by (iii), $c_1(\mathscr{M})^{\cdot \dim Y} > 0$. Since Y is arbitrary, $\mathscr{L}$ is horizontally positive.

Conversely, assume that $\mathscr{L}$ is horizontally positive. We will show by noetherian induction that $h_{\mathscr{L}}$ has a positive lower bound on $Y(\overline{\mathbb{Q}})$ for all Zariski-closed subsets Y of X, and therefore it holds for X.

Let Y be a Zariski-closed subset of X. If $Y = \varnothing$ then there is nothing to show. If Y is reducible, then write $Y = Y_1 \cup \cdots \cup Y_n$ with all Y_i irreducible. By the inductive hypothesis, $h_{\mathscr{L}}$ has a positive lower bound on Y_i for all i, so the same is true for Y.

Assume now that Y is irreducible. If Y is not horizontal, then $Y(\overline{\mathbb{Q}})$ is empty, and there is nothing to prove. Otherwise, we apply the above result of Zhang. Note that, since $\mathscr{L}$ is horizontally positive, the hypothesis that $h_{\mathscr{M}}(y) \ge 0$ for all $y \in Y(\overline{\mathbb{Q}})$ holds, and so does condition (iii) of Zhang's corollary. Therefore, by condition (ii) of the corollary, there is a nonempty open $U \subseteq Y$ such that $h_{\mathscr{L}}$ has a positive lower bound on U. Also $h_{\mathscr{L}}$ has a positive lower bound on $Y \setminus U$ by the inductive hypothesis, so $h_{\mathscr{L}}$ has a positive lower bound on Y.

It follows by taking $Y = X$ that $h_{\mathscr{L}}$ has a positive lower bound on X. □

Definition 2.11. For all $a \in \mathbb{R}$ let $\mathscr{V}_a$ be the smoothly metrized line sheaf on X such that $(\mathscr{V}_a)_{\mathrm{fin}}$ is the structure sheaf of X and the constant section 1 of $\mathscr{V}_a$ has constant metric e^{-a}. (Here $\mathscr{V}$ stands for *vertical.*)

We are now ready to prove the main result of this subsection.

Proposition 2.12. *Assume that X is generically smooth and that $\mathscr{L}$ is ample. Then there is a $c > 0$ such that $\mathscr{L} \otimes \mathscr{V}_{-\epsilon}$ is ample for all $\epsilon < c$.*

Proof. For all $\epsilon \in \mathbb{R}$, the properties $(\mathscr{L} \otimes \mathscr{V}_{-\epsilon})_{\mathbb{Q}}$ ample and $\mathscr{L} \otimes \mathscr{V}_{-\epsilon}$ vertically nef follow trivially from the same properties of $\mathscr{L}$. Therefore it will suffice to find $c > 0$ such that $\mathscr{L} \otimes \mathscr{V}_{-\epsilon}$ is horizontally positive for all $\epsilon < c$.

Let

$$c = \inf_{x \in X(\overline{\mathbb{Q}})} h_{\mathscr{L}}(x).$$

By Lemma 2.10, $c > 0$. Fix $\epsilon < c$. We need to show that $\mathscr{L} \otimes \mathscr{V}_{-\epsilon}$ is horizontally positive. To see this, we note that

$$h_{\mathscr{L} \otimes \mathscr{V}_{-\epsilon}}(x) = h_{\mathscr{L}}(x) - \epsilon$$

for all $x \in X(\overline{\mathbb{Q}})$. Then $h_{\mathscr{L} \otimes \mathscr{V}_{-\epsilon}}$ has the positive lower bound $c - \epsilon$, and therefore $\mathscr{L} \otimes \mathscr{V}_{-\epsilon}$ is ample by Lemma 2.10. □

3. Arithmetic function fields

An *arithmetic function field* is a finitely generated extension field of $\mathbb{Q}$. Such fields have a diophantine theory that contains the number field case as a special case.

This theory was originally developed in [Moriwaki 2000]. See also the survey article [Moriwaki 2002].

3A. *Polarizations, places, and heights.*

Definition 3.1. Let K be an arithmetic function field, and let $d = \operatorname{tr.deg}_{\mathbb{Q}} K$. Then a *polarization* of K consists of

(i) an arithmetic variety B, given with an isomorphism $\kappa(B) \xrightarrow{\sim} K$, and

(ii) nef smoothly metrized line sheaves $\mathscr{M}_1, \dots \mathscr{M}_d$ on B.

Such a polarization will be denoted $M = (B; \mathscr{M}_1, \dots, \mathscr{M}_d)$. A polarization will be said to be *big* if $\mathscr{M}_1, \dots, \mathscr{M}_d$ are all big.

We now define a set of places of K to replace the set M_k of places of a number field k recalled in Section 1. This description follows [Burgos Gil et al. 2016, Section 1] as well as [Moriwaki 2000].

We assume from now on that B is normal.

We start with the nonarchimedean places. Let $B^{(1)}$ denote the set of prime (Weil) divisors on B; i.e., the set of integral closed subschemes of B of codimension 1. (These may be horizontal or vertical.)

Let Y be a prime divisor on B, and let

$$h_M(Y) = c_1(\mathscr{M}_1|_Y) \cdots c_1(\mathscr{M}_d|_Y). \tag{3.2}$$

By Proposition 1.12(a), $h_M(Y) \ge 0$. For nonzero $x \in K$, we then define a nonarchimedean absolute value associated to Y as

$$\|x\|_Y = \exp(-h_M(Y) \operatorname{ord}_Y(x)). \tag{3.3}$$

(Note that if $d = 0$ then K is a number field k, Y is a closed point on $\operatorname{Spec} \mathscr{O}_k$, and the intersection product (3.2) is just the cycle Y, whose degree is the logarithm of the number of elements in the residue field. Therefore $\|x\|_Y$ coincides with $\|x\|_v$ for the place $v \in M_k$ that corresponds to Y.)

The set $B^{(1)}$ will be the set of nonarchimedean places of K. We write $M_K^0 = B^{(1)}$ and let μ_{fin} be the counting measure on $B^{(1)} = M_K^0$.

For archimedean places, we define the *set of generic points* of $B(\mathbb{C})$ as

$$B(\mathbb{C})^{\text{gen}} = B(\mathbb{C}) \setminus \bigcup_{Y \in B^{(1)}} Y(\mathbb{C}).$$

For such a generic point $b \in B(\mathbb{C})^{\text{gen}}$, we define an absolute value

$$\|x\|_b = |x|_b = |x(b)|$$

for all $x \in K$. Note that $x(b) \in \mathbb{C}$, because b does not lie on a pole of the function x: all such poles lie in elements of $B^{(1)}$.

The set $B(\mathbb{C})^{\text{gen}}$ will be the set of archimedean places of K, and we will usually denote it M_K^∞. In sharp contrast to the number field case, if $d > 0$ then there are uncountably many archimedean places.

We let μ_∞ be the Lebesgue measure on $B(\mathbb{C})$ associated to the (d,d)-form $c_1(\|\cdot\|_{\mathscr{M}_1}) \cdots c_1(\|\cdot\|_{\mathscr{M}_d})$. This form is semipositive by Proposition 1.5(b). The set $B(\mathbb{C}) \setminus B(\mathbb{C})^{\text{gen}}$ is a countable union of the sets

$Y(\mathbb{C})$, all of which have measure zero, so $B(\mathbb{C}) \setminus B(\mathbb{C})^{\mathrm{gen}}$ has measure zero. We also regard μ_∞ as a measure on $B(\mathbb{C})^{\mathrm{gen}}$. We then have

$$\mu_\infty(B(\mathbb{C})^{\mathrm{gen}}) = c_1((\mathscr{M}_1)_{\mathbb{Q}}) \cdots c_1((\mathscr{M}_d)_{\mathbb{Q}}) < \infty. \tag{3.4}$$

One can then let M_K be the disjoint union

$$M_K = M_K^\infty \amalg M_K^0 = B(\mathbb{C})^{\mathrm{gen}} \amalg B^{(1)},$$

and combine the measures μ_∞ on $B(\mathbb{C})$ and μ_{fin} on $B^{(1)}$ to give a measure μ on $B(\mathbb{C}) \amalg B^{(1)} \supseteq M_K$. As in [Moriwaki 2000, Section 3.2], this then leads to a *product formula*

$$\int_{M_K} \log\|x\|_v \, d\mu(v) = 0 \quad \text{for all } x \in K^* \tag{3.5}$$

and a "*naïve height*"

$$\begin{aligned} h_K(x) &= \int_{M_K} \log^+\|x\|_v \, d\mu(v) \\ &= \int_{B(\mathbb{C})^{\mathrm{gen}}} \log^+|x(b)| \, d\mu_\infty(b) + \sum_{Y \in B^{(1)}} \max\{0, -\operatorname{ord}_Y(x)\} h_M(Y) \end{aligned} \tag{3.6}$$

for all $x \in K$; here we take $\max\{0, -\operatorname{ord}_Y(x)\} = 0$ if $x = 0$. Note that $h_K(x) \ge 0$ for all $x \in K$.

Remark 3.7. The set of archimedean places of K can be canonically identified with the set of embeddings of K into $\mathbb{C}$, in such a way that if an archimedean place v of K corresponds to $\sigma : K \to \mathbb{C}$, then

$$\|x\|_v = |\sigma(x)| \tag{3.7.1}$$

for all $x \in K$. So this is just like the number field case. The construction using $B(\mathbb{C})^{\mathrm{gen}}$ is necessary to define the measure.

To see this identification, recall from [Hartshorne 1977, II, Exercise 2.7] that giving an element of $B(\mathbb{C})$ is equivalent to giving a point $P \in B$ and an injection $\kappa(P) \hookrightarrow \mathbb{C}$. The elements of $B(\mathbb{C})^{\mathrm{gen}}$ are exactly those for which the point P is the generic point of B. Thus $B(\mathbb{C})^{\mathrm{gen}}$ is in natural bijection with $\operatorname{Hom}(K, \mathbb{C})$, and (3.7.1) is true.

Definition 3.8. For all $v \in M_K$ we define a field extension $\mathbb{C}_v/K$ as follows. If v is archimedean, then let $\mathbb{C}_v = \mathbb{C}$, viewed as an extension of K by the embedding $K \hookrightarrow \mathbb{C}$ of Remark 3.7. If v is nonarchimedean, then we let $\mathbb{C}_v$ be the completion of the algebraic closure $\overline{K}_v$ of the completion K_v of K at v. This field is algebraically closed [Bosch et al. 1984, Proposition 3.4.1/3].

3B. *Finite extensions of arithmetic function fields.* Let K be an arithmetic function field of transcendence degree d over $\mathbb{Q}$, and let K' be a finite extension of K. Then K' is also an arithmetic function field of transcendence degree d.

Definition 3.9. Let $M = (B; \mathscr{M}_1, \dots, \mathscr{M}_d)$ be a polarization of K. We define a polarization M' of K' as follows. Let B' be the normalization of B in K', and let $\pi : B' \to B$ be the associated map. Then π is a finite morphism of degree $[K' : K]$, and of course B' is normal. Let $\mathscr{M}'_i = \pi^*\mathscr{M}_i$ for all i; these are nef line sheaves on B' by Proposition 2.2. Thus $M' := (B'; \mathscr{M}'_1, \dots, \mathscr{M}'_d)$ is a polarization of K', and is called the polarization of K' *induced by* M, or the *induced polarization* of K' if M is clear from the context.

The absolute values of K' are related to those of K as follows.

Definition 3.10. Let M, M', and $\pi : B' \to B$ be as in Definition 3.9, let $v \in M_K$, and let $w \in M_{K'}$. Then we say that w *lies over* v, and write $w \mid v$, if one of the following holds:

(i) Both w and v are archimedean, corresponding to $b' \in B'(\mathbb{C})^{\text{gen}}$ and $b \in B(\mathbb{C})^{\text{gen}}$, respectively, and $\pi(b') = b$.

(ii) Both w and v are nonarchimedean, corresponding to prime divisors Y' on B' and Y on B, respectively, and $\pi(Y') = Y$.

As in [Moriwaki 2000, Section 3.2], we then have:

Proposition 3.11. *Let $v \in M_K$. For each $w \in M_{K'}$ lying over v there is a canonical injection $i : \mathbb{C}_v \to \mathbb{C}_w$ of fields, and a canonical integer $n_{w/v}$ such that*

$$\|i(x)\|_w = \|x\|_v^{n_{v/w}} \tag{3.11.1}$$

for all $x \in \mathbb{C}_v$. Moreover,

$$\sum_{w|v} n_{w/v} = [K' : K], \tag{3.11.2}$$

$$\prod_{w\,|\,v} \|i(x)\|_w = \|x\|_v^{[K':K]} \quad \textit{for all } x \in \mathbb{C}_v, \tag{3.11.3}$$

and

$$h_{K'}(x) = [K' : K]h_K(x) \quad \textit{for all } x \in K. \tag{3.11.4}$$

Proof. If v is archimedean, then let $\sigma : K \to \mathbb{C}$ and $\sigma' : K' \to \mathbb{C}$ be injections as in Remark 3.7 for v and w, respectively. Then $i : \mathbb{C}_v \to \mathbb{C}_w$ is just the identity map on $\mathbb{C}$ via the identifications $\mathbb{C}_v = \mathbb{C} = \mathbb{C}_w$, and the diagram

$$\begin{array}{ccc} K & \xrightarrow{\ \sigma\ } & \mathbb{C} \\ \big\downarrow & & \big\downarrow i \\ K' & \xrightarrow{\ \sigma'\ } & \mathbb{C} \end{array}$$

commutes; therefore (3.11.1) holds with $n_{w/v} = 1$. Moreover, since K'/K is separable, there are exactly $[K' : K]$ places w lying over v, and this gives (3.11.2).

If v is nonarchimedean, then it corresponds to a prime divisor Y on B, and $\pi^*Y = \sum_i e_i Y_i$, where the Y_i are the irreducible components of $\pi^{-1}(Y)$. These correspond to the places w of K' lying over v. Let f_i

be the residue degree $[K(Y_i):K(Y)]$ for all i. Then, for all i, $h_{M'}(Y_i)=f_i h_M(Y)$ and $\operatorname{ord}_{Y_i} x = e_i \operatorname{ord}_Y x$ for all $x\in K^*$. Therefore (3.11.1) holds with $n_{w/v}=e_i f_i$ if w corresponds to Y_i. Also (3.11.2) holds by the basic theory of Dedekind rings applied to the local ring $\mathscr{O}_{B,Y}$ and its integral closure in K'.

Finally, in both cases (3.11.3) and (3.11.4) follow immediately from (3.11.1) and (3.11.2). □

3C. *Models and Arakelov heights.* For higher generality, Roth's theorem over arithmetic function fields is best formulated using Arakelov theory, using a model for $\mathbb{P}^1_K$.

Throughout this subsection let $(B;\mathscr{M}_1,\dots,\mathscr{M}_d)$ be a polarization of K.

Definition 3.12. Let V be a projective variety over K. A *model* for V over B consists of an arithmetic variety X, a morphism $X\to B$, and an isomorphism $i: V \xrightarrow{\sim} X_K$ over K. We say that a given line sheaf $\mathscr{L}$ (resp. Cartier divisor D) on V *extends to* X if there is a smoothly metrized line sheaf $\mathscr{L}'$ (resp. Arakelov Cartier divisor D') on X such that $i^*\mathscr{L}'_{\text{fin}}\cong\mathscr{L}$ (resp. $i^*D'_{\text{fin}}=D$).

Remark 3.13. Let V, X, and π be as above. Not every line sheaf $\mathscr{L}$ or Cartier divisor D on V extends to X, but there is always a model for V to which $\mathscr{L}$ or D extends. For existence of a model X, we may take an embedding of V into $\mathbb{P}^n_K$, and let X be the closure of the image in $\mathbb{P}^n_B$. To see that for any given Cartier divisor D on V there is a model to which D extends, it will suffice for our purposes to assume that V is nonsingular. Take any model X_0 for V, extend each irreducible component of $\operatorname{Supp} D$ to X_0 as a Weil divisor, and blow up the sheaves of ideals of the closure in X_0 of each such irreducible component. The resulting scheme X will then be a model for V to which D extends as a Cartier divisor. Given any line sheaf $\mathscr{L}$ on V, one can then find a model to which $\mathscr{L}$ extends by writing $\mathscr{L}\cong\mathscr{O}(D)$ for a Cartier divisor D, and finding a model to which D extends. For more general situations; see [Vojta 2007].

We can now define height functions in terms of Arakelov theory.

Definition 3.14 [Moriwaki 2000, Section 3.3]. Let $\pi: X\to B$ be a model for a variety V over K, and let $\mathscr{L}$ be a continuously metrized line sheaf on X. Then the *Arakelov height* of a point $x\in V(\bar K)$ (or, equivalently, $x\in X(\bar K)$) is given by

$$h_{\mathscr{L}}(x)=\frac{c_1(\pi^*\mathscr{M}_1|_{\bar x})\cdots c_1(\pi^*\mathscr{M}_d|_{\bar x})\cdot c_1(\mathscr{L}|_{\bar x})}{[\boldsymbol{\kappa}(x):K]}. \tag{3.14.1}$$

Here, as usual, $\bar x$ denotes the closure of x in X. (Compare with (1.6).)

We will use the following results of Moriwaki.

Proposition 3.15 [Moriwaki 2000, Proposition 3.3.1]. *Let V, X, π, and $\mathscr{L}$ be as above. Let K' be a finite extension of K, and let $(B';\mathscr{M}'_1,\dots,\mathscr{M}'_d)$ be the polarization of K' induced by the polarization $(B;\mathscr{M}_1,\dots,\mathscr{M}_d)$ of K. Let X' be the main component of $X\times_B B'$ (the latter may have many components if V is not geometrically integral over K). Let $f: X'\to X$ be the projection morphism, and let $\mathscr{L}'=f^*\mathscr{L}$. Here X' is a model over B' for the main component $V_{K'}$ of $V\times_K K'$. For all $x\in X(\bar K)$, pick $x'\in X'(\bar K')$ lying over x. Then*

$$h_{\mathscr{L}'}(x')=[K':K]h_{\mathscr{L}}(x). \tag{3.15.1}$$

Theorem 3.16 (Northcott's finiteness theorem, [Moriwaki 2000, Theorem 4.3]). *Let V, X, and π be as above, and let $\mathscr{L}$ be a continuously metrized line sheaf on X. Assume that the polarization $(B; \mathscr{M}_1, \dots, \mathscr{M}_d)$ of K is big; i.e., that all $\mathscr{M}_i$ are big (see* [Yuan 2008, Corollary 2.4]; *it suffices if the $\mathscr{M}_i$ are ample). Assume also that $\mathscr{L}_K$ is ample. Then for all $C \in \mathbb{R}$ and all $n \in \mathbb{Z}_{>0}$, the set*

$$\{x \in X(\bar{K}) : h_{\mathscr{L}}(x) \le C \text{ and } [\boldsymbol{\kappa}(x) : K] \le n\}$$

is finite.

Proposition 3.17 [Moriwaki 2000, Proposition 3.3.2]. *Let $\mathscr{L}$ be the continuously metrized line sheaf on $\mathbb{P}^1_B$ such that $\mathscr{L}_{\text{fin}}$ is the tautological line sheaf $\mathscr{O}(1)$ on $\mathbb{P}^1_B$ and the metric is uniquely defined by the condition that for all global sections $s = a_0 z_0 + a_1 z_1$, where z_0, z_1 are the standard homogeneous coordinates on $\mathbb{P}^1$,*

$$\|s\|(p_0 : p_1) = \frac{\|a_0 p_0 + a_1 p_1\|_v}{\max\{\|p_0\|_v, \|p_1\|_v\}}.$$

Then the Arakelov height $h_{\mathscr{L}}$ is equal to the "naïve height" h_K of (3.6).

This then gives a Northcott finiteness theorem for the naïve height as an immediate corollary.

3D. *M_K-constants and Weil functions.* This paper will rely heavily on Weil functions (also called local heights). As far as I know, they have not been developed in the context of arithmetic function fields, but their construction from the number field case carries over directly, once the definitions have been chosen.

Throughout this subsection, K is an arithmetic function field, with polarization $(B; \mathscr{M}_1, \dots, \mathscr{M}_d)$. Models over B of varieties are not necessary for the theory of Weil functions itself, although they can be used to construct examples of Weil functions. We do need the polarization, though, because it determines M_K.

Definition 3.18. An M_K-constant is a measurable, L^1 function from M_K to $\mathbb{R}$, whose support has finite measure. An M_K-constant is usually denoted $v \mapsto c_v$ or $(c_v)_v$. Equivalently, an M_K-constant is a measurable, L^1 function $v \mapsto c_v$ such that, when restricted to nonarchimedean places, $c_v = 0$ for all but finitely many v.

The sum and maximum of two M_K-constants is an M_K-constant, and a (real) constant multiple of an M_K-constant is an M_K-constant.

Since an M_K-constant $(c_v)_v$ is L^1, we have

$$\int_{M_K} |c_v| \, d\mu(v) < \infty \quad \text{and} \quad -\infty < \int_{M_K} c_v \, d\mu(v) < \infty. \tag{3.19}$$

Remark 3.20. Since $-\log|z|$ has finite integral on the unit disc $\mathbb{D}$, the function $v \mapsto -\log\|\alpha\|_v$ is an M_K-constant for all $\alpha \in K^*$. Note, however, that if α is transcendental, then $-\log\|\alpha\|_v$ is not bounded in the usual sense: for all $c \in \mathbb{R}$ there is a $v \in M_K$ such that $-\log\|\alpha\|_v > c$. (This happens near zeroes of α on $B(\mathbb{C})$.)

The reliance on integration and measure theory makes it necessary to assume that the sets and functions encountered are measurable (this trivially holds for the counting measure). Therefore:

In this paper, subsets of M_K of finite measure are always assumed to be measurable.

Also, we define the following.

Definition 3.21. Let V be a variety over K, and let S be a measurable subset of M_K:

(a) The set $V(S)$ is the disjoint union

$$V(S) = \coprod_{v \in S} V(\mathbb{C}_v).$$

In particular,

$$V(M_K) = \coprod_{v \in M_K} V(\mathbb{C}_v).$$

(b) A function $\alpha\colon V(S) \to \mathbb{R}$ is *$\overline{K}$-measurable* if the following condition is true. For all finite extensions L of K, let $\pi_L\colon B_L \to B$ be the normalization of B in L, let π_L^{gen} denote the induced map $B_L(\mathbb{C})^{\text{gen}} \to B(\mathbb{C})^{\text{gen}}$, let $S_L = (\pi_L^{\text{gen}})^{-1}(S)$, and (as usual) let $V_L = V \times_K L$. A rational point $P \in V(L)$ induces a function $\beta_P\colon S_L \to V_L(S_L)$; for all $w \in S_L$, we have a canonical identification of $V_L(\mathbb{C}_w)$ with $V(\mathbb{C}_v)$, where $v = \pi_L^{\text{gen}}(w) \in S$. This identification associates β_P with a function $\beta'_P\colon S_L \to V(S)$. Then the condition is that $\alpha \circ \beta'_P\colon S_L \to \mathbb{R}$ is a measurable function for all L and P as above. (Note that S_L does not contain any nonarchimedean places, but that removing nonarchimedean places from a given set does not affect whether the set is measurable.)

(c) A function $\alpha\colon V(S) \to \mathbb{R}$ is *M-continuous* if it is $\overline{K}$-measurable and if, for all $v \in S$, its restriction to $V(\mathbb{C}_v)$ is continuous in the topology induced by the metric on $\mathbb{C}_v$.

(d) Let $U = \operatorname{Spec} A$ be an open affine in V, let $x_1, \dots, x_n$ be elements of A such that $A = K[x_1, \dots, x_n]$, and let γ be an M_K-constant. Then

$$B_S(U, x_1, \dots, x_n, \gamma) = \{P \in U(S) : \log\|x_i\| \le \gamma_{v(P)} \text{ for all } i\},$$

where $v(P)$ denotes the (unique) $v \in S$ for which $P \in V(\mathbb{C}_v)$.

(e) Let U be as in (d). Then a subset E of $V(S)$ is *affine M-bounded* with respect to U if there exist $x_1, \dots, x_n \in A$ and an M_K-constant γ such that $A = K[x_1, \dots, x_n]$ and $E \subseteq B_S(U, x_1, \dots, x_n, \gamma)$. (This implies $E \subseteq U(S)$.)

(f) A set $E \subseteq V(S)$ is *M-bounded* if there exist open affine subsets $U_1, \dots, U_n$ of V and a decomposition $E = E_1 \cup \dots \cup E_n$ such that E_i is affine M-bounded with respect to U_i for all i.

(g) A function $\alpha\colon V(S) \to \mathbb{R}$ is *locally M-bounded* if it is bounded above and below by M_K-constants on all M-bounded subsets of $V(S)$.

Then Weil functions can be defined, following [Lang 1983, Chapter 10]; see also [Gubler 1997, Section 2]:[2]

Definition 3.22. Let V be a complete variety over K, and let D be a Cartier divisor on V. Then a *Weil function* for D is a function $\lambda_D \colon (V \setminus \operatorname{Supp} D)(M_K) \to \mathbb{R}$ such that, for all open $U \subseteq V$ and all $f \in K(V)^*$ for which $D|_U = \operatorname{div}(f)|_U$, there is an M-continuous, locally M-bounded function $\alpha \colon U(M_K) \to \mathbb{R}$ such that

$$\lambda_D(P) = -\log\|f(P)\|_v + \alpha(P) \quad \text{for all } P \in (U \setminus \operatorname{Supp} D)(M_K),$$

where v is the (unique) place of K for which $P \in U(\mathbb{C}_v)$.

Similarly, for a subset $S \subseteq M_K$, a *partial Weil function* for D over S is a function $\lambda_D \colon (V \setminus \operatorname{Supp} D)(S) \to \mathbb{R}$ that satisfies a similar condition.

For $v \in S$, the restriction of λ_D to $(V \setminus \operatorname{Supp} D)(\mathbb{C}_v)$ is denoted $\lambda_{D,v}$.

The following lemma will be needed in the proof of Proposition 3.28.

Lemma 3.23. *Let V be a variety over K, and let S be a measurable subset of M_K:*

(a) *Let $U = \operatorname{Spec} A$ be an open affine subset of V, and let E be a subset of $U(S)$ which is affine M-bounded with respect to U. Then the condition of Definition 3.21(e) is satisfied for every choice of $x_1, \dots, x_n \in A$ such that $A = K[x_1, \dots, x_n]$.*

(b) *If $U' \subseteq U$ are open affine subsets of V, and if $E \subseteq V(S)$ is affine M-bounded with respect to U', then E is also affine M-bounded with respect to U.*

(c) *Let E be an M-bounded subset of $V(S)$. Then, for all (finite) open affine covers $U_1, \dots, U_n$ of V, there is a decomposition $E = E_1 \cup \dots \cup E_n$ such that E_i is affine M-bounded with respect to U_i for all i.*

(d) *If V is affine, then a subset of $V(S)$ is M-bounded if and only if it is affine M-bounded with respect to V.*

(e) *Let $V_1, \dots, V_n$ be a covering of V by arbitrary open subsets V_i. Then any M-bounded subset E of $V(S)$ has a decomposition $E = E_1 \cup \dots \cup E_n$, in which each E_i is an M-bounded subset of $V_i(S)$. Therefore a function $V(S) \to \mathbb{R}$ is locally M-bounded if and only if its restriction to $V_i(S)$ is locally M-bounded on V_i for all i.*

(f) *Let D be a Cartier divisor on V. Let $\{U_1, \dots, U_n\}$ be a covering of V by open affines, and let $f_1, \dots, f_n \in K(V)^*$ be rational functions such that $D|_{U_i} = \operatorname{div}(f_i)|_{U_i}$ for all i. Then a function $\lambda_D \colon (V \setminus \operatorname{Supp} D)(M_K) \to \mathbb{R}$ is a partial Weil function for D over S if (and only if) for all i it satisfies the condition of Definition 3.22 with U and f replaced by U_i and f_i, respectively.*

[2] Gubler does not require M_K-constants to have support of finite measure. This condition can be omitted for the purposes of this paper.

Proof (sketch). Part (a) amounts to showing that if $x_1, \dots, x_n$ and $y_1, \dots, y_m$ are two systems of generators for A over K, then for each M_K-constant γ there is an M_K-constant γ' such that $B_S(U, x_1, \dots, x_n, \gamma) \subseteq B_S(U, y_1, \dots, y_m, \gamma')$.

For part (b), if $U' = \operatorname{Spec} A'$, $U = \operatorname{Spec} A$, and $A = K[x_1, \dots, x_n]$, then since $A' \supseteq A$, we may use $A' = K[x'_1, \dots, x'_m]$ with $\{x_1, \dots, x_n\} \subseteq \{x'_1, \dots, x'_m\}$.

For part (c), we first claim that the conclusion holds if V is affine and E is affine M-bounded with respect to V. It suffices to prove this case when all U_i are principal open affines $D(f_i)$ in V, in which case we use the existence of $a_1, \dots, a_n \in \mathscr{O}_V(V)$ such that $a_1 f_1 + \dots + a_n f_n = 1$. The general case then follows by reducing to finitely many instances of this special case.

Parts (d) and (e) are immediate from (c).

Finally, part (f) follows from (e), together with the fact that $-\log|f|$ is an M-bounded function on $V(S)$ for all $f \in \mathscr{O}(V)^*$, and the fact that finite sums of M-bounded functions on $V(S)$ are M-bounded. □

For details on parts of the above proof, see [Lang 1983, Chapter 10] or [Gubler 1997, Section 2].

With the definitions from the number field case extended to arithmetic function fields in the above way, the theory of Weil functions follows from [Lang 1983, Chapter 10], where one replaces references to a finite subset of M_K with a subset of M_K of finite measure, and similarly references to "almost all $v \in M_K$" with "all $v \in M_K$ outside a set of finite measure."

In particular, we have the following, in which $O_{M_K}(1)$ refers to a function whose absolute value is bounded by an M_K-constant.

Theorem 3.24. *Let V be a complete variety over an arithmetic function field K. Then*:

(a) ***Additivity***: *If λ_1 and λ_2 are Weil functions for Cartier divisors D_1 and D_2, respectively, on V, then $\lambda_1 + \lambda_2$ (on the intersection of their domains) extends uniquely to a Weil function for $D_1 + D_2$.*

(b) ***Functoriality***: *If λ is a Weil function for a Cartier divisor D on V, and if $f : V' \to V$ is a morphism of varieties over K whose image is not contained in* Supp D, *then $\lambda \circ f$ is a Weil function for f^*D on V'.*

(c) ***Normalization***: *If $V = \mathbb{P}^n_K$ (with $n \in \mathbb{Z}_{>0}$), then the function λ_D defined by*

$$\lambda_{D,v}([x_0 : \dots : x_n]) = -\log \frac{\|x_0\|_v}{\max\{\|x_0\|_v, \dots, \|x_n\|_v\}}$$

for all $v \in M_K$ is a Weil function for the divisor D given by $x_0 = 0$.

(d) ***Uniqueness***: *If both λ_1 and λ_2 are Weil functions for the same Cartier divisor D on V, then $\lambda_1 = \lambda_2 + O_{M_K}(1)$.*

(e) ***Boundedness from below***: *If λ is a Weil function for an effective Cartier divisor D, then λ is bounded from below by an M_K-constant.*

(f) ***Existence***: *If V is projective, then every Cartier divisor on V has a Weil function.* (*For the case in which V is complete, see Remark 3.29.*)

(g) ***Principal divisors***: *For all $f \in K(V)^*$, the function $-\log\|f\|_v$ is a Weil function for the principal divisor (f) on V.*

Proof. Parts (a)–(c) and (g) are easy to see from the definitions. For parts (d) and (e), see [Lang 1983, Chapter 10, Propositions 2.2 and 3.1], together with Chow's lemma. For (f), see [Lang 1983, Chapter 10, Theorem 3.5]. □

Next we show that nonzero rational sections of certain line sheaves can be used to define Weil functions for the associated divisors. We start by defining the construction of such functions in more detail.

Definition 3.25. Let V be a projective variety over K, let $\pi\colon X \to B$ be a model for V with isomorphism $i\colon V \to X_K$, let $\mathscr{L}$ be a continuously metrized line sheaf on X, let s be a nonzero rational section of $\mathscr{L}$, and let $D = i^* \operatorname{div}(s_K)$. Then we define a function

$$\lambda_s\colon (V \setminus \operatorname{Supp} D)(M_K) \to \mathbb{R}$$

as follows:

(i) If v is an infinite place, then it corresponds to a point $b \in B(\mathbb{C})^{\text{gen}}$. Furthermore, $\mathbb{C}_v \cong \mathbb{C}$; up to this choice of isomorphism, we have a canonical isomorphism of $V(\mathbb{C}_v)$ with $\pi^{-1}(b)$. This identifies $V(M_K^\infty)$ with $\pi^{-1}(B(\mathbb{C})^{\text{gen}})$. So if $v \in M_K^\infty$ and $P \in (V \setminus \operatorname{Supp} D)(\mathbb{C}_v)$, then P corresponds to a point $x \in \pi^{-1}(b) \cap (V \setminus \operatorname{Supp} D)(\mathbb{C})$, and we define $\lambda_s(P) = -\log|s(x)|$ (using the metric on $\mathscr{L}$).

(ii) If v is a finite place, then it corresponds to a prime divisor Y on B. Let η be the generic point of Y. Since B is normal, the local ring $\mathscr{O}_{B,\eta}$ is a dvr, whose valuation determines the valuation used to define $\mathbb{C}_v$. A point $P \in (V \setminus \operatorname{Supp} D)(\mathbb{C}_v)$ corresponds to a point $x \in X$ and an injection from its residue field $\kappa(x)$ to $\mathbb{C}_v$ compatible with the injections $\mathscr{O}_{B,\eta} \hookrightarrow K \hookrightarrow \mathbb{C}_v$. (Therefore x actually lies on the generic fiber X_K.) By the valuative criterion of properness, the morphism $\operatorname{Spec} \mathbb{C}_v \to X$ extends to a morphism $h\colon \operatorname{Spec} \mathscr{O}_v \to X$, where $\mathscr{O}_v$ is the valuation ring of $\mathbb{C}_v$. Let $x_0 \in X$ be the image of the closed point of $\operatorname{Spec} \mathscr{O}_v$ under this morphism. Then x_0 is a specialization of x in X.

Now let U be an open neighborhood of x_0 in X such that $\mathscr{L}|_U$ is trivial, and let $s_0 \in \mathscr{L}(U)$ be a section that generates $\mathscr{L}$ over U. Then h^*s_0 generates $h^*\mathscr{L}$ (over all of $\operatorname{Spec} \mathscr{O}_v$), and h^*s is a well-defined nonzero section of $h^*\mathscr{L}$ (because $x \notin \operatorname{Supp} D$). In particular, $h^*s/h^*s_0 \in \mathbb{C}_v^*$, and so we define $\lambda_s(P) = -\log\|h^*s/h^*s_0\|_v$.

This value is independent of the choices of U and s_0. Indeed, suppose that U' and s_0' are a different set of such choices. Then h^*s_0 and h^*s_0' both generate $h^*\mathscr{L}$ at the special point, so $h^*s_0'/h^*s_0 \in \mathscr{O}_v^*$, so $\|h^*s_0'/h^*s_0\|_v = 1$ and therefore $\|h^*s/h^*s_0\|_v = \|h^*s/h^*s_0'\|_v$.

We also let $-\log\|s\|$ denote λ_s, so $\lambda_s(P) = -\log\|s(P)\|_v$ for all $v \in M_K$ and all $P \in (V \setminus \operatorname{Supp} D)(\mathbb{C}_v)$.

Lemma 3.26. *Let n be a positive integer, let $V = \mathbb{P}^n_K$, and let $X = \mathbb{P}^n_B$, so that X is a model for V. Let $\mathscr{L}'$ be the line sheaf $\mathscr{O}(1)$ on X, with continuous metric uniquely determined by the condition that the* *metric of a global section* $s = a_0x_0 + \cdots + a_nx_n$ *at a point* $P = [p_0 : \cdots : p_n]$ *is given by*

$$\|s\|(P) = \frac{\|a_0p_0 + \cdots + a_np_n\|_v}{\max\{\|p_0\|_v, \ldots, \|p_n\|_v\}} \tag{3.26.1}$$

(*this generalizes the metric of Proposition 3.17*). *Let* s' *be the global section* x_0 *of* $\mathscr{L}'$. *Let* $D = \operatorname{div}(s')_K$ (*the hyperplane at infinity on* $V = \mathbb{P}^n_K$). *Then* $\lambda_{s'} = -\log\|s'\|$ *is a Weil function for* D.

Proof. By Lemma 3.23(f), it suffices to check the condition of Definition 3.22 on the standard open affines $U_i = D_+(x_i)$ with $f_i = x_0/x_i$, for $i = 0, \ldots, n$.

First we consider $i = 0$. Then f_0 is the constant function 1, and (in the notation of Definition 3.22) $\alpha = \lambda_{s'}$ (note that $U_0 \setminus \operatorname{Supp} D = U_0$). We write $U_0 = \operatorname{Spec} K[y_1, \ldots, y_n]$, where $y_i = x_i/x_0$ for all i. For all $v \in M_K$, the value of $\lambda_{s'}$ at a point $P = [p_0 : \cdots : p_n] \in U_0(\mathbb{C}_v)$ is

$$\lambda_{s'}(P) = -\log\frac{\|p_0\|_v}{\max\{\|p_0\|_v, \ldots, \|p_n\|_v\}} = \log\max\{1, \|y_1(P)\|_v, \ldots, \|y_n(P)\|_v\}. \tag{3.26.2}$$

Indeed, for infinite v this holds by (3.26.1). For finite v, choose j such that

$$\max\{\|p_0\|_v, \ldots, \|p_n\|_v\} = \|p_j\|_v.$$

Then, in the notation of Definition 3.25, we may take $s_0 = x_j$, so

$$\|s'/s_0\|_v = \|(x_0/x_j)(P)\|_v = \frac{\|p_0\|_v}{\max\{\|p_0\|_v, \ldots, \|p_n\|_v\}},$$

and again we obtain (3.26.2).

The right-hand side of (3.26.2) is obviously continuous on $U_0(\mathbb{C}_v)$ for all v, and it is M-bounded below because it is always nonnegative. It is M-bounded above because for all M_K-constants γ we have $\lambda_{s'}$ bounded above by γ on $B_{M_K}(U_0, 1, y_1, \ldots, y_n, \gamma)$, by (3.26.2), Definition 3.21(d), and Lemma 3.23(a).

For $i \ne 0$, by symmetry it suffices to consider the case $i = n$. We have

$$U_n = \operatorname{Spec} K[y_0, y_1, \ldots, y_{n-1}],$$

where $y_i = x_i/x_n$ for all $i = 0, \ldots, n-1$. We have $f_n = x_0/x_n = y_0$, so

$$\begin{aligned}\alpha(P) &= \lambda_{s'}(P) + \log\|y_0(P)\|_v \\ &= -\log\frac{\|p_0\|_v}{\max\{\|p_0\|_v, \ldots, \|p_n\|_v\}} + \log\frac{\|p_0\|_v}{\|p_n\|_v} \\ &= \log\max\{\|y_0(P)\|_v, \ldots, \|y_{n-1}(P)\|_v, 1\}\end{aligned}$$

for all $P = [p_0 : \cdots : p_n] \in U_n(\mathbb{C}_v)$ and all $v \in M_K$. This is M-continuous and M-bounded for the same reasons as before.

Thus $\lambda_{s'}$ is a Weil function for D. □

Lemma 3.27. *Let* V *be a projective variety over* K, *and let* $\mathscr{L}$ *be a line sheaf on* X. *Then there exist a model* $\pi : X \to B$ *for* V *with isomorphism* $i : V \to X_K$, *a continuously metrized line sheaf* $\mathscr{L}'$ *on* X *that extends* $\mathscr{L}$, *and a nonzero rational section* s' *of* $\mathscr{L}'$, *such that* $\lambda_{s'}$ *is a Weil function for* $i^*\operatorname{div}(s')_K$.

Proof. We first prove this in the case where $\mathscr{L}$ is very ample.

Let $j\colon V \to \mathbb{P}^n_K$ be a closed immersion over K such that $\mathscr{L} \cong j^*\mathscr{O}(1)$. We may assume that $n > 0$ and the image of j is not contained in the hyperplane $x_0 = 0$. Let X be the closure of the image of j in $\mathbb{P}^n_B$, and let $\mathscr{L}'$ be the sheaf $\mathscr{O}(1)$ on X. Then X is a model for V and $\mathscr{L}'$ extends $\mathscr{L}$. Finally, let s' be the restriction of the section x_0 of $\mathscr{O}(1)$ to X. Since $j(V)$ is not contained in the hyperplane at infinity, s' is nonzero.

Then the lemma holds in this case by Theorem 3.24(b) and compatibility of $\operatorname{div}(\cdot)_K$ with pull-back.

We now consider the general case.

An arbitrary line sheaf $\mathscr{L}$ on V can be written as $\mathscr{L} \cong \mathscr{L}_1 \otimes \mathscr{L}_2^\vee$, where $\mathscr{L}_1$ and $\mathscr{L}_2$ are very ample on V. By the previous special case, for $\ell = 1, 2$ there exist projective models $\pi_\ell\colon X_\ell \to B$ for V over B, continuously metrized line sheaves $\mathscr{L}'_\ell$ on X_ℓ extending $\mathscr{L}_\ell$, and nonzero rational sections s'_ℓ of $\mathscr{L}'_\ell$ such that $-\log\|s'_\ell\|_v$ are Weil functions for $i^* \operatorname{div}(s'_\ell)_K$.

Let X be a projective model for V that dominates X_1 and X_2 (e.g., one can let X be the closure of the graph of the isomorphism $(X_1)_K \xrightarrow{\sim} (X_2)_K$ in $X_1 \times_B X_2$). After pulling back the $\mathscr{L}'_\ell$ to X, we may assume that $X_1 = X_2 = X$. Letting $\mathscr{L}' = \mathscr{L}'_1 \otimes \mathscr{L}_2'^\vee$ and $s' = s'_1/s'_2$, we have that $\mathscr{L}'$ extends $\mathscr{L}$, s' is a nonzero rational section of $\mathscr{L}'$, and $-\log\|s'\|_v = -\log\|s'_1\|_v + \log\|s'_2\|_v$ is a Weil function for $i^* \operatorname{div}(s')_K = i^* \operatorname{div}(s'_1)_K - i^* \operatorname{div}(s'_2)_K$ on V, by Theorem 3.24(a). ☐

Proposition 3.28. *Let $\pi : X \to B$ be a dominant morphism of arithmetic varieties (i.e., a model for X_K), let $\mathscr{L}$ be a continuously metrized line sheaf on X, and let s be a nonzero rational section of $\mathscr{L}$. Then $\lambda_s = -\log\|s\|$ is a Weil function for* $\operatorname{div}(s)_K$.

Proof. Let $V = X_K$. By Lemma 3.27 there exist a model X' for V, a line sheaf $\mathscr{L}'$ on X' extending $\mathscr{L}_K$, and a nonzero rational section s' of $\mathscr{L}'$, such that $\lambda_{s'}$ is a Weil function for $\operatorname{div}(s')_K$.

We may assume that X' dominates X (replace X' with the closure of the graph of $V \xrightarrow{\sim} X'_K$ in $X' \times_B X$), so there exists a proper birational morphism $p\colon X' \to X$ inducing an isomorphism $X'_K \xrightarrow{\sim} X_K$.

Then $\mathscr{L}'_K \cong p^*\mathscr{L}_K$, so the nonzero rational section s'/p^*s of $\mathscr{L}' \otimes p^*\mathscr{L}^\vee$ corresponds to an element $\alpha \in K(V)^*$. Moreover,

$$\operatorname{div}(s')_K - \operatorname{div}(p^*s)_K = (\alpha) \quad \text{on } X'_K. \tag{3.28.1}$$

Let $\mathscr{M}$ be the metrized line sheaf $\mathscr{L}' \otimes p^*\mathscr{L}^\vee$ on X', and let $t = s'/\alpha p^*s$. Then t is a nonzero rational section of $\mathscr{M}$ whose restriction to $\mathscr{M}_K$ is a global section that generates $\mathscr{M}_K$ everywhere. Therefore we have $\mathscr{M}_K \cong \mathscr{O}_{X'_K}$, hence $\mathscr{M}_{\text{fin}} \cong \mathscr{O}(E)$ for some Cartier divisor E on X' supported only on fibers of $X' \to B$. In particular there are M_K-constants γ and γ' such that $\gamma \le \lambda_t \le \gamma'$ everywhere on $V(M_K)$ (these M_K-constants may be taken to be constants on M_K^∞, by compactness of $X'(\mathbb{C})$).

Thus λ_t is a Weil function for the trivial divisor on V. By additivity, we have

$$\lambda_{p^*s} = \lambda_{s'} + \log\|\alpha\| - \lambda_t,$$

and this is a Weil function for $\operatorname{div}(s')_K - (\alpha)_K = \operatorname{div}(p^*s)_K$ by (3.28.1), Theorem 3.24(a) and (g). Since p induces an isomorphism $X'_K \xrightarrow{\sim} X_K$, we may identify X'_K with X_K to obtain $X'_K(M_K) = X_K(M_K)$, $\lambda_{p^*s} = \lambda_p$, and $\operatorname{div}(p^*s)_K = \operatorname{div}(s)_K$; thus λ_s is a Weil function for $\operatorname{div}(s)_K$. □

Remark 3.29. More generally, let X be an integral scheme, let $\pi : X \to B$ be a proper morphism, let $\mathscr{L}$ be a continuously metrized line sheaf on X, and let s be a nonzero rational section of $\mathscr{L}$. Definition 3.25 extends easily to this situation, giving a real-valued function $\lambda_s = -\log\|s\|$ on $(X_K \setminus \operatorname{Supp} D)(M_K)$, where $D = \operatorname{div}(s)_K$. Then the above proposition can be extended to this situation. Indeed, by Chow's lemma there is a proper birational morphism $\phi\colon X' \to X$ such that X' is projective over B, so $\phi^*\lambda_s$ is a Weil function for ϕ^*D. It then follows that λ_s is a Weil function for D, because if $f\colon X_K(M_K) \to \mathbb{R}$ is a function such that $f \circ \phi$ is M-bounded, then f is also M-bounded. Then Theorem 3.24(f) can be generalized to complete varieties V over K, as follows. Given a complete variety V over K, there exists X as above with $X_K \cong V$ over K by Nagata's embedding theorem; moreover X can be chosen such that $\mathscr{O}(D)$ extends to a line sheaf $\mathscr{L}$ on X; see [Vojta 2007]. Let s be the extension of the canonical section of $\mathscr{O}(D)$ to $\mathscr{L}$. Then λ_s is a Weil function for D. This fact is not needed in this paper, though, so the details are left to the reader.

Weil functions can be extended to finite extensions of arithmetic function fields (with polarizations as in Definition 3.9) in much the same way as for number fields. Indeed, let K' be a finite extension of K, and let $M = (B; \mathscr{M}_1, \dots, \mathscr{M}_d)$ and $M' := (B'; \mathscr{M}'_1, \dots, \mathscr{M}'_d)$ be as in Definition 3.9. Let $w \in M_{K'}$, and let $v \in M_K$ be the place lying under it. Let V be a complete variety over K, and recall that $V_{K'} = V \times_K K'$. Then there is a natural bijection $\iota_{w/v}\colon V_{K'}(\mathbb{C}_w) \xrightarrow{\sim} V(\mathbb{C}_v)$. Let D be a Cartier divisor on V, let λ_D be a Weil function for D, and let

$$\lambda_{D',w} = n_{w/v}\lambda_{D,v} \circ \iota_{w/v} \tag{3.30}$$

for all $w \in M_{K'}$ and all $v \in M_K$ with $w \mid v$, where $n_{w/v}$ is as in Proposition 3.11. Then $\lambda_{D'}$ is a Weil function for the pull-back D' of D to $V_{K'}$. Moreover, by (3.11.3) and functoriality of pull-back of polarizations to finite extension fields, this construction is functorial in towers of finite extensions of K.

This allows us to define proximity and counting functions for complete varieties over arithmetic function fields, as follows.

Definition 3.31. Let S be a subset of M_K of finite measure, let K' be a finite extension of K with the polarization M' induced by the polarization M of K, and let

$$S' = \{w \in M_{K'} : w \mid v \text{ for some } v \in S\}.$$

Let D be a Cartier divisor on a complete variety V over K. Let $V_{K'}$, D', λ_D, and $\lambda_{D'}$ be as above. Then the *proximity function* and *counting function* for D relative to S are defined by

$$m_S(D, P) = \frac{1}{[K' : K]} \int_{S'} \lambda_{D',w}(P)\, d\mu(w) \tag{3.31.1}$$

and

$$N_S(D, P) = \frac{1}{[K' : K]} \int_{M_{K'} \setminus S'} \lambda_{D',w}(P)\, d\mu(w), \tag{3.31.2}$$

respectively, for all $P \in (V \setminus \operatorname{Supp} D)(K')$. By functoriality of (3.30) in towers, these quantities are independent of the choice of K'.

Combining these definitions leads to a height function

$$h_\lambda(P) = m_S(D, P) + N_S(D, P) = \frac{1}{[K' : K]} \int_{M_{K'}} \lambda_{D',w}(P)\, d\mu(w) \tag{3.31.3}$$

for all $P \in (V \setminus \operatorname{Supp} D)(K')$. By the method of [Lang 1983, Chapter 10, Section 4], this can be extended to give a height function $h_\lambda \colon V(\bar{K}) \to \mathbb{R}$. Indeed, choose a function $f \in K(V)^*$ such that $P \notin \operatorname{Supp}(D + (f))$, and let $\lambda_f = \lambda_D - \log\|f\|$. Then λ_f is a Weil function for $D + (f)$, so we define $h_\lambda(P) = h_{\lambda_f}(P)$, where the latter is defined as in (3.31.3). This is independent of the choice of f, because if $g \in K(V)^*$ also satisfies $P \notin \operatorname{Supp}(D + (g))$, then the rational function f/g extends to a rational function $\alpha \in K(V)^*$ which is regular and nonzero at P, and $\lambda_f - \lambda_g = -\log\|\alpha\|$, so $h_{\lambda_f}(P) - h_{\lambda_g}(P) = 0$ by the product formula (3.5) applied to $\alpha(P) \in K'^*$.

As is true in the number field case, Theorem 3.24(d) and (3.19) imply that the above definitions are independent of the choice of Weil functions, up to $O(1)$.

The next two propositions show that this height is the same (up to $O(1)$) as the height defined by Moriwaki (Definition 3.14), and relate the height defined by Weil functions on $\mathbb{P}^1$ to the naïve height (3.6).

Proposition 3.32. *Let V be a projective variety over K, and let $\mathscr{L}$ be a line sheaf on V. Let X be a model for V over B such that $\mathscr{L}$ extends to a continuously metrized line sheaf $\mathscr{L}'$ on X:*

(a) *Let s be a nonzero rational section of $\mathscr{L}'$, and let $\lambda = \lambda_s$ (Definition 3.25). Then*

$$h_\lambda(P) = h_{\mathscr{L}'}(P) \quad \text{for all } P \in V(\bar{K}). \tag{3.32.1}$$

(b) *If D is a Cartier divisor on V such that $\mathscr{O}(D) \cong \mathscr{L}$, and λ_D is a Weil function for D, then*

$$h_{\lambda_D}(P) = h_{\mathscr{L}'}(P) + O(1) \quad \text{for all } P \in V(\bar{K}). \tag{3.32.2}$$

Proof. We first consider part (a). By Definition 3.31 and Proposition 3.15, it suffices to prove (3.32.1) for all $P \in X(K)$.

Let t be a nonzero rational section of $\mathscr{L}'$ which is regular and nonzero at P, let $\sigma \colon B \dashrightarrow X$ be the rational section of $\pi \colon X \to B$ corresponding to P, and let $\bar{P}$ denote the closure of P in X. By Definition 3.14, the projection formula, and Lemma 1.11,

$$\begin{aligned} h_{\mathscr{L}'}(P) &= c_1(\pi^*\mathscr{M}_1|_{\bar{P}}) \cdots c_1(\pi^*\mathscr{M}_d|_{\bar{P}}) \cdot c_1(\mathscr{L}'|_{\bar{P}}) \\ &= c_1(\mathscr{M}_1) \cdots c_1(\mathscr{M}_d) \cdot \pi_* c_1(\mathscr{L}'|_{\bar{P}}) \\ &= \sum_{Y \in B^{(1)}} \operatorname{ord}_Y(\sigma^* t) c_1(\mathscr{M}_1|_Y) \cdots c_1(\mathscr{M}_1|_Y) + \int_{B(\mathbb{C})^{\text{gen}}} (-\log\|\sigma^* t\|) c_1(\mathscr{M}_1) \wedge \cdots \wedge c_1(\mathscr{M}_d). \end{aligned}$$

Note that, since B is normal, the rational section σ is regular at the generic points of all prime divisors Y on B, so $\operatorname{ord}_Y(\sigma^*t)$ is defined. Moreover, if $v \in M_K^0$ corresponds to Y, then $\|\cdot\|_Y$ as defined in (3.3) agrees with $\|\cdot\|_v$ on $\mathbb{C}_v$ (by definition of $\mathbb{C}_v$). Therefore, by (3.2), (3.3), and Definition 3.25,

$$\operatorname{ord}_Y(\sigma^*t)c_1(\mathscr{M}_1|_Y) = -\log\|(t/t_0)(P)\|_v = \lambda_{t,v}(P),$$

where t_0 is a local generator of $\sigma^*\mathscr{L}'$ at the generic point of Y. By (3.4), Definition 3.25, and (3.31.3), we then have

$$h_{\mathscr{L}'}(P) = \int_{M_K} \lambda_{t,v}(P)\,d\mu(v) = h_{\lambda_t}(P).$$

Since $h_{\lambda_t} = h_{\lambda_s}$ (see the end of Definition 3.31), this gives (3.32.1).

To prove (3.32.2), it suffices by (3.32.1) to show that $h_{\lambda_D}(P) = h_\lambda(P) + O(1)$ for all $P \in V(\bar{K})$, where λ is defined by letting s be the rational section of $\mathscr{L}'$ corresponding to the canonical section of $\mathscr{O}(D)$.

With this choice of s, λ is a Weil function for the same divisor D, so $|\lambda_D - \lambda| \le \gamma$ for some M_K-constant γ by Theorem 3.24(d). Then, for all finite extensions K' of K and all $P \in V(K')$,

$$|h_{\lambda_D}(P) - h_\lambda(P)| \le \frac{1}{[K':K]}\int_{M_{K'}} \gamma = \int_{M_K} \gamma = O(1),$$

where γ is extended to an $M_{K'}$-constant as in (3.30). This implies (3.32.2). □

Proposition 3.33. *Let λ_D be a Weil function for a divisor D on $\mathbb{P}^1_K$. Then*

$$h_{\lambda_D}(P) = (\deg D)h_K(P) + O(1)$$

for all $P \in \mathbb{P}^1(\bar{K})$.

Proof. Let $\mathscr{L} = \mathscr{O}(D)$ on $\mathbb{P}^1_K$. Let $X = \mathbb{P}^1_B$, and let $\mathscr{L}'$ be the line sheaf $\mathscr{O}(\deg D)$ on X, with metric obtained from the metric of Proposition 3.17 by the isomorphism $\mathscr{O}(\deg D) \cong \mathscr{O}(1)^{\otimes(\deg D)}$. Then $\mathscr{L}'$ extends $\mathscr{L}$ to X.

Therefore, by Proposition 3.32(b), (3.14.1), multilinearity of the intersection product, and by Proposition 3.17,

$$h_{\lambda_D}(P) = h_{\mathscr{L}'}(P) + O(1) = (\deg D)h_{\mathscr{O}(1)}(P) + O(1) = (\deg D)h_K(P) + O(1). \qquad \square$$

4. Roth's theorem

This section discusses several equivalent formulations of Roth's theorem, as well as the reasons why certain choices have been made in extending Roth's theorem to arithmetic function fields.

We also show that all of these variants are equivalent (i.e., can be proved from one another by relatively short arguments).

4.1. Throughout this section, K is an arithmetic function field, $M := (B; \mathscr{M}, \dots, \mathscr{M})$ is a big polarization of K with all metrized line sheaves equal to the same smoothly metrized line sheaf $\mathscr{M}$, M_K is derived from this polarization, and S is a subset of M_K with finite measure. We also write M as $(B; \mathscr{M})$.

Note that, by Proposition 2.5, if a polarization of a field K is big, then so is the induced polarization of a finite extension K' of K. Also, the set of places of K' lying over places in S has finite measure. Therefore 4.1 is preserved under passing to the induced polarization of a finite extension.

We start with a definition.

Definition 4.2. Let D be an effective divisor on a nonsingular variety V over a field K. We say that D is *reduced* if all components in Supp D occur with multiplicity 1.

The first version of Roth's theorem is stated using notation from Nevanlinna theory.

Theorem 4.3. *Let K, M_K, and S be as in 4.1; let D be a reduced effective divisor on $\mathbb{P}^1_K$; let $m_S(D, \cdot)$ be the proximity function associated to some choice of Weil function for D; let $\epsilon > 0$; and let $c \in \mathbb{R}$. Then the inequality*

$$m_S(D, \xi) \le (2+\epsilon)h_K(\xi) + c \tag{4.3.1}$$

holds for all but finitely many $\xi \in K$.

The next version of the theorem is close to the above formulation (see the equivalence proof, below) but avoids Weil functions.

Theorem 4.4. *Let K, M_K, and S be as in 4.1; let $\alpha_1, \dots, \alpha_q$ be distinct elements of K; let $\epsilon > 0$; and let $c \in \mathbb{R}$. Then the inequality*

$$\int_S \left(\sum_{j=1}^{q} -\log^- \|\xi - \alpha_j\|_v \right) d\mu(v) \le (2+\epsilon)h_K(\xi) + c \tag{4.4.1}$$

holds for all but finitely many $\xi \in K$.

Next, the following version is close to the preceding version, and is the statement that will be proved in this paper.

Theorem 4.5. *Let K, M_K, and S be as in 4.1; let $\alpha_1, \dots, \alpha_q$ be distinct elements of K; let $\epsilon > 0$; and let $c \in \mathbb{R}$. Then the inequality*

$$\int_S \max\{0, -\log\|\xi - \alpha_1\|_v, \dots, -\log\|\xi - \alpha_q\|_v\} \, d\mu(v) \le (2+\epsilon)h_K(\xi) + c \tag{4.5.1}$$

holds for all but finitely many $\xi \in K$.

Finally, we consider a version that is close to Roth's original theorem.

Theorem 4.6. *Let K, M_K, and S be as in 4.1, and let $\alpha_1, \dots, \alpha_q$ be distinct elements of $\overline{K}$. Choose embeddings $\iota_{v,j}\colon K(\alpha_j) \hookrightarrow \overline{K}_v$ over K for all $j = 1, \dots, q$ and all $v \in S$ in such a way that the function $v \mapsto -\log^- \|\iota_{v,j}(\xi - \alpha_j)\|_v$ is a measurable function for all j and all $\xi \in K \setminus \{\alpha_1, \dots, \alpha_q\}$. Assume also that $\iota_{v,j}(\alpha_j) \ne \iota_{v,j'}(\alpha_{j'})$ for all v and all $j \ne j'$ (this is automatically true unless α_j and $\alpha_{j'}$ are conjugate over K). Then, for all $\epsilon > 0$ and all $c \in \mathbb{R}$, the inequality*

$$\int_S \left(\sum_{j=1}^{q} -\log^- \|\iota_{v,j}(\xi - \alpha_j)\|_v \right) d\mu(v) \le (2+\epsilon)h_K(\xi) + c \tag{4.6.1}$$

holds for all but finitely many $\xi \in K$.

Remark 4.7. Roth's theorem over number fields is often stated in the form of Theorem 0.1, involving choices of $\alpha_v \in \overline{\mathbb{Q}}$ for all $v \in S$. This leads to the question of whether a more natural generalization would be to choose a function $\alpha\colon S \to \overline{K}$ and then bound $\int_S(-\log^-\|\xi - \alpha(v)\|_v)\,d\mu(v)$. I doubt that this is true, although I do not have a counterexample. I believe that Theorems 4.3–4.6 represent a more natural generalization, because they correspond more closely to Nevanlinna theory, and because they are sufficient to prove Siegel's theorem on integral points (Corollary 4.11). (If the image of α is required to be finite, then this is strictly weaker than Theorem 4.6, since the function does not depend on ξ. One can fix a finite subset T of $\overline{K}$, though, and allow α to be a function from S to T depending on ξ. This would then be equivalent to Theorem 4.6.)

We now show that these four theorems are all equivalent, and therefore proving any one of them suffices to prove all four.

Proposition 4.8. *Theorems 4.3–4.6 are equivalent.*

Proof. We first show that Theorems 4.3 and 4.4 are equivalent. Let $\alpha_1, \dots, \alpha_q$ be as in the statement of Theorem 4.4. By Proposition 3.28, for fixed $\alpha \in K$ the function $\xi \mapsto -\log^-\|\xi - \alpha\|_v$ defines a Weil function for the divisor $[\alpha]$ on $\mathbb{P}^1$. By additivity of Weil functions, the integrand in (4.4.1) defines a Weil function for a divisor $D := [\alpha_1] + \cdots + [\alpha_q]$; hence the left-hand side of (4.4.1) equals $m_S(D, \xi)$ for this choice of Weil function, so (4.4.1) and (4.3.1) are equivalent.

This shows that Theorem 4.3 implies Theorem 4.4. It does not (yet) show the converse, though, since not all reduced effective divisors D on $\mathbb{P}^1_K$ are of the above form.

To show the converse, let D be a reduced effective divisor on $\mathbb{P}^1_K$. We first consider the case in which $\infty \notin \operatorname{Supp} D$.

Let K' be a finite Galois extension of K such that all points in $\operatorname{Supp} D$ are rational over K', let S' be the subset of $M_{K'}$ lying over S (as in Definition 3.31), and let D' be the pull-back of D to $\mathbb{P}^1_{K'}$. The proximity function $m_S(D, \xi)$ in (4.3.1) was defined using a specific choice of Weil function for D; let this be extended to a Weil function for D' on $\mathbb{P}^1_{K'}$ as in (3.30). We then have $m_{S'}(D', \xi) = [K' : K] m_S(D, \xi)$ and $h_{K'}(\xi) = [K' : K] h_K(\xi)$ for all $\xi \in K \setminus \operatorname{Supp} D$. Therefore Theorem 4.3 for D' on $\mathbb{P}^1_{K'}$ implies Theorem 4.3 for D on $\mathbb{P}^1_K$. Since all points in $\operatorname{Supp} D'$ are rational over K', Theorem 4.3 for D' follows from Theorem 4.4 applied over K'. Therefore Theorem 4.3 also holds for D.

To drop the assumption $\infty \notin \operatorname{Supp} D$, let ϕ be an automorphism of $\mathbb{P}^1_K$ such that $\phi(\infty) \notin \operatorname{Supp} D$. One can use the pull-back via ϕ of a Weil function for D to give a Weil function for $\phi^* D$; we then have $m_S(\phi^* D, \xi) = m_S(D, \phi(\xi))$ for all $\xi \in \mathbb{P}^1_K \setminus \operatorname{Supp} \phi^* D$. Also $h_K(\xi) = h_K(\phi(\xi)) + O(1)$ for all ξ by Proposition 3.33 (let D' be any divisor on $\mathbb{P}^1$ of degree 1, and note that $\deg \phi^* D' = 1$ also). Therefore Theorem 4.3 for $\phi^* D$ implies Theorem 4.3 for D. Since the former follows from Theorem 4.4, it follows that Theorems 4.3 and 4.4 are equivalent.

We next show that Theorems 4.4 and 4.5 are equivalent. Let $\alpha_1, \dots, \alpha_q$ be distinct elements of K, and let $D = [\alpha_1] + \cdots + [\alpha_q]$. For fixed $\alpha \in K$,

$$\max\{0, -\log\|\xi - \alpha\|_v\} = -\log^-\|\xi - \alpha\|_v$$

gives a Weil function for the divisor $[\alpha]$. Therefore, by [Lang 1983, Chapter 10, Proposition 3.2] (applied with $Y = -D$, and using the fact that the theory of Weil functions carries over directly to arithmetic function fields), the integrand in (4.5.1) is a Weil function for D, and therefore the left-hand sides of (4.4.1) and (4.5.1) differ by $O(1)$. Thus Theorems 4.4 and 4.5 are equivalent.

Finally, we show that Theorem 4.6 is equivalent to the other three. Clearly Theorem 4.6 reduces to Theorem 4.4 in the case when all α_i lie in K, so Theorem 4.6 implies Theorem 4.4.

For the converse, let $\alpha_1, \dots, \alpha_q \in \bar{K}$ and $\iota_{j,v}\colon K(\alpha_j) \hookrightarrow \bar{K}_v$ ($1 \le j \le q$, $v \in S$) be as in the statement of Theorem 4.6. Since the inequality (4.6.1) is strengthened by adding more elements to $\{\alpha_1, \dots, \alpha_q\}$, we may assume that this set is invariant under $\operatorname{Gal}(\bar{K}/K)$. (When doing this, it is possible to choose the embeddings for the added elements in a way that satisfies the condition on measurability.) Then $K' := K(\alpha_1, \dots, \alpha_q)$ is a finite Galois extension of K. The map $\{1, \dots, q\} \to \{\alpha_1, \dots, \alpha_q\}$ given by $j \mapsto \iota_{v,j}(\alpha_j)$ is injective, hence bijective; therefore

$$\sum_{j=1}^{q} -\log^- \|\iota_{v,j}(\xi - \alpha_j)\|_v = \sum_{j=1}^{q} -\log^- \|\iota_v(\xi - \alpha_j)\|_v \tag{4.8.1}$$

for all $v \in S$, all $\iota_v\colon K' \to K_v$ over K, and all $\xi \in K \setminus \{\alpha_1, \dots, \alpha_q\}$.

We will show that Theorem 4.4, with K replaced by K', S replaced by the set S' of all places of K' lying over places of S, and c replaced by $[K' : K]c$, implies Theorem 4.6 (with no replacements). Indeed, let μ' denote the measure on $M_{K'}$ associated to the polarization of K' induced by the polarization of K. This is compatible with the measure μ on M_K; combining this with (3.11.3) and (4.8.1) gives

$$\begin{aligned}\int_{S'} \biggl(\sum_{j=1}^{q} -\log^- \|\xi - \alpha_j\|_w\biggr)\, d\mu'(w) &= \int_{S} \biggl(\sum_{j=1}^{q} \sum_{w \mid v} -\log^- \|\xi - \alpha_j\|_w\biggr)\, d\mu(v)\\ &= [K' : K] \int_{S} \biggl(\sum_{j=1}^{q} -\log^- \|\iota_v(\xi - \alpha_j)\|_v\biggr)\, d\mu(v)\\ &= [K' : K] \int_{S} \biggl(\sum_{j=1}^{q} -\log^- \|\iota_{v,j}(\xi - \alpha_j)\|_v\biggr)\, d\mu(v).\end{aligned}$$

Combining this with (3.11.4) then gives that (4.6.1) is equivalent to (4.4.1) (with the above replacements). □

Remark 4.9. For the equivalence of Theorems 4.4 and 4.5, something stronger was actually proved. The above proof additionally showed that, for any given K, M_K, S, $\alpha_1, \dots, \alpha_q$, and ϵ, Theorem 4.4 for all c is equivalent to Theorem 4.5 for all c. This fact will be used in the proof of Proposition 5.7, below.

As is true over number fields, Roth's theorem and Mordell's conjecture imply the author's "Main Conjecture" [Vojta 1987, Conjecture 3.4.3] in the special case of (rational points on) curves. This is proved by essentially the same proof as over number fields, so the proof will only be sketched.

Corollary 4.10. *Let X be a smooth projective curve over K of genus g; let D be a reduced effective divisor on X; let $\mathscr{A}$ be a line sheaf of degree 1 on X; let $m_S(D, \cdot)$ and $h_{\mathscr{A}}(\cdot)$ be the proximity and height* *functions, respectively, determined by some fixed choice of Weil function for* D *and* $\mathscr{A}$, *respectively*; *let* $\epsilon > 0$; *and let* $c \in \mathbb{R}$. *Then the inequality*

$$m_S(D, \xi) \le (2 - 2g + \epsilon) h_{\mathscr{A}}(\xi) + c \tag{4.10.1}$$

holds for all but finitely many $\xi \in X(K)$.

Proof (sketch). When $g = 0$ this is Theorem 4.3, and when $g > 1$ this follows from Mordell's conjecture over K (see the Introduction) since $X(K)$ is finite. This leaves the case $g = 1$. In this case, (4.10.1) reduces to $m_S(D, \xi) \le \epsilon\, h_{\mathscr{A}}(\xi) + c$.

As in the proof of Proposition 4.8, we may assume that all points of D are rational over K. We may also assume that $D \ne 0$, so in particular $X(K) \ne \varnothing$. Thus X is an elliptic curve.

Assume that the statement is false. Then the inequality

$$m_S(D, \xi) > \epsilon\, h_{\mathscr{A}}(\xi) + c \tag{4.10.2}$$

holds for infinitely many $\xi \in X(K)$.

Following [Lang 1960], fix an integer $n > 2/\sqrt{\epsilon}$. Since the Mordell–Weil theorem is known for $X(K)$ (see the Introduction), the subgroup $nX(K)$ is of finite index in $X(K)$. Therefore some coset $\xi_0 + nX(K)$ contains infinitely many points ξ for which (4.10.2) holds. Let $\phi\colon X \to X$ be the K-morphism $\xi \mapsto n\xi + \xi_0$. Then, for some constant c', the inequality

$$m_S(\phi^* D, \xi') > \epsilon\, h_{\phi^*\mathscr{A}}(\xi') + c'$$

holds for infinitely many $\xi' \in X(K)$.

Pick a morphism $\psi\colon X \to \mathbb{P}^1_K$ over K of degree 2. Let D' be the reduced divisor on $\mathbb{P}^1_K$ whose support is $\psi(\operatorname{Supp}\phi^* D)$. Since ϕ is étale, the divisor $\phi^* D$ is reduced (as well as effective). Therefore the divisor $\psi^* D' - \phi^* D$ is effective, so $m_S(\psi^* D', \xi') \ge m_S(\phi^* D, \xi') + O(1)$ for all $\xi' \in X(K)$. In addition, $\phi^*\mathscr{A}$ and $\psi^*\mathscr{O}(1)$ have degrees n^2 and 2, respectively; therefore, for any ϵ'' such that $2 + \epsilon'' < n^2\epsilon/2$, standard properties of heights (which extend straightforwardly to arithmetic function fields) give

$$\epsilon\, h_{\phi^*\mathscr{A}}(\xi') \ge (2 + \epsilon'') h_{\psi^*\mathscr{O}(1)}(\xi') + O(1) = (2 + \epsilon'') h_K(\psi(\xi')) + O(1).$$

By choice of n, we may take $\epsilon'' > 0$. Thus, up to $O(1)$ at each step,

$$m_S(D', \psi(\xi')) = m_S(\psi^* D', \xi') \ge m_S(\phi^* D, \xi') > \epsilon\, h_{\phi^*\mathscr{A}}(\xi') \ge (2 + \epsilon'') h_K(\psi(\xi')).$$

This holds for infinitely many points $\psi(\xi')$ in $\mathbb{P}^1(K)$, which contradicts Theorem 4.3. □

This leads, in the usual way, to Siegel's theorem on integral points on curves, due to [Lang 1960, Theorem 4]; see also [Lang 1991, Chapter IX Theorem 3.1]:

Corollary 4.11. *Let K be a field finitely generated over $\mathbb{Q}$, let R be a subring of K finitely generated over $\mathbb{Z}$, and let C be an affine curve over K. Assume that either none of the irreducible components of $C \times_K \overline{K}$ are rational, or that there exists a projective completion of C having at least three points at infinity. Then, for any closed embedding $i\colon C \hookrightarrow \mathbb{A}^n_K$ over K, the set $i^{-1}(R^n)$ of integral points on C is finite.*

Proof. The proof follows the classical proof over number fields very closely.

By enlarging K, we may assume that C is geometrically integral. Fix a big polarization $M = (B; \mathscr{M})$ of K such that B is normal and generically smooth. Let $S \subseteq M_K$ be the union of M_K^∞ and the set of all prime divisors Y on B such that some generator of R has a pole along Y. Then S has finite measure, and R is contained in the ring of S-integers of K. Let X_0 be a projective closure of C.

Let $i\colon C \to \mathbb{A}^n_K$ be a closed embedding over K, and let $x_1, \dots, x_n$ be the pull-backs to C of the coordinate functions on $\mathbb{A}^n_K$. Then, for each $v \in M_K$ the function $X_0(\mathbb{C}_v) \to \mathbb{R} \cup \{\infty\}$ given by

$$\lambda_v(\xi) = \max\{0, \log\|x_1(\xi)\|_v, \dots, \log\|x_n(\xi)\|_v\}$$

defines a Weil function λ on X_0 for an effective divisor D_0 such that $\operatorname{Supp} D_0 = X_0 \setminus C$.

Let $\Sigma = i^{-1}(R^n)$, and assume that this is an infinite set. By construction we have $\lambda_v(\xi) = 0$ for all $v \in M_K \setminus S$ and all $\xi \in \Sigma$.

Let $\pi\colon X \to X_0$ be the normalization of X_0, let D be the reduced divisor on X such that $\operatorname{Supp} D = X \setminus \pi^{-1}(C)$, and choose a Weil function λ_D for D on X. Since $\pi^* D_0 - D$ is an effective divisor and $\lambda_v(\xi) = 0$ for all $v \notin S$ and all $\xi \in \Sigma$, Theorem 3.24(e) implies that there is an M_K-constant (c_v) such that $\lambda_{D,v}(\xi) \le c_v$ for all $v \notin S$ and all $\xi \in \pi^{-1}(\Sigma)$. It then follows that

$$m_S(D, \xi) = h_D(\xi) + O(1) \tag{4.11.1}$$

for all $\xi \in \pi^{-1}(\Sigma)$, where $m_S(D, \cdot)$ and h_D are proximity and height functions defined using λ_D.

Let g be the genus of X. The hypotheses on C imply that $\deg D > 2 - 2g$, so (4.11.1) contradicts Corollary 4.10 by basic properties of heights (which still hold over arithmetic function fields). □

We conclude this section with two examples showing that Theorem 0.1 does not extend straightforwardly to arithmetic function fields without requiring $\{\alpha_v : v \in S\}$ to be a finite set.

These two examples use the standard notation $B_r(z_0) = \{z \in \mathbb{C} : |z - z_0| < r\}$.

Example 4.12. Let $K = \mathbb{Q}(t)$ with t an indeterminate, let $B = \mathbb{P}^1_{\mathbb{Z}}$, let $\mathscr{M} = \mathscr{O}(1)$ with Fubini–Study metric, and let $S = M_K^\infty$. Identify $B(\mathbb{C})$ with $\mathbb{C} \cup \{\infty\}$ in the usual way, so that S is identified with $\mathbb{C} \setminus \overline{\mathbb{Q}}$ by associating $\tau\colon K \to \mathbb{C}$ to $\tau(t) \in \mathbb{C} \setminus \overline{\mathbb{Q}}$.

For each $n \in \mathbb{N}$ let S_n be the subset of S corresponding to $B_{\frac{1}{2}}(n) \cap (\mathbb{C} \setminus \overline{\mathbb{Q}})$. Note that these subsets are mutually disjoint (but do not cover S).

Since $\mathbb{Q}(\sqrt{-1})$ is dense in $\mathbb{C}$ (in the classical topology), for each $n \in \mathbb{N}$ and each $v \in S_n$ we may choose $\beta_v \in \mathbb{Q}(\sqrt{-1})$ to be arbitrarily close to $v - n$. This can be done so that the function $v \mapsto \beta_v$ is a measurable function (for example, partition S_n into finitely many measurable subsets and let β_v be constant on each of these subsets). Let $\beta_v = 0$ for all $v \notin S_0 \cup S_1 \cup \cdots$, and let $\alpha_v = t - \beta_v$ for all $v \in S$.

If we choose β_v such that $-\log|\beta_v + n - v| \ge 3h_K(n)/\mu(S_n)$ for all $n \in \mathbb{N}$ and all $v \in S_n$, then we will have

$$\int_S -\log^- \|n - \alpha_v\|_v \, d\mu(v) \ge \int_{S_n} -\log^- |n - v + \beta_v| \, d\mu(v) \ge 3h_K(n)$$

for all $n \in \mathbb{N}$. Thus, taking $\epsilon = 1$ and $c = 0$, we have constructed an infinite subset $\mathbb{N} \subseteq K$ and a system of choices of $\alpha_v \in \overline{K}$ for all $v \in S$ such that

$$\int_S -\log^- \|\xi - \alpha_v\|_v \, d\mu(v) \ge (2+\epsilon)h_K(\xi) + c \tag{4.12.1}$$

for all $\xi \in \mathbb{N}$.

In this example, the elements $\alpha_v \in \overline{K}$ all have finite degrees over K, and in fact they all lie in the same arithmetic function field $\mathbb{Q}(\sqrt{-1}, t)$. However, their heights are unbounded.

This next example is very similar, except that the heights are bounded but the degrees are not. (Bounding both the degrees and the heights amounts to requiring that $\{\alpha_v : v \in S\}$ be a finite set.)

Example 4.13. Let $F = \mathbb{Q}(\sqrt{-1})$, let $K = F(t)$, let $B = \mathbb{P}^1_{\mathbb{Z}[\sqrt{-1}]}$, and let $M = \mathscr{O}(1)$ with Fubini–Study metric. Fix an embedding $i : F \to \mathbb{C}$, and let $S \subseteq M_K^\infty$ be the subset of maps $\tau : K \hookrightarrow \mathbb{C}$ that satisfy $\tau|_F = i$. Again identify S with $\mathbb{C} \setminus \overline{\mathbb{Q}}$ as in Example 4.12.

This example will use the fact that the set $\{\zeta + \zeta' : \zeta \text{ and } \zeta' \text{ are roots of unity}\}$ is dense in the closed ball $|z| \le 2$.

Choose $\xi_n \in F$ and $r_n > 0$ for all $n \in \mathbb{N}$ such that $S_n := B_{r_n}(\xi_n) \cap (\mathbb{C} \setminus \overline{\mathbb{Q}})$ are mutually disjoint subsets of $B_2(0)$. Then, as noted above, for each n and each $v \in S_n$ one can choose roots of unity ζ_v and ζ'_v whose sum is arbitrarily close to $v - \xi_n$.

Then, proceeding as before, we construct a collection of choices $\alpha_v \in \overline{K}$ for all $v \in S$ such that (4.12.1) with $\epsilon = 1$ and $c = 0$ holds for all ξ in the infinite subset $\Xi := \{\xi_0, \xi_1, \ldots\}$ of K. In addition, $h_K(\alpha_v) \le h_K(t) + \mu_\infty(B_2(0)) \log 4$ for all $v \in S$.

5. Reductions

In this section we begin the main line of the proof of Roth's theorem over arithmetic function fields. Specifically, Theorem 4.5 will be proved in the remaining sections of the paper (and the other variations will then follow, by Proposition 4.8).

The purpose of this section is to show that it will suffice to prove Theorem 4.5 under the following additional hypotheses:

5.1. The set S contains all of the archimedean places.

5.2. B is generically smooth.

5.3. $\mathscr{M}$ is ample.

5.4. The metric on $\mathscr{M}$ is positive.

We start by noting that the integrand of (4.5.1) is nonnegative, so enlarging the set S will only strengthen the theorem. In particular, we may assume that 5.1 holds.

Next, consider the condition 5.2. Recall from 4.1 that $M = (B; \mathscr{M})$ is a big polarization of K. Let $\pi : B' \to B$ be a generic resolution of singularities of B, and let $\mathscr{M}' = \pi^* \mathscr{M}$. Then $M' := (B'; \mathscr{M}')$ is also a big polarization of K.

The map π induces a bijection $\pi^{\text{gen}} \colon B'(\mathbb{C})^{\text{gen}} \to B(\mathbb{C})^{\text{gen}}$ which preserves measures and absolute values.

As for nonarchimedean places, let $Y' \in (B')^{(1)}$, and let $Y = \pi(Y')$. First consider the case in which $\operatorname{codim} Y = 1$. Then $Y \in B^{(1)}$, and $h_M(Y) = h_{M'}(Y')$ by (1.9). Also $\operatorname{ord}_{Y'}(\xi) = \operatorname{ord}_Y(\xi)$ for all $\xi \in K^*$, so we have $\|\xi\|_{Y'} = \|\xi\|_Y$ for all $\xi \in K$.

Next consider Y' for which $\operatorname{codim} Y > 1$. Then $\pi_*(Y, 0) = 0$ in $\mathrm{Z}^1_{\mathrm{D}}(B)$, so $h_{M'}(Y') = 0$ by (1.8). Therefore $\|\xi\|_{Y'} = 1$ for all $\xi \in K^*$.

Therefore, it is clear from (3.6) that $h_M(\xi)$ remains the same when one changes the polarization from M to M'.

Next let S' be the subset of M'_K defined by

$$S' = B'(\mathbb{C})^{\text{gen}} \cup \{Y' \in (B')^{(1)} : \pi(Y') \in S \cap B^{(1)}\}.$$

Since $S \supseteq B(\mathbb{C})^{\text{gen}}$ by 5.1, the integral in (4.5.1) is unchanged when S is replaced by S'. Therefore, for each $\xi \in K$, (4.5.1) is true for the polarization M if and only if it is true for M', and therefore it suffices to prove Theorem 4.5 under the additional conditions 5.1 and 5.2.

This leaves 5.3 and 5.4. We begin with a result from Arakelov theory.

In the remainder of this section, it will be convenient to work with slightly different notation. For an integral scheme X, projective over $\operatorname{Spec} \mathbb{Z}$, let $\widehat{\operatorname{Pic}}(X)$ denote the group of smoothly metrized line sheaves on X, whose group operation is tensor product. A *smoothly metrized* $\mathbb{Q}$*-line sheaf* on X is an element of $\widehat{\operatorname{Pic}}(X) \otimes \mathbb{Q}$. The previous definitions of "nef," "big," and "ample" extend to this group. For simplicity, elements of $\widehat{\operatorname{Pic}}(X) \otimes \mathbb{Q}$ will be written additively.

Since the intersection number $c_1(\mathscr{L}_1) \cdots c_1(\mathscr{L}_n)$ on an arithmetic variety X is multilinear, its definition extends to allow the $\mathscr{L}_i$ to be smoothly metrized $\mathbb{Q}$-line sheaves, and correspondingly we allow smoothly metrized $\mathbb{Q}$-line sheaves to be used in polarizations.

Lemma 5.5. *Let B be a generically smooth arithmetic variety, and let $\mathscr{M}$ and $\mathscr{A}$ be smoothly metrized line sheaves on B. Assume that $\mathscr{M}$ is big and nef, that $\mathscr{A}$ is ample, and that the metric on $\mathscr{A}$ is positive. Then*:

(a) *For all rational $\delta > 0$, $\mathscr{M} + \delta\mathscr{A}$ is ample, and its metric is positive.*

(b) *Let $\delta \in \mathbb{Q}_{>0}$. Let $K = \boldsymbol{\kappa}(B)$, and let h_K and h'_K denote the naïve heights computed using the polarizations $(B; \mathscr{M})$ and $(B; \mathscr{M} + \delta\mathscr{A})$, respectively. Then $h'_K(\xi) \geq h_K(\xi)$ for all $\xi \in K$.*

(c) *For any given $\epsilon'' > 0$ there is a rational $\delta > 0$ such that the inequality*

$$c_1((\mathscr{M} + \delta\mathscr{A})Y)^{\cdot d} \leq (1 + \epsilon'')c_1(\mathscr{M}|_Y)^{\cdot d} \tag{5.5.1}$$

holds for all but finitely many $Y \in B^{(1)}$.

Proof. First, we claim that the inequality

$$c_1((\mathscr{M}+\delta\mathscr{A})|_Y)^{\cdot\dim Y} \ge c_1(\mathscr{M}|_Y)^{\cdot\dim Y} \tag{5.5.2}$$

holds for all rational $\delta > 0$ and all integral closed subschemes Y of X. Indeed,

$$c_1((\mathscr{M}+\delta\mathscr{A})|_Y)^{\cdot\dim Y} - c_1(\mathscr{M}|_Y)^{\cdot\dim Y} = \sum_{i=1}^{\dim Y}\binom{\dim Y}{i}\delta^i c_1(\mathscr{M}|_Y)^{\cdot(\dim Y-i)}\cdot c_1(\mathscr{A}|_Y)^{\cdot i},$$

and each term on the right-hand side is nonnegative.

By a similar argument,

$$c_1((\mathscr{M}+\delta\mathscr{A})|_Y)^{\cdot\dim Y} \ge \delta^{\dim Y} c_1(\mathscr{A}|_Y)^{\cdot\dim Y}. \tag{5.5.3}$$

Now consider (a). The metric on $\mathscr{M}+\delta\mathscr{A}$ is positive because the metrics on $\mathscr{A}$ and $\mathscr{M}$ are positive and semipositive, respectively. Also $\mathscr{M}+\delta\mathscr{A}$ is vertically nef because both $\mathscr{M}$ and $\mathscr{A}$ are.

Since the metric on $\mathscr{M}$ is semipositive, $\mathscr{M}_{\mathbb{Q}}$ is nef, and therefore $(\mathscr{M}+\delta\mathscr{A})_{\mathbb{Q}}$ is ample (by either Kleiman's or Seshadri's criterion for ampleness).

Finally, $\mathscr{M}+\delta\mathscr{A}$ is horizontally positive by (5.5.3) and horizontal positivity of $\mathscr{A}$. Thus $\mathscr{M}+\delta\mathscr{A}$ is ample.

Next consider (b). Let $\xi \in K$. By (3.6),

$$h'_K(\xi) - h_K(\xi) = \int_{B(\mathbb{C})^{\text{gen}}} \log^+|\xi(b)|\,d(\mu'_\infty(b)-\mu_\infty(b)) + \sum_{Y\in B^{(1)}} \max\{0,-\operatorname{ord}_Y(\xi)\}(h_{M'}(Y)-h_M(Y)), \tag{5.5.4}$$

where μ_∞ and μ'_∞ are the measures on $B(\mathbb{C})^{\text{gen}}$ defined using M and M', respectively. The signed measure $\mu'_\infty - \mu_\infty$ is associated to the (d,d)-form

$$c_1(\|\cdot\|_{\mathscr{M}+\delta\mathscr{A}})^{\wedge d} - c_1(\|\cdot\|_{\mathscr{M}})^{\wedge d} = \sum_{i=1}^{d}\binom{d}{i}\delta^i c_1(\|\cdot\|_{\mathscr{M}})^{\wedge(d-i)} \wedge c_1(\|\cdot\|_{\mathscr{A}})^{\wedge i},$$

and this is nonnegative because each term on the right is nonnegative. Also, by (5.5.2), $h_{M'}(Y) \ge h_M(Y)$ for all $Y \in B^{(1)}$. Therefore, the right-hand side of (5.5.4) is nonnegative, and this gives (b).

Finally, consider (c).

By [Moriwaki 2014, Proposition 5.43], there is a rational $\eta > 0$ such that some positive integer multiple of $\mathscr{M}-\eta\mathscr{A}$ has a nonzero strictly small global section.

Let s be such a global section. Let $Y \in B^{(1)}$, and assume that Y does not occur in the support of $\operatorname{div}(s)_{\text{fin}}$. This excludes only finitely many Y.

Since $s|_Y$ is nonzero and both $\mathscr{M}$ and $\mathscr{A}$ are nef, Proposition 1.12(b) gives

$$c_1(\mathscr{M}|_Y)^{\cdot(d-1-j)}\cdot c_1(\mathscr{A}|_Y)^{\cdot j}\cdot c_1((\mathscr{M}-\eta\mathscr{A})|_Y) \ge 0 \tag{5.5.5}$$

for all $j = 0, \ldots, d-1$.

Let $\epsilon'' > 0$ be given. Choose a rational $\delta > 0$ such that

$$(1+\epsilon'')\eta^d \geq (\eta+\delta)^d.$$

Since

$$(1+\epsilon'')\eta^d - (\eta+\delta)^d = \eta^d\epsilon'' - \sum_{i=1}^{d}\binom{d}{i}\eta^{d-i}\delta^i,$$

we have

$$\eta^j\epsilon'' - \sum_{i=1}^{j}\binom{d}{i}\eta^{j-i}\delta^i \geq 0 \tag{5.5.6}$$

for all $j = 0, \ldots, d$.

For $j = 0, \ldots, d$ let

$$C_j = \left(\eta^j\epsilon'' - \sum_{i=1}^{j}\binom{d}{i}\eta^{j-i}\delta^i\right)c_1(\mathscr{M}|_Y)^{\cdot(d-j)}\cdot c_1(\mathscr{A}|_Y)^{\cdot j} - \sum_{i=j+1}^{d}\binom{d}{i}\delta^i c_1(\mathscr{M}|_Y)^{\cdot(d-i)}\cdot c_1\big(\mathscr{A}|_Y\big)^{\cdot i}.$$

We claim that $C_j \geq 0$ for all j. This will be proved by descending induction on j. When $j = d$, we have

$$C_d = \left(\eta^d\epsilon'' - \sum_{i=1}^{d}\binom{d}{i}\eta^{d-i}\delta^i\right)c_1(\mathscr{A}|_Y)^{\cdot d},$$

and this is nonnegative by (5.5.6) and Proposition 1.12(a). For $j = 0, \ldots, d-1$, we have

$$C_{j+1} = \eta(\eta^j\epsilon'' - \sum_{i=1}^{j}\binom{d}{i}\eta^{j-i}\delta^i)c_1(\mathscr{M}|_Y)^{\cdot(d-j-1)}\cdot c_1(\mathscr{A}|_Y)^{\cdot(j+1)} - \sum_{i=j+1}^{d}\binom{d}{i}\delta^i c_1(\mathscr{M}|_Y)^{\cdot(d-i)}\cdot c_1(\mathscr{A}|_Y)^{\cdot i},$$

and therefore

$$C_j - C_{j+1} = \left(\eta^j\epsilon'' - \sum_{i=1}^{j}\binom{d}{i}\eta^{j-i}\delta^i\right)\cdot\big(c_1(\mathscr{M}|_Y)^{\cdot(d-j)}\cdot c_1(\mathscr{A}|_Y)^{\cdot j} - \eta c_1(\mathscr{M}|_Y)^{\cdot(d-j-1)}\cdot c_1(\mathscr{A}|_Y)^{\cdot(j+1)}\big).$$

By (5.5.6) and (5.5.5), the right-hand side is nonnegative; hence $C_j \geq C_{j+1}$.

We then have $C_0 \geq 0$. Since

$$C_0 = \epsilon'' c_1(\mathscr{M}|_Y)^{\cdot d} - \sum_{i=1}^{d}\binom{d}{i}\delta^i c_1(\mathscr{M}|_Y)^{\cdot(d-i)}\cdot c_1(\mathscr{A}|_Y)^{\cdot i} = (1+\epsilon'')c_1(\mathscr{M}|_Y)^{\cdot d} - c_1((\mathscr{M}+\delta\mathscr{A})|_Y)^{\cdot d},$$

we have (5.5.1). □

This sets the stage for the main result of this section.

Remark 5.6. In the proof of the following proposition, it will be convenient to consider polarizations $(B; \mathscr{M}')$ in which $\mathscr{M}'$ is a smoothly metrized $\mathbb{Q}$-line sheaf. This can be justified as follows.

Let $(B; \mathscr{M})$ be a polarization of K, and let n be a positive integer. Then $(B; n\mathscr{M})$ is also a polarization, with the same set M_K of places. The archimedean absolute values of this new polarization are the same as those of the original polarization, but the measure μ_∞ is multiplied by n^d. For nonarchimedean places, the counting measure is of course unchanged, but the absolute values for $(B; n\mathscr{M})$ are the n^d powers of the absolute values for $(B; \mathscr{M})$. Therefore the naïve height is multiplied by n^d by this change. Similarly, let D be a Cartier divisor on a variety V over K, and let λ be a Weil function for D using the polarization $(B; \mathscr{M})$. Define a function λ' by letting $\lambda'_v = \lambda_v$ for all archimedean v and $\lambda' = n^d \lambda_v$ for all nonarchimedean v. Then λ' is a Weil function for D relative to $(B; n\mathscr{M})$. It then follows that the proximity and counting functions obtained using λ' and $(B; n\mathscr{M})$ are equal to n^d times those obtained using λ and $(B; \mathscr{M})$.

Therefore, one obtains well-defined notions of absolute value, naïve height, Weil functions, proximity functions, and counting functions for polarizations with smoothly metrized $\mathbb{Q}$-line sheaves. And, if Theorem 4.5 holds for polarizations as defined earlier, then it is also true for polarizations using smoothly metrized $\mathbb{Q}$-line sheaves.

Proposition 5.7. *It suffices to prove Theorem 4.5 under the additional hypotheses 5.1–5.4.*

Proof. As noted earlier, we may already assume that 5.1 and 5.2 hold, so it remains to show that if Theorem 4.5 holds under 5.1–5.4 then it holds when only 5.1 and 5.2 are assumed to be true. By Remark 4.9, we may work with Theorem 4.4 instead of 4.5.

So let K, M_K, and S be as in 4.1, where S contains all archimedean places, and the polarization $M = (B; \mathscr{M})$ satisfies 5.2; i.e., B is generically smooth. Also let $\alpha_1, \dots, \alpha_q$, ϵ, and c be as in the statement of Theorem 4.4.

Pick $\epsilon' > 0$ and $\epsilon'' > 0$ such that

$$\frac{q-2-\epsilon'}{1+\epsilon''} = q-2-\epsilon. \tag{5.7.1}$$

Choose an ample smoothly metrized line sheaf $\mathscr{A}$ on B with positive metric, and let $\delta > 0$ be as in Lemma 5.5(c). We may assume that S contains all of the (finitely many) $Y \in B^{(1)}$ for which (5.5.1) fails to hold.

Let $D = [\alpha_1] + \cdots + [\alpha_q]$, and let λ_D be the Weil function for D defined by

$$\lambda_{D,v} = -\sum_{i=1}^{q} \log^- \|\xi - \alpha_i\|_v. \tag{5.7.2}$$

Also let $m_S(D, \xi)$ and $N_S(D, \xi)$ be as in Definition 3.31. By Proposition 3.33,

$$m_S(D, \xi) + N_S(D, \xi) = q\, h_K(\xi) + O(1) \tag{5.7.3}$$

for all $\xi \in K \setminus \{\alpha_1, \dots, \alpha_q\}$.

Let $\mathscr{M}' = \mathscr{M} + \delta\mathscr{A}$, and let $M' = (B; \mathscr{M}')$. Note that M_K depends only on B, so it is the same for both polarizations M and M'. Define $h'_K(\xi)$, λ'_D, $m'_S(D, \xi)$, and $N'_S(D, \xi)$ similarly to $h_K(\xi)$, λ_D, etc., but using M' instead of M. Again, we have

$$m'_S(D, \xi) + N'_S(D, \xi) = q\, h'_K(\xi) + O(1) \tag{5.7.4}$$

for all $\xi \in K \setminus \{\alpha_1, \dots, \alpha_q\}$.

By Lemma 5.5(a), $\mathscr{M}'$ is ample with positive metric. Therefore, we can apply Theorem 4.4 to get that, for all $c' \in \mathbb{R}$, the inequality

$$m'_S(D,\xi) \le (2+\epsilon')h'_K(\xi) + c'$$

holds for all but finitely many $\xi \in K$ (where the excluded set depends on c' as well as all other data here). By (5.7.4) there is a constant a', independent of c', such that

$$N'_S(D,\xi) \ge (q-2-\epsilon')h'_K(\xi) - c' - a'.$$

By Lemma 5.5(b) and (5.5.1), we have $h'_K(\xi) \ge h_K(\xi)$ and $N'_S(D,\xi) \le (1+\epsilon'')N_S(D,\xi)$ for all $\xi \in K \setminus \{\alpha_1,\dots,\alpha_q\}$. Therefore

$$(1+\epsilon'')N_S(D,\xi) \ge (q-2-\epsilon')h_K(\xi) - c' - a'$$

for all but finitely many $\xi \in K$. By (5.7.1) and (5.7.3), there is a constant a, independent of c', such that

$$m_S(D,\xi) \le (2+\epsilon)h_K(\xi) + \frac{c'+a'}{1+\epsilon''} + a.$$

We can then take c' small enough so that $(c'+a')/(1+\epsilon'') + a \le c$ to get (4.4.1). □

6. Reduction to simultaneous approximation: The main analytic part

The proof of Theorem 4.5 follows the classical proof over number fields very closely. Most parts carry over directly without difficulty. The main exception to this is the part of the proof that is often called "reduction to simultaneous approximation". This is briefly described in the Introduction; see also [Lang 1983, Chapter 7, Section 2], [Hindry and Silverman 2000, Theorem D.2.2], or [Bombieri and Gubler 2006, 6.4.2–6.4.4].

In more detail, reduction to simultaneous approximation is as follows. In the special case of number fields, (4.5.1) reduces to the inequality

$$\sum_{v\in S} \max_{1\le j\le q} -\log^-\|\xi-\alpha_j\|_v \le (2+\epsilon)h_K(\xi) + c,$$

where S is a finite set. Reduction to simultaneous approximation consists of showing that, to prove Roth's theorem, it suffices to prove the following statement. For all functions $j\colon S \to \{1,\dots,q\}$ and all $(c_v)_{v\in S} \in \mathbb{R}^{\#S}$ such that $\sum c_v > 2$, only finitely many $\xi \in K$ simultaneously satisfy

$$-\log^-\|\xi-\alpha_{j(v)}\|_v > c_v h_K(\xi)$$

for all $v \in S$.

In the number field case S is finite, so this is proved by a simple compactness argument combined with the pigeonhole principle. In the case of arithmetic function fields, though, $S \cap M_K^\infty$ is a subset of a complex manifold and $v \mapsto -\log^-\|\xi-\alpha_j\|_v$ is a smooth function (with singularities on the manifold outside of M_K^∞). This becomes a question in analysis, reminiscent of the Arzelà–Ascoli theorem. In fact, the proof presented here is motivated by the proof of the Arzelà–Ascoli theorem. The singularities can be handled by removing a subset T of bounded measure from M_K^∞. It is possible to do this, for basically the same reason as in [Wirsing 1971]. Since $M_K^\infty \setminus T$ may now be locally disconnected, though, it is necessary to work with differences instead of derivatives.

Another challenge in reducing to simultaneous approximation is the fact that the analytic estimates in the proof need to be uniform in the rational points. Simple compactness arguments will not work here. For example, in the $d = 1$ case the degree of the rational function can be arbitrarily large. Instead, we can use the fact that $-\log\|\xi - \alpha_j\|_v$ is a Green function for the principal divisor $(\xi - \alpha_j)$, and use properties of Green forms and functions from Arakelov theory to write this function as an integral whose integrand can be treated using compactness arguments (see Proposition 6.3).

The proof of reduction to simultaneous approximation for arithmetic function fields takes up the next three sections of this paper. They form the core of this paper.

This section carries out the main analytic arguments leading up to Proposition 6.16, which is motivated by a part of the proof of the Arzelà–Ascoli theorem. Section 7 gives an upper bound on what is lost by removing the set T; this is Proposition 7.3. Section 8 then carries these two results over to the arithmetical setting, and proves the main result on reduction to simultaneous approximation (Proposition 8.12). This is the part that uses the pigeonhole argument.

Ultimately the proof of Proposition 6.16 relies on the following elementary lemma on integration (which is used in proving Lemma 6.13).

Lemma 6.1. *Let X be a space with measure μ, let $g\colon X \to [0, \infty]$ be a measurable function with finite integral, and let $c > 0$. Then*

$$\mu(\{x \in X : g(x) \ge c\}) \le \tfrac{1}{c} \int_X g\, d\mu.$$

Proof. Let $\chi\colon X \to [0, c]$ be the function defined by $\chi(x) = c$ if $g(x) \ge c$ and $\chi(x) = 0$ otherwise. Then

$$\int_X g\, d\mu - c\mu(\{x \in X : g(x) \ge c\}) = \int_X (g - \chi)\, d\mu \ge 0$$

because the integrand is nonnegative. □

Wirsing's proof also uses this lemma (via its reliance on Chebyshev's inequality).

Results in this section and the next will be phrased in terms of a smooth complex projective variety X. The topology on X will be the classical topology. In Section 8 we will apply these results as X varies over all connected components of $B(\mathbb{C})$, where B is the arithmetic variety in some polarization of K. Note that K is a subfield of $\boldsymbol{\kappa}(X)$ (in fact, $\boldsymbol{\kappa}(X)$ is the compositum of K and $\mathbb{C}$ over the algebraic closure F of $\mathbb{Q}$ in K, for some choice of embedding of F into $\mathbb{C}$).

Definition 6.2. Let X be a smooth complex projective variety and let $Y \subseteq X$ be an irreducible closed subvariety of X of codimension $p > 0$. Then a *Green form* for Y is a smooth $(p-1, p-1)$-form on $X \setminus Y$ whose associated current on X is a Green current for Y. A *Green form of log type* for Y is a Green form for Y that is of logarithmic type along Y [Soulé 1992, Definition II.3].

Proposition 6.3. *Let X be a smooth complex projective variety of dimension $d \geq 1$. Let Δ be the diagonal in $X \times X$, let $\pi\colon W \to X \times X$ be the blowing-up of $X \times X$ along Δ, let E be the exceptional divisor, choose a smooth metric on the line sheaf $\mathscr{O}(E)$, and let s be the canonical section of this line sheaf. Then there exist smooth $(d-1, d-1)$-forms α and β on W for which the following statements are true*:

(a) *There is a Green form g_Δ of log type for Δ on $X \times X$ such that*

$$\pi^* g_\Delta = (-\log\|s\|^2)\alpha + \beta \quad \text{on } W \setminus E. \tag{6.3.1}$$

(b) *For each prime divisor D on X, define g_D as follows. Let $j\colon \tilde{D} \to X$ be a proper map with image D such that $\tilde{D} \to D$ is a desingularization of D, let $q\colon \tilde{D} \times X \to X$ be the projection to the second factor, and let*

$$g_D = q_*(j \times \mathrm{Id}_X)^* g_\Delta. \tag{6.3.2}$$

Then g_D is a Green form of log type for D on X.

(c) *For each $\xi \in \boldsymbol{\kappa}(X)^*$, write the principal divisor (ξ) as a (finite) sum $(\xi) = \sum_D n_D D$, where each D is a prime divisor and $n_D \in \mathbb{Z}$ for all D. Then there is a constant c such that*

$$-\log|\xi|^2 = \sum_D n_D g_D + c. \tag{6.3.3}$$

Proof. Part (a) is proved in Step 2 of the proof of [Soulé 1992, Theorem II.3], where f is taken to be the identity map on X.

For part (b), note that $(j \times \mathrm{Id}_X)^{-1}(\Delta)$ is the graph Γ_j of j. Since

$$\operatorname{codim}_{\tilde{D} \times X} \Gamma_j = \operatorname{codim}_{X \times X} \Delta,$$

it follows from [Soulé 1992, Section II.3.2] that $(j \times \mathrm{Id}_X)^* g_\Delta$ is a Green form of log type for Γ_j.

Since the push-forward $q_*\Gamma_j$ equals D (as cycles on X), it follows from [Soulé 1992, II, Lemma 2(ii) and proof of III, Theorem 3(ii)] that $q_*\,(j \times \mathrm{Id}_X)^* g_\Delta$ is a Green form of log type for D on X. This gives part (b).

For part (c), we note that both $-\log|\xi|^2$ and $\sum n_D g_D$ are Green forms for the same divisor (ξ). Therefore, by [Gillet and Soulé 1990, Lemma 1.2.4], there is a smooth function $f\colon X \to \mathbb{R}$ such that

$$-\log|\xi|^2 = \sum n_D g_D + \log f \tag{6.3.4}$$

everywhere outside of the support of the divisor (ξ).

Since g_Δ is a Green form for Δ on $X \times X$, the (d, d)-form $dd^c g_\Delta$ extends to a smooth form ω_Δ on $X \times X$. Similarly, if D is a prime divisor then $dd^c g_D$ extends to a smooth form ω_D on X. By functoriality, $dd^c((j \times \mathrm{Id}_X)^* g_\Delta)$ extends to the smooth form $(j \times \mathrm{Id}_X)^*\omega_\Delta$ on $\tilde{D} \times X$, and by [Soulé 1992, proof of III, Theorem 3(ii)] we have

$$q_*\,(j \times \mathrm{Id}_X)^*\omega_\Delta = \omega_D. \tag{6.3.5}$$

Let $\mathscr{H}^{i,j}(M)$ denote the set of harmonic (i, j)-forms on M for some fixed choice of Kähler (or Riemannian) metric on a complex manifold M [Griffiths and Harris 1978, page 82]. Fix such a metric on X and use the induced metric on $X \times X$. By the construction in Step 2 of the proof of [Soulé 1992, Theorem II.3], we may choose g_Δ such that ω_Δ is any given representative of Δ in $H^{d,d}_{\bar\partial}(X \times X)$. By the Hodge decomposition [Griffiths and Harris 1978, page 116], each cohomology class is represented by a unique harmonic form. Therefore we may assume that ω_Δ is harmonic.

By the Künneth formula [Griffiths and Harris 1978, page 104],

$$\mathscr{H}^{d,d}(X \times X) = \bigoplus_{\substack{i+j=d \\ i'+j'=d}} \mathscr{H}^{i,j}(X) \otimes \mathscr{H}^{i',j'}(X).$$

Applying this decomposition to ω_Δ, the only component that affects the value of $q_*(j \times \mathrm{Id}_X)^*\omega_\Delta$ is the one with $j = j' = 1$. Therefore there are forms

$$u_1, \dots, u_n \in \mathscr{H}^{d-1,d-1}(X) \quad \text{and} \quad v_1, \dots, v_n \in \mathscr{H}^{1,1}(X)$$

such that if $\tilde p, \tilde q \colon X \times X \to X$ are the first and second projections and if $p \colon \tilde D \times X \to \tilde D$ is the first projection, then

$$\begin{aligned} q_*(j \times \mathrm{Id}_X)^*\omega_\Delta &= \sum_{i=1}^n q_*(j \times \mathrm{Id}_X)^*(\tilde p^* u_i \otimes \tilde q^* v_i) \\ &= \sum_{i=1}^n q_*(p^* j^* u_i \otimes q^* v_i) \\ &= \sum_{i=1}^n \biggl(\int_{\tilde D} j^* u_i\biggr) q^* v_i. \end{aligned}$$

In particular, by (6.3.5), ω_D is harmonic.

Therefore, $\sum n_D \omega_D$ is also harmonic. Since it represents the (trivial) cohomology class of the principal divisor (ξ), it must be zero. By (6.3.4), we then have

$$dd^c \log f = -dd^c \log|\xi|^2 = 0,$$

and therefore f is constant. □

Remark 6.4. In part (b), we may assume that j maps a Zariski-open subset U of $\tilde D$ isomorphically to the smooth locus D_{reg} of D. Since $\tilde D \setminus U$ has measure zero and g_Δ is a form (as opposed to a current), we can compute g_D by integrating over D_{reg}:

$$g_D(x) = \int_{D_{\mathrm{reg}} \times \{x\}} g_\Delta \quad \text{for all } x \in X \setminus D. \tag{6.4.1}$$

The following construction will often be used to obtain analytic estimates.

Lemma 6.5. *Let M_1 and M_2 be complex manifolds of dimension $d \geq 1$, and let ψ be a positive smooth $(d-1, d-1)$-form on M_1. Let $\operatorname{Gr}^1 TM_1$ be the Grassmannnian of hyperplanes in fibers of the tangent bundle TM_1, and let $\tau_1\colon \operatorname{Gr}^1 TM_1 \to M_1$ be the structural morphism. Let $G = (\operatorname{Gr}^1 TM_1) \times M_2$ and $\tau = \tau_1 \times \mathrm{Id}_{M_2}\colon G \to M_1 \times M_2$. This can be regarded as the Grassmannian of hyperplanes in fibers of the relative tangent bundle of $M_1 \times M_2$ over M_2, taken relative to the projection $q\colon M_1 \times M_2 \to M_2$ to the second factor.*

Then, for each open subset U of $M_1 \times M_2$ and each smooth $(d-1, d-1)$-form α on U, there is a unique smooth function $\chi_\alpha\colon \tau^{-1}(U) \to \mathbb{C}$, depending only on M_1, M_2, ψ, U, and α, such that the following is true.

Let N be a locally closed submanifold of M_1 of dimension $d-1$. At each $w \in N$, the tangent space $T_w N$ is a hyperplane in $T_w M_1$, and this gives smooth sections $\sigma_{N,1}\colon N \to \operatorname{Gr}^1 TM_1$ and $\sigma_N := \sigma_{N,1} \times \mathrm{Id}_{M_2}\colon N \times M_2 \to G$ of $\tau_1^{-1}(N) \to N$ and $\tau^{-1}(N \times M_2) \to N \times M_2$, respectively. Then we have

$$\alpha|_{(N\times M_2)\cap U_x} = ((\chi_\alpha \circ \sigma_N)\cdot(p^*\psi))|_{(N\times M_2)\cap U_x} \quad \textit{for all } x \in M_2, \tag{6.5.1}$$

where $U_x = (M_1 \times \{x\}) \cap U$ and $p\colon M_1 \times M_2 \to M_1$ is the projection to the first factor.

Proof. Let N be as above. For dimension reasons, there is a smooth function

$$\rho_{\alpha,N}\colon (N \times M_2) \cap U \to \mathbb{C}$$

such that

$$\alpha|_{(N\times M_2)\cap U_x} = \rho_{\alpha,N}\cdot(p^*\psi)|_{(N\times M_2)\cap U_x} \quad \text{for all } x \in M_2. \tag{6.5.2}$$

For each $(w, x) \in U$ and each N passing through w, the value of this function at (w, x) depends only on $T_w N$; in other words, if N and N' both pass through a point $w \in M_1$ and are tangent at w, then $\rho_{\alpha,N}(w, x) = \rho_{\alpha,N'}(w, x)$ for all $x \in M_2$ such that $(w, x) \in U$.

We claim that there is a function $\chi_\alpha\colon \tau^{-1}(U) \to \mathbb{C}$ such that

$$\rho_{\alpha,N}|_{(N\times M_2)\cap U_x} = \chi_\alpha \circ (\sigma_N|_{(N\times M_2)\cap U_x}) \quad \text{for all } x \in M_2. \tag{6.5.3}$$

Indeed, we first note that the lemma is local on M_1, so we may assume that M_1 is an open subset of $\mathbb{C}^d$. Then $\operatorname{Gr}^1 TM_1$ can be canonically identified with the set of pairs (w, H), where $w \in M_1$ and H is a hyperplane in $\mathbb{C}^d$ passing through w.

For all $(w, H) \in \operatorname{Gr}^1 TM_1$ and all $x \in M_2$ such that $(w, x) \in U$, let

$$\chi_\alpha(w, H, x) = \left(\frac{\alpha|_{(H\times\{x\})\cap U}}{p^*\psi|_{(H\times\{x\})\cap U}}\right)(w, x),$$

where the quotient refers to (6.5.2). Let $(w, x) \in (N \times M_2) \cap U$ and let H be the hyperplane in $\mathbb{C}^d$ tangent to N at w. Then $\sigma_{N,1}(w) = (w, H)$; combining this with (6.5.2) gives

$$\begin{aligned}(\chi_\alpha \circ \sigma_N)(w, x) &= \chi_\alpha(w, H, x) \\ &= \left(\frac{\alpha|_{(H\times\{x\})\cap U}}{p^*\psi|_{(H\times\{x\})\cap U}}\right)(w, x) \\ &= \left(\frac{\alpha|_{(N\times\{x\})\cap U}}{p^*\psi|_{(N\times\{x\})\cap U}}\right)(w, x) \\ &= \rho_{\alpha,N}(w, x).\end{aligned}$$

This gives (6.5.3).

Then (6.5.1) follows by combining (6.5.2) and (6.5.3). □

Corollary 6.6. *Let X, Δ, and g_Δ be as in Proposition 6.3, and let $d = \dim X$. Then, for each positive smooth $(d-1, d-1)$-form ψ on X, there is a χ_{g_Δ} such that*

$$g_D(x) = \int_{w\in D_{\mathrm{reg}}} \chi_{g_\Delta}(\sigma_{D_{\mathrm{reg}},1}(w), x)\cdot \psi(w) \tag{6.6.1}$$

for all D and g_D as in Proposition 6.3(b) and all $x \in X \setminus D$.

Proof. This follows from (6.4.1), by applying Lemma 6.5 with $M_1 = M_2 = X$, $U = (X \times X)\setminus\Delta$, $\alpha = g_\Delta$, and ψ as above. □

The first application of this construction will be to give bounds on the behavior of α and β in Proposition 6.3 near Δ.

Lemma 6.7. *Let V be an open subset of $\mathbb{C}^d$ with $d \ge 1$, and let ψ be a positive smooth $(d-1, d-1)$-form on V. Let $\pi\colon W_V \to V \times V$ be the (analytic) blowing-up of $V \times V$ along the diagonal Δ, and let α be a smooth $(d-1, d-1)$-form on $U := (V \times V)\setminus\Delta$ that extends to a smooth form on W_V. Let $\tau\colon (\mathrm{Gr}^1 TV)\times V \to V \times V$ and $\chi_\alpha\colon \tau^{-1}(U) \to \mathbb{C}$ be as in Lemma 6.5.*

We have $TV \cong V \times \mathbb{C}^d$ and therefore $\mathrm{Gr}^1 TV \cong V \times (\mathbb{P}^{d-1})^*$, canonically (where $(\mathbb{P}^{d-1})^*$ is taken to be a point if $d = 1$). Thus we let pairs $(\boldsymbol{w}, H) \in V \times (\mathbb{P}^{d-1})^*$ denote points in $\mathrm{Gr}^1 TV$.

Let L_1 and L_2 be compact subsets of V. Then, for all $(\boldsymbol{w}, \boldsymbol{z}) \in (L_1 \times L_2) \cap U$ and all $H \in (\mathbb{P}^{d-1})^*$, we have

$$|\chi_\alpha(\boldsymbol{w}, H, \boldsymbol{z})| \le O\left(\frac{1}{|\boldsymbol{w}-\boldsymbol{z}|^{2d-2}}\right), \tag{6.7.1}$$

$$\left|\frac{\partial\chi_\alpha(\boldsymbol{w}, H, \boldsymbol{z})}{\partial z_i}\right| \le O\left(\frac{1}{|\boldsymbol{w}-\boldsymbol{z}|^{2d-1}}\right), \quad i = 1, \dots, d, \tag{6.7.2}$$

and

$$\left|\frac{\partial\chi_\alpha(\boldsymbol{w}, H, \boldsymbol{z})}{\partial \bar z_i}\right| \le O\left(\frac{1}{|\boldsymbol{w}-\boldsymbol{z}|^{2d-1}}\right), \quad i = 1, \dots, d, \tag{6.7.3}$$

where $z_1, \dots, z_d$ are the coordinates of $\boldsymbol{z}$. Moreover, the implicit constants in $O(\cdot)$ are uniform over $\tau^{-1}((L_1 \times L_2) \cap U)$.

Proof. If $d = 1$, then π is an isomorphism and α is a smooth function on $V \times V$, so (6.7.1)–(6.7.3) are trivial.

Therefore, we assume from now on that $d \geq 2$.

For points $(\boldsymbol{w}, \boldsymbol{z}) \in U$, write $\boldsymbol{w} = (w_1, \dots, w_d)$ and $\boldsymbol{z} = (z_1, \dots, z_d)$. Let $v_i = w_i - z_i$ for $i = 1, \dots, d$; then $(v_1, \dots, v_d, z_1, \dots, z_d)$ is a (global) coordinate system on $V \times V$ in which Δ is given by $v_1 = \cdots = v_d = 0$.

For each $l = 1, \dots, d$ let U_l be the subset of points $P \in U$ such that

$$\max\{|v_1(P)|, \dots, |v_d(P)|\} = |v_l(P)|. \tag{6.7.4}$$

Note that $U_1 \cup \cdots \cup U_d = U$ (and that the sets U_l are not open). From now on, for convenience of notation, we assume that $l = 1$ unless otherwise specified.

Let $u_1 = v_1$ and $u_i = v_i / v_1$ for $i = 2, \dots, d$. Then $(u_1, \dots, u_d, z_1, \dots, z_d)$ is a local coordinate system on W_V near all points of $\pi^{-1}(U_1)$. Let W_1 be the largest open subset of W_V on which the functions u_i are regular for all $i \neq 1$. Then $(u_1, \dots, u_d, z_1, \dots, z_d)$ is a coordinate system on W_1, and

$$\pi^{-1}(U_1) = \{P \in W_1 : |u_i(P)| \leq 1 \text{ for all } i \neq 1\}.$$

As l varies, the similarly defined sets W_l cover all of W_V.

Let $q\colon V \times V \to V$ denote the projection to the second factor. Then, on fibers of $q \circ \pi$, we have $du_1 = dv_1 = dw_1$ and

$$du_i = d\left(\frac{v_i}{v_1}\right) = \frac{v_1\, dv_i - v_i\, dv_1}{v_1^2} = \frac{v_1\, dw_i - v_i\, dw_1}{v_1^2}, \qquad i = 2, \dots, d. \tag{6.7.5}$$

By (6.7.4), we have $|v_1| \leq |\boldsymbol{w} - \boldsymbol{z}| \leq \sqrt{d}|v_1|$ over U_1. Then all coefficients 1, $1/v_1$, and $-v_i/v_1^2$ above are bounded in absolute value by $\max\{1, \sqrt{d}/|\boldsymbol{w} - \boldsymbol{z}|\}$ over U_1 (again using $|v_i| \leq |v_1|$). The same estimates hold for the coefficients obtained when writing $d\bar{u}_i$ in terms of $d\bar{w}_1$ and $d\bar{w}_i$ for all $i = 1, \dots, d$.

Next, for all $\boldsymbol{z} \in V$, let $W_{\boldsymbol{z}}$ denote the fiber of $q \circ \pi$ over $\boldsymbol{z}$; it is isomorphic to the blowing-up of V at $\boldsymbol{z}$. For all $\boldsymbol{z} \in V$, we have

$$\alpha|_{W_1 \cap W_{\boldsymbol{z}}} = \sum_{i=1}^{d} \sum_{j=1}^{d} \alpha_{ij} du_1 \wedge \cdots \wedge \widehat{du_i} \wedge \cdots \wedge du_d \wedge d\bar{u}_1 \wedge \cdots \wedge \widehat{d\bar{u}_j} \wedge \cdots \wedge d\bar{u}_d,$$

where $\widehat{\ }$ denotes omission and $\alpha_{ij} : W_1 \to \mathbb{C}$ are smooth. Using the above substitutions for du_i in terms of dw_1 and dw_i, and letting $U_{\boldsymbol{z}} = (V \setminus \{\boldsymbol{z}\}) \times \{\boldsymbol{z}\}$, we then have

$$\alpha|_{W_1 \cap \pi^{-1}(U_{\boldsymbol{z}})} = \sum_{\tilde{\imath}=1}^{d} \sum_{\tilde{\jmath}=1}^{d} \tilde{\alpha}_{\tilde{\imath}\tilde{\jmath}}\, dw_1 \wedge \cdots \wedge \widehat{dw_{\tilde{\imath}}} \wedge \cdots \wedge dw_d \wedge d\bar{w}_1 \wedge \cdots \wedge \widehat{d\bar{w}_{\tilde{\jmath}}} \wedge \cdots \wedge d\bar{w}_d,$$

where

$$\tilde{\alpha}_{\tilde{\imath}\tilde{\jmath}} = \sum_{i,j} \alpha_{ij} \cdot P_{ij\tilde{\imath}\tilde{\jmath}}(1/v_1, 1/\bar{v}_1, v_2/v_1^2, \bar{v}_2/\bar{v}_1^2, \dots, v_d/v_1^2, \bar{v}_d/\bar{v}_1^2) \tag{6.7.6}$$

and each $P_{ij\tilde{\imath}\tilde{\jmath}}$ is a polynomial of degree $2d-2$ with constant coefficients, which depends only on d, i, j, $\tilde{\imath}$, and $\tilde{\jmath}$.

Now we restrict to a hyperplane $H \subseteq T_{\boldsymbol{w}}V$. This hyperplane is given by the vanishing of a nontrivial linear combination of $dw_1, \dots, dw_d$. Therefore there is an index m such that H is given by

$$dw_m = \sum_{i \neq m} c_i \, dw_i \quad \text{with } c_i \in \mathbb{C} \text{ and } |c_i| \leq 1 \text{ for all } i \neq m. \tag{6.7.7}$$

Then, for any locally closed submanifold N of U_z of dimension $d-1$ tangent to H at $\boldsymbol{w}$, we have

$$\alpha|_N = \alpha_m(\boldsymbol{w}, H, z) \, dw_1, \dots \wedge \widehat{dw_m}, \dots \wedge dw_d \wedge d\overline{w}_1, \dots \wedge \widehat{d\overline{w}_m}, \dots \wedge d\overline{w}_d \quad \text{at } \boldsymbol{w}, \tag{6.7.8}$$

where

$$\alpha_m = \sum_{\tilde{\imath}=1}^{d} \sum_{\tilde{\jmath}=1}^{d} \sigma_{\tilde{\imath}m} \sigma_{\tilde{\jmath}m} c_{\tilde{\imath}} \bar{c}_{\tilde{\jmath}} (\tilde{\alpha}_{\tilde{\imath}\tilde{\jmath}} \circ \tau), \tag{6.7.9}$$

$\sigma_{km} = \pm 1$ depending on k and m, and $c_m = 1$.

Now let $K_1 = \{P \in W_1 \cap \pi^{-1}(L_1 \times L_2) : |u_i(P)| \leq 1 \text{ for all } i \neq 1\}$. This set is compact. For all i and j let M_{ij} be the maximum value of $|\alpha_{ij}|$ over K_1. By (6.7.6) there is a constant C_d, depending only on d, such that

$$|\tilde{\alpha}_{\tilde{\imath}\tilde{\jmath}}| \leq \frac{C_d}{\min\{1, |\boldsymbol{w} - z|\}^{2d-2}} \sum_{i,j} M_{ij} \quad \text{on } K_1 \cap \pi^{-1}(U). \tag{6.7.10}$$

Let $K_{1,m}$ be the set of elements of $\tau^{-1}(\pi(K_1) \cap U)$ such that the hyperplane H satisfies (6.7.7). By (6.7.9) and (6.7.10), we then have

$$|\alpha_m| \leq \frac{d^2 C_d}{\min\{1, |\boldsymbol{w} - z|\}^{2d-2}} \sum_{i,j} M_{ij} \quad \text{on } K_{1,m}. \tag{6.7.11}$$

Let K'_m be the set of all points $(\boldsymbol{w}, H) \in \tau_1^{-1}(L_1)$ for which H satisfies (6.7.7). This set is compact and $K'_m \times L_2$ contains $K_{1,m}$. For all $(\boldsymbol{w}, H) \in K'_m$ and all locally closed submanifolds N of V of dimension $d-1$ tangent to H at $\boldsymbol{w}$, we have

$$\psi|_N = (\sqrt{-1})^{d-1} \psi_m(\boldsymbol{w}, H) \, dw_1 \wedge d\overline{w}_1, \dots \wedge \widehat{dw_m \wedge d\overline{w}_m}, \dots \wedge dw_d \wedge d\overline{w}_d \quad \text{at } \boldsymbol{w}, \tag{6.7.12}$$

where $\psi_m : K'_m \to \mathbb{R}$ is continuous and positive. Let $D_m > 0$ be the minimum value of ψ_m on K'_m.

Combining (6.7.8) and (6.7.12) gives

$$\alpha|_N = \frac{(-1)^{(d-2)(d-1)/2} \alpha_m(\boldsymbol{w}, H, z)}{(\sqrt{-1})^{d-1} \psi_m(\boldsymbol{w}, H)} \psi|_N \quad \text{at } (\boldsymbol{w}, H, z)$$

for all $(\boldsymbol{w}, H, z) \in K_{1,m}$. By (6.5.1) and the fact that $\sigma_N(\boldsymbol{w}, z) = (\boldsymbol{w}, H, z)$, we have

$$\chi_\alpha(\boldsymbol{w}, H, z) = \frac{(-1)^{(d-2)(d-1)/2} \alpha_m(\boldsymbol{w}, H, z)}{(\sqrt{-1})^{d-1} \psi_m(\boldsymbol{w}, H, z)},$$

and therefore, by (6.7.11) and the definition of D_m,

$$|\chi_\alpha(\boldsymbol{w}, H, \boldsymbol{z})| \le \frac{d^2 C_d}{D_m} \sum_{i,j} M_{ij} \cdot \frac{1}{\min\{1, |\boldsymbol{w}-\boldsymbol{z}|\}^{2d-2}}$$

for all $(\boldsymbol{w}, H, \boldsymbol{z}) \in K_{1,m}$.

Combining these estimates for all l and all m then gives (6.7.1), uniformly over $\tau^{-1}((L_1 \times L_2) \cap U)$.

Now consider (6.7.2) and (6.7.3).

First of all, it is important to note that the notation $\partial/\partial z_k$ is ambiguous. If taken with respect to the coordinate system $u_1, \dots, u_d, z_1, \dots, z_d$, then $\boldsymbol{u} = (u_1, \dots, u_d)$ is kept fixed (as well as all z_h with $h \ne k$), whereas if taken with respect to the coordinate system $w_1, \dots, w_d, z_1, \dots, z_d$ then $\boldsymbol{w}$ is kept fixed. We denote these (different) partials $\partial_{\boldsymbol{u}}/\partial z_k$ and $\partial_{\boldsymbol{w}}/\partial z_k$, respectively, and define $\partial_{\boldsymbol{u}}/\partial \bar{z}_k$ and $\partial_{\boldsymbol{w}}/\partial \bar{z}_k$ similarly.

The proof of (6.7.2) and (6.7.3) is similar to that of (6.7.1), but is more complicated due to the presence of partial derivatives.

First look at (6.7.5). Recalling that $v_i = w_i - z_i$, we have $\partial_{\boldsymbol{w}} v_i/\partial z_k = -\delta_{ik}$ (using the Kronecker delta), and therefore

$$\frac{\partial_{\boldsymbol{w}}}{\partial z_k}\left(\frac{1}{v_1}\right) = \begin{cases} 1/v_1^2 & \text{if } k = 1, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad \frac{\partial_{\boldsymbol{w}}}{\partial z_k}\left(-\frac{v_i}{v_1^2}\right) = \begin{cases} -v_i/v_1^3 & \text{if } k = 1, \\ 1/v_1^2 & \text{if } k = i, \\ 0 & \text{otherwise.} \end{cases}$$

This gives

$$\max\left\{\left|\frac{\partial_{\boldsymbol{w}}}{\partial z_k}(1)\right|, \left|\frac{\partial_{\boldsymbol{w}}}{\partial z_k}\left(\frac{1}{v_1}\right)\right|, \left|\frac{\partial_{\boldsymbol{w}}}{\partial z_k}\left(-\frac{v_i}{v_1^2}\right)\right|\right\} \le \frac{2}{d \cdot |\boldsymbol{w}-\boldsymbol{z}|^2}. \tag{6.7.13}$$

A similar bound holds for $(\partial_{\boldsymbol{w}}/\partial \bar{z}_k)(1/\bar{v}_1)$ and for $(\partial_{\boldsymbol{w}}/\partial \bar{z}_k)(-v_i/\bar{v}_1^2)$. (Of course we also have that $(\partial_{\boldsymbol{w}}/\partial \bar{z}_k)(1/v_1) = 0$, etc.).

Next we need bounds for $|\partial_{\boldsymbol{w}} \alpha_{ij}/\partial z_k|$ and $|\partial_{\boldsymbol{w}} \alpha_{ij}/\partial \bar{z}_k|$.

From the formulas $u_1 = v_1 = w_1 - z_1$ and $u_h = v_h/v_1 = (w_h - z_h)/(w_1 - z_1)$ for all $h \ne 1$ and the multivariable chain rule, we have

$$\frac{\partial_{\boldsymbol{w}} \alpha_{ij}}{\partial z_k} = \frac{\partial_{\boldsymbol{u}} \alpha_{ij}}{\partial z_k} + \begin{cases} -\frac{\partial \alpha_{ij}}{\partial u_1} + \sum_{h=2}^d \frac{v_h}{v_1^2} \frac{\partial \alpha_{ij}}{\partial u_h} & \text{if } k = 1, \\ -\frac{1}{v_1} \frac{\partial \alpha_{ij}}{\partial u_k} & \text{if } k \ne 1 \end{cases}$$

on $W_1 \cap \pi^{-1}(U)$. Using bounds for $|\partial \alpha_{ij}/\partial u_h|$ and $|\partial_{\boldsymbol{u}} \alpha_{ij}/\partial z_k|$ on K_1, we find constants M_{ijk} such that

$$|\partial_{\boldsymbol{w}} \alpha_{ij} \partial z_k| \le \frac{M_{ijk}}{\min\{1, |\boldsymbol{w}-\boldsymbol{z}|\}} \tag{6.7.14}$$

on $K_1 \cap \pi^{-1}(U)$. A similar argument gives the same bound for $|\partial_{\boldsymbol{w}} \alpha_{ij}/\partial \bar{z}_k|$ (after possibly enlarging M_{ijk}).

By (6.7.13) and (6.7.14), we then have

$$\left|\frac{\partial_{\boldsymbol{w}} \tilde{\alpha}_{\tilde{i}\tilde{j}}}{\partial z_k}\right| \le \frac{1}{\min\{1, |\boldsymbol{w}-\boldsymbol{z}|\}^{2d-1}} \left(C'_d \sum_{i,j} M_{ij} + C''_d \sum_{i,j} M_{ijk}\right) \quad \text{on } K_1 \cap \pi^{-1}(U) \tag{6.7.15}$$

(corresponding to (6.7.10)), where again C_d' and C_d'' depend only on d. Again, the same bound is true for $|\partial_{\boldsymbol{w}}\tilde\alpha_{\tilde\imath\tilde\jmath}/\partial\bar z_k|$ by the same argument.

(Note that the bounds (6.7.13) and (6.7.14) are worse than the corresponding bounds used when proving (6.7.1) by a factor $1/|\boldsymbol{w}-z|$ or $1/\min\{1, |\boldsymbol{w}-z|\}$, so the bound in (6.7.15) is worse than (6.7.10) by that same amount since each term in Leibniz's rule contains only one derivative.)

The rest of the proofs of (6.7.2) and (6.7.3) proceed as for (6.7.1). □

The following lemma applies the preceding lemma to give local information on forms of type (6.3.1).

Lemma 6.8. *Let $V'' \Subset V' \Subset V$ be open subsets of $\mathbb{C}^d$ with $d \ge 1$ and V'' convex. Let Δ, U, and $\pi\colon W_V \to V \times V$ be as in Lemma 6.7. Let α, β, and γ be smooth $(d-1, d-1)$-forms on U such that*

$$\gamma = (-\log|z-\boldsymbol{w}|^2)\alpha + \beta \tag{6.8.1}$$

at all $(\boldsymbol{w}, z) \in U$, and such that $\pi^*\alpha$ and $\pi^*\beta$ extend to smooth forms on W_V. Let $\tau\colon (\mathrm{Gr}^1\, TV) \times V \to V \times V$, ψ, and $\chi_\gamma\colon \tau^{-1}(U) \to \mathbb{C}$ be as in Lemma 6.5. Then there exist real constants $r_0 \in (0, 1]$, c_2, and c_3, depending only on V'', V', ψ, α, and β, such that the bound

$$|\chi_\gamma(\boldsymbol{w}, H, z) - \chi_\gamma(\boldsymbol{w}, H, z')| \le \max\left\{\frac{c_2 + c_3(-\log\rho)}{\rho^{2d-1}}, \frac{c_2 + c_3(-\log\rho')}{(\rho')^{2d-1}}\right\}|z-z'|$$

holds for all $(\boldsymbol{w}, H) \in \mathrm{Gr}^1\, TV$ and all $z, z' \in \overline{V''}$ such that $\boldsymbol{w} \in \overline{V'} \setminus \{z, z'\}$, where

$$\rho = \min\{r_0, |z-\boldsymbol{w}|\} \quad \textit{and} \quad \rho' = \min\{r_0, |z'-\boldsymbol{w}|\}.$$

Proof. Fix $r_0 > 0$ such that $r_0 \le 1$ and r_0 is at most the distance between $\overline{V''}$ and $\mathbb{C}^d \setminus V'$.

Let $\boldsymbol{w}$, H, z, and z' be as in the statement of the lemma. We may assume that $|z-\boldsymbol{w}| \le |z'-\boldsymbol{w}|$. Then $\rho \le \rho'$.

Let $\mathbb{B}$ be the open ball of radius ρ centered at $\boldsymbol{w}$.

We first claim that there is a piecewise smooth path from z to z' of length at most $(\pi/2)|z-z'|$ and lying entirely in $\overline{V'} \setminus \mathbb{B}$. Indeed, start with the straight-line path from z to z'. It lies entirely in $\overline{V''}$. If it does not pass through $\mathbb{B}$, then we are done. Otherwise, replace the segment in $\mathbb{B}$ with a path along a great circle on $\partial\mathbb{B}$ of minimal length that joins the endpoints of that segment. This increases the length of that segment by a factor of at most $\pi/2$, so the revised path has length at most $(\pi/2)|z-z'|$. Also, no point on the great circle is further than $\rho \le r_0$ from a point on the original line segment, so the revised path stays entirely in $\overline{V'}$. (This rerouting can take place within a plane in $\mathbb{C}^d = \mathbb{R}^{2d}$ that contains the three points $\boldsymbol{w}$, z, and z'.)

Let $\boldsymbol{y}\colon [0, \ell] \to \overline{V'} \setminus \mathbb{B}$ be this path, parametrized by arc length. It will then suffice to show that

$$\left|\frac{d}{dt}\chi_\gamma(\boldsymbol{w}, H, \boldsymbol{y}(t))\right| \le \frac{2}{\pi}\cdot\frac{c_2 + c_3(-\log\rho)}{\rho^{2d-1}} \tag{6.8.2}$$

at smooth points of the path, since that would give

$$|\chi_\gamma(\boldsymbol{w}, H, z) - \chi_\gamma(\boldsymbol{w}, H, z')| \le \frac{c_2 + c_3(-\log\rho)}{\rho^{2d-1}}|z-z'|.$$

To see (6.8.2), let χ_α and χ_β be as in Lemma 6.5. Then

$$\chi_\gamma(\boldsymbol{w}, H, \boldsymbol{y}) = (-\log|\boldsymbol{y}-\boldsymbol{w}|^2)\chi_\alpha(\boldsymbol{w}, H, \boldsymbol{y}) + \chi_\beta(\boldsymbol{w}, H, \boldsymbol{y})$$

for all $\boldsymbol{y} \in V$. Then (6.8.2) follows from the bounds (6.7.1)–(6.7.3) applied to χ_α and χ_β, together with the inequality $|(d/dt)(-\log|\boldsymbol{y}(t)-\boldsymbol{w}|)| \le 1/\rho$ at smooth points of the path. □

This can be translated to a result on the complex manifold X.

Corollary 6.9. *Let X be a smooth complex projective variety of dimension $d \ge 1$. Let ψ be a positive smooth $(d-1, d-1)$-form on X. Let (U, ϕ) be a coordinate chart on X, and let $U'' \Subset U$ be a nonempty open subset such that $\phi(U'')$ is convex. Then there is a measurable function $f\colon X \times \overline{U''} \to [0, \infty]$ such that*

(i) *for all $\xi \in \boldsymbol{\kappa}(X)^*$, the inequality*

$$\big|-\log|\xi(x)| + \log|\xi(x')|\big| \le \frac{|\phi(x)-\phi(x')|}{2} \sum_D |n_D| \int_{D_{\mathrm{reg}}} (f(w,x) + f(w,x')) \cdot \psi(w) \qquad (6.9.1)$$

holds for all $x, x' \in \overline{U''}$, where $(\xi) = \sum_D n_D D$ as in Proposition 6.3(c); and

(ii) *there exists a constant c_4, depending only on X, γ, ψ, U, U'', and ϕ, such that*

$$\int_{\overline{U''}} f(w,x)\, d\phi^*\mu(x) \le c_4 \qquad (6.9.2)$$

for all $w \in X$, where μ is the standard measure on $\mathbb{C}^d$.

Proof. Let g_Δ be as in Proposition 6.3, write $\gamma = g_\Delta$, and let χ_γ be as in Lemma 6.5 (applied with $M_1 = M_2 = X$ and $\alpha = \gamma$).

We first claim that there exists a function f for which the inequality

$$|\chi_\gamma(w, H, x) - \chi_\gamma(w, H, x')| \le 2\max\{f(w,x), f(w,x')\}|\phi(x)-\phi(x')| \qquad (6.9.3)$$

holds for all $w \in X$ and all $x, x' \in \overline{U''} \setminus \{w\}$.

Pick an open subset U' such that $U'' \Subset U' \Subset U$. Let $\tau_1\colon \operatorname{Gr}^1 TX \to X$ be as in Lemma 6.5.

Note that

$$\chi_{(\phi^{-1})^*\gamma}(\phi(w), H, \phi(x)) = \chi_\gamma(w, H, x) \qquad (6.9.4)$$

for all $(w, H) \in \tau_1^{-1}(U)$ and all $x \in U \setminus \{w\}$, and that $(\phi^{-1})^*\gamma$ is of the form (6.8.1) (using the fact that if s and $\|\cdot\|$ are as in Proposition 6.3 then the function $(\boldsymbol{w}, z) \mapsto -\log\|s(\phi^{-1}(\boldsymbol{w}), \phi^{-1}(z))\|^2 + \log|z-\boldsymbol{w}|^2$ extends to a smooth function on $\phi(U) \times \phi(U)$).

Then, by Lemma 6.8, there are real constants $r_0 > 0$, c_2, and c_3, such that, letting

$$f(w,x) = \frac{c_2 + c_3(-\log\min\{r_0, |\phi(x)-\phi(w)|\})}{2\min\{r_0, |\phi(x)-\phi(w)|\}^{2d-1}}$$

for all $w \in U'$ and all $x \in \overline{U''} \setminus \{w\}$, (6.9.3) holds whenever $w \in U'$.

Since the set $\tau_1^{-1}(X \setminus U') \times \overline{U''}$ is compact and χ_γ is smooth on an open neighborhood of this set, there is a constant c_4 such that

$$|D_z((w, H, z) \mapsto \chi_\gamma(w, H, \phi^{-1}(z)))| \le c_4$$

for all $(w, H) \in \tau_1^{-1}(X \setminus U')$ and all $z \in \phi(\overline{U''})$. Here D_z means the vector consisting of all partial derivatives in the coordinates of z. Then, letting

$$f(w, x) = c_4$$

for all $w \notin U'$, it now follows that (6.9.3) holds without additional restrictions on w.

By (6.3.3), (6.6.1), and (6.9.3), we then have

$$\begin{aligned} 2\big|-\log|\xi(x)| + \log|\xi(x')|\big| &= \Big|\sum_D n_D(g_D(x) - g_D(x'))\Big| \\ &\le \sum_D |n_D| \int_{D_{\text{reg}}} |\chi_\gamma(\sigma_{D_{\text{reg}},1}(w), x) - \chi_\gamma(\sigma_{D_{\text{reg}},1}(w), x')| \cdot \psi(w) \\ &\le 2|\phi(x) - \phi(x')| \sum_D |n_D| \int_{D_{\text{reg}}} \max\{f(w, x), f(w, x')\} \cdot \psi(w), \end{aligned}$$

and this gives (6.9.1).

Finally, (6.9.2) follows from the fact that $\phi(\overline{U''})$ is bounded and that the integrals

$$\int_{\mathbb{D}^d} \frac{d\mu(z)}{|z|^{2d-1}} \quad \text{and} \quad \int_{\mathbb{D}^d} \frac{\log|z|}{|z|^{2d-1}}\, d\mu(z)$$

converge. □

The next lemma combines Corollaries 6.6 and 6.9 to show that $-\log|\xi|$ obeys a Lipshitz condition after removing a set of arbitrarily small (but nonzero) measure, with prescribed uniformities.

We start with a definition.

Definition 6.10. Let X be a smooth projective variety of dimension $d \ge 1$, and let $\mathscr{M}$ be an ample line sheaf on X:

(a) For all divisors D on X, let

$$\deg_{\mathscr{M}} D = c_1(\mathscr{M})^{\cdot(d-1)} \cdot D.$$

(b) For all $\xi \in \boldsymbol{\kappa}(X)^*$, write $(\xi) = \sum_D n_D D$ as in Proposition 6.3(c). Then we let

$$\deg_{\mathscr{M}} \xi = \tfrac{1}{2} \sum_D |n_D| \deg_{\mathscr{M}} D. \tag{6.10.1}$$

If, moreover, X is a variety over $\mathbb{C}$ and if $\mathscr{M}$ is a smoothly metrized line sheaf on X such that $\mathscr{M}_{\text{fin}}$ is ample, then $\deg_{\mathscr{M}}$ is defined to be $\deg_{\mathscr{M}_{\text{fin}}}$ in the above two contexts.

Remark 6.11. Let X, $\mathscr{M}$, and ξ be as above. Then the divisors

$$(\xi)_0 := \sum_D \max\{0, n_D\} D \quad \text{and} \quad (\xi)_\infty := \sum_D \max\{0, -n_D\} D$$

are linearly equivalent, so

$$\deg_{\mathscr{M}} \xi = \deg_{\mathscr{M}} (\xi)_\infty. \tag{6.11.1}$$

In particular, if $X = \mathbb{P}^1$ and $\mathscr{M} = \mathscr{O}(1)$, then $\deg_{\mathscr{M}} \xi$ coincides with the degree of ξ as a rational function.

Remark 6.12. Let X be a smooth complex projective variety of dimension $d \ge 1$, let $\mathscr{M}$ be a smoothly metrized line sheaf on X such that $\mathscr{M}_{\text{fin}}$ is ample, and let D be a prime divisor on X. Then

$$\deg_{\mathscr{M}} D = \int_D c_1(\|\cdot\|_{\mathscr{M}})^{\wedge(d-1)} = \int_{D_{\text{reg}}} c_1(\|\cdot\|_{\mathscr{M}})^{\wedge(d-1)}. \tag{6.12.1}$$

Therefore if $\xi \in \boldsymbol{\kappa}(X)^*$, then by (6.10.1)

$$\deg_{\mathscr{M}} \xi = \tfrac{1}{2} \sum_D |n_D| \int_{D_{\text{reg}}} c_1(\|\cdot\|_{\mathscr{M}})^{\wedge(d-1)}. \tag{6.12.2}$$

Lemma 6.13. *Let X, U, U'' and ϕ be as in Corollary 6.9, and let $\mathscr{M}$ be a smoothly metrized line sheaf on X with positive metric. Then for all $\epsilon_1 > 0$ there is a constant c_5 such that the following is true. For each $\xi \in \boldsymbol{\kappa}(X)^*$ there is a closed subset T of $\overline{U''}$ such that $\mu(\phi(T)) \le \epsilon_1$ and such that the inequality*

$$\big| -\log|\xi(x)| + \log|\xi(x')| \big| \le c_5 (\deg_{\mathscr{M}} \xi) |\phi(x) - \phi(x')| \tag{6.13.1}$$

holds for all $x, x' \in \overline{U''} \setminus T$.

Proof. We apply Corollary 6.9 with $\psi = c_1(\|\cdot\|_{\mathscr{M}})^{\wedge(d-1)}$ (note that ψ is positive by Proposition 1.5(a)). This gives a function $f : X \times \overline{U''} \to [0, \infty]$ and a constant c_4 that satisfy (6.9.1) and (6.9.2). Let

$$c_5 = \frac{4c_4}{\epsilon_1}.$$

Let $\xi \in \boldsymbol{\kappa}(X)^*$, and write $(\xi) = \sum_D n_D D$ as in Proposition 6.3(c). By (6.9.1), it then suffices to construct a suitable set T such that

$$\sum_D |n_D| \int_{D_{\text{reg}}} (f(w, x) + f(w, x')) \cdot \psi(w) \le c_5 \deg_{\mathscr{M}} \xi$$

for all $x, x' \in \overline{U''} \setminus T$. For this, in turn, it suffices to find T such that

$$\sum_D |n_D| \int_{D_{\text{reg}}} f(w, x) \cdot \psi(w) \le \frac{c_5}{2} \deg_{\mathscr{M}} \xi \tag{6.13.2}$$

for all $x \in \overline{U''} \setminus T$.

Let $g : \overline{U''} \to [0, \infty]$ be the function defined by

$$g(x) = \sum_D |n_D| \int_{D_{\text{reg}}} f(w, x) \cdot \psi(w).$$

Then (6.13.2) holds with

$$T = \left\{ x \in \overline{U''} : g(x) \ge \frac{c_5}{2} \deg_{\mathscr{M}} \xi \right\}.$$

It remains only to show that $\mu(\phi(T)) \le \epsilon_1$. Indeed, by Tonelli's theorem, (6.9.2), and (6.12.2), we have

$$\begin{aligned}\int_{\overline{U''}} g(x)\, d\phi^*\mu &= \sum_D |n_D| \int_{x\in\overline{U''}} \int_{w\in D_{\mathrm{reg}}} f(w,x)\cdot\psi(w)\, d\phi^*\mu(x) \\ &= \sum_D |n_D| \int_{w\in D_{\mathrm{reg}}} \int_{x\in\overline{U''}} f(w,x)\, d\phi^*\mu(x)\cdot\psi(w) \\ &\le \sum_D |n_D| \int_{D_{\mathrm{reg}}} c_4\psi \\ &= 2c_4 \deg_{\mathscr{M}} \xi \\ &= \frac{\epsilon_1 c_5}{2} \deg_{\mathscr{M}} \xi.\end{aligned}$$

Then $\mu(\phi(T)) \le \epsilon_1$ by Lemma 6.1. □

Coordinate charts as in Corollary 6.9 and Lemma 6.13 will now be used to obtain global results on X, via the following construction.

Let X be a smooth complex projective variety of dimension $d \ge 1$. Since X is compact, there exists a finite collection

$$\{(U_i, \phi_i, U_i'') : i = 1, \dots, n\} \tag{6.14}$$

with $U_1'', \dots, U_n''$ covering X, such that for each i, (U_i, ϕ_i) is a coordinate chart on X, $U_i'' \Subset U_i$ is a nonempty open subset, and $\phi_i(U_i'')$ is convex.

Let $\mathscr{M}$ be a smoothly metrized line sheaf on X with positive metric, and let $\theta = c_1(\|\cdot\|_{\mathscr{M}})^{\wedge d}$. This is a positive (d,d)-form by Proposition 1.5(a), so it defines a measure μ_θ on X. For all i, the measures μ_θ and $\phi_i^*\mu$ on U_i are related by $\mu_\theta = \rho_i \cdot \phi_i^*\mu$, where $\rho_i : U_i \to \mathbb{R}_{>0}$ is smooth. Since $\overline{U_i''}$ is compact, there are constants $c_{6,i}$ and $c_{7,i}$ such that

$$c_{6,i}\phi_i^*\mu \le \mu_\theta \le c_{7,i}\phi_i^*\mu \tag{6.15}$$

on $\overline{U_i''}$.

This construction then leads to the main result of this section.

Proposition 6.16. *Let X be a smooth complex projective variety of dimension $d \ge 1$, let $\mathscr{M}$ be a smoothly metrized line sheaf on X with positive metric, let $\theta = c_1(\|\cdot\|_{\mathscr{M}})^{\wedge d}$, and let μ_θ be the corresponding measure on X. Then, for all $\epsilon_2 > 0$ and $\epsilon_3 > 0$ there is a finite collection of subsets $C_1, \dots, C_\Lambda$ of X such that $\bigcup_l C_l = X$ and such that the following is true. For each $\xi \in \boldsymbol{\kappa}(X)^*$ there is a measurable subset T of X such that $\mu_\theta(T) \le \epsilon_2$ and such that*

$$\left| -\log^-|\xi(x)| + \log^-|\xi(x')| \right| \le \epsilon_3 \deg_{\mathscr{M}} \xi \tag{6.16.1}$$

for all $x, x' \in C_l \setminus T$ and all $l = 1, \dots, \Lambda$.

Proof. Choose triples (U_i, ϕ_i, U_i'') as in (6.14), and fix for now an index i. Let $c_{7,i}$ be as in (6.15).

By Lemma 6.13, there is a constant $c_{5,i}$ such that for each $\xi \in \boldsymbol{\kappa}(X)^*$ there is a subset $T_i \subseteq \overline{U_i''}$ such that $\mu(\phi_i(T_i)) \le \epsilon_2/nc_{7,i}$ and such that (6.13.1) holds for all $x, x' \in \overline{U_i''} \setminus T_i$.

Choose subsets $C_{i,1}, \dots, C_{i,\Lambda_i}$ of $\overline{U_i''}$ such that $\bigcup_l C_{i,l} = \overline{U_i''}$ and such that $\phi(C_{i,l})$ has diameter at most $\epsilon_3/c_{5,i}$ for all l. Let $\xi \in \boldsymbol{\kappa}(X)^*$. The function $f(y) = \min\{0, y\}$ satisfies $|f(y) - f(y')| \le |y - y'|$ for all $y, y' \in \mathbb{R}$. Combining this with (6.13.1) and the above diameter bound, we have

$$\begin{aligned}\big|-\log^-|\xi(x)| + \log^-|\xi(x')|\big| &\le \big|-\log|\xi(x)| + \log|\xi(x')|\big| \\ &\le c_{5,i}(\deg_{\mathscr{M}} \xi)|\phi(x) - \phi(x')| \\ &\le \epsilon_3 \deg_{\mathscr{M}} \xi\end{aligned}$$

for all $l = 1, \dots, \Lambda_i$ and all $x, x' \in C_{i,l} \setminus T_i$, where T_i is the subset chosen above for the given ξ.

Now, letting i vary, let $C_1, \dots, C_\Lambda$ be the collection of all $C_{i,l}$. Given ξ as above, let $T = \bigcup_i T_i$; then

$$\mu_\theta(T) \le \sum_{i=1}^n c_{7,i}\,\mu(\phi_i(T_i)) \le \sum_{i=1}^n \frac{\epsilon_2}{n} = \epsilon_2,$$

and (6.16.1) holds for T. □

7. Reduction to simultaneous approximation: The excluded set T

Proposition 6.16 in the previous section involved excluding a set T, which can be chosen to have arbitrarily small measure. This section provides the key estimate needed in order to show that excluding this set does not affect the diophantine estimates excessively.

7.1. Throughout this section, X is a smooth complex projective variety of dimension $d \ge 1$, $\mathscr{M}$ is a smoothly metrized line sheaf on X with positive metric, $\theta = c_1(\|\cdot\|_{\mathscr{M}})^{\wedge d}$, $\psi = c_1(\|\cdot\|_{\mathscr{M}})^{\wedge(d-1)}$, and μ_θ is the measure on X associated to θ.

We start with some definitions.

Definition 7.2. Let

$$\deg_{\mathscr{M}} X = c_1(\mathscr{M}_{\text{fin}})^{\cdot d} = \int_X \theta,$$

and let

$$h_X(\xi) = \int_X -\log^-|\xi|^2 \cdot \theta$$

for all $\xi \in \boldsymbol{\kappa}(X)^*$.

The main result of this section is then the following.

Proposition 7.3. *Let $X, d, \mathscr{M}, \theta$, and μ_θ be as in 7.1. Then for all $\epsilon_4 > 0$ there is an $\epsilon_5 > 0$ such that the inequality*

$$\int_T -\log^-|\xi|^2 \cdot \theta \le \epsilon_4 \deg_{\mathscr{M}} \xi + \frac{\mu_\theta(T)}{\deg_{\mathscr{M}} X}(2h_X(\xi) + c_8 \deg_{\mathscr{M}} \xi) \tag{7.3.1}$$

holds for all $\xi \in \boldsymbol{\kappa}(X)^*$ and all measurable $T \subseteq X$ with $\mu_\theta(T) \le \epsilon_5$. Here c_8 is a constant that depends only on X and $\mathscr{M}$.

To prove the proposition, we write

$$-\log|\xi|^2 = \sum_D n_D g_D + c_\xi \tag{7.4}$$

for all $\xi \in \boldsymbol{\kappa}(X)^*$ as in (6.3.3), and bound the integrals of each term on the right-hand side separately.

Lemma 7.5. *Let (U, ϕ) be a coordinate chart on X, and let $U'' \Subset U$ be a nonempty open subset. Then for all $\epsilon_6 > 0$ there is an $\epsilon_7 > 0$ such that the following is true. Let $\xi \in \boldsymbol{\kappa}(X)^*$, and write $(\xi) = \sum_D n_D D$ in the notation of (7.4). Then for all measurable subsets $T \subseteq \overline{U''}$ such that $\mu(\phi(T)) \le \epsilon_7$, we have*

$$\sum_D n_D \int_T g_D(x)\, d\phi^*\mu(x) \le \epsilon_6 \deg_{\mathscr{M}} \xi. \tag{7.5.1}$$

Proof. Let $\gamma = g_\Delta$ and let χ_γ be as in Lemma 6.5. By (6.6.1) and Tonelli's theorem, (7.5.1) is equivalent to

$$\sum_D n_D \int_{D_{\mathrm{reg}}} \int_T \chi_\gamma(\sigma_{D_{\mathrm{reg}},1}(w), x)\, d\phi^*\mu(x) \cdot \psi(w) \le \epsilon_6 \deg_{\mathscr{M}} \xi. \tag{7.5.2}$$

To prove this, it suffices to show that the inequality

$$\int_T |\chi_\gamma(w, H, x)|\, d\phi^*\mu(x) \le \frac{\epsilon_6}{2} \tag{7.5.3}$$

holds for all $w \in X$, all H, and all $T \subseteq \overline{U''}$ with $\mu(\phi(T)) \le \epsilon_7$ (where ϵ_7 is to be chosen later). Indeed, integrating (7.5.3) and applying (6.12.2) implies (7.5.2).

To show (7.5.3), choose an open subset $U' \subseteq U$ such that $U'' \Subset U' \Subset U$, and let $V = \phi(U)$, $V' = \phi(U')$, and $V'' = \phi(U'')$. Fix $r_0 \in (0, 1]$ such that r_0 is at most the distance between $\overline{V''}$ and $\mathbb{C}^d \setminus V'$. By (6.9.4) and the fact that $(\phi^{-1})^*\gamma$ is of the form (6.8.1), we obtain from (6.7.1) that there are constants c and c', depending only on X, γ, ψ, U, U', U'', ϕ, and r_0, such that

$$|\chi_\gamma(w, H, x)| \le \frac{c + c'(-\log \rho)}{\rho^{2d-2}} \tag{7.5.4}$$

for all $w \in \overline{U'}$, all H, and all $x \in \overline{U''} \setminus \{w\}$, where

$$\rho = \min\{r_0, |\phi(w) - \phi(x)|\}.$$

We may assume that $c, c' \ge 0$.

Next, we claim that for all $\epsilon_6 > 0$ there is an $\epsilon_7 > 0$ such that, for all $\tilde{T} \subseteq \overline{V''}$ with $\mu(\widetilde{T}) \le \epsilon_7$ and for all $\boldsymbol{w} \in V'$, we have

$$\int_{\tilde{T}} \frac{c + c'(-\log \min\{r_0, |\boldsymbol{w} - z|\})}{\min\{r_0, |\boldsymbol{w} - z|\}^{2d-2}}\, d\mu(z) \le \frac{\epsilon_6}{2}. \tag{7.5.5}$$

Basically, this follows from the fact that the integrand is a function of $\boldsymbol{w} - z$, and that the latter function is locally L^1.

In more detail, let $\mathbb{D}_r^d = \{z \in \mathbb{C}^d : |z| < r\}$. The integral in (7.5.6) converges for all $r > 0$; therefore there is a number $r > 0$ such that

$$\int_{\mathbb{D}_r^d} \frac{c + c'(-\log \min\{r_0, |z|\})}{\min\{r_0, |z|\}^{2d-2}} \, d\mu(z) \le \frac{\epsilon_6}{2}. \tag{7.5.6}$$

Pick such an r and let $\epsilon_7 = \mu(\mathbb{D}_d^r)$. Then

$$\int_{\tilde{T}} \frac{c + c'(-\log \min\{r_0, |z|\})}{\min\{r_0, |z|\}^{2d-2}} \, d\mu(z) \le \int_{\mathbb{D}_r^d} \frac{c + c'(-\log \min\{r_0, |z|\})}{\min\{r_0, |z|\}^{2d-2}} \, d\mu(z) \le \frac{\epsilon_6}{2}$$

for all $\tilde{T} \subseteq \mathbb{C}^d$ with $\mu(\tilde{T}) \le \epsilon_7$. This then gives (7.5.5) by translation.

Combining (7.5.5) with (7.5.4) then gives (7.5.3) for all $w \in U'$.

Next consider $w \notin U'$. Let $\tau_1 \colon \operatorname{Gr}^1 TX \to X$ be as in Lemma 6.5, and let c'' be the maximum of $|\chi_\gamma|$ over the compact set $\tau_1^{-1}(X \setminus U') \times \overline{U''}$. We then have

$$\int_T |\chi_\gamma(w, H, x)| \, d\phi^*\mu(x) \le c''\epsilon_7$$

for all $w \in X \setminus U'$, all H, and all $T \subseteq \overline{U''}$ for which $\mu(\phi(T)) \le \epsilon_7$.

Assume now that ϵ_7 has been chosen so that $c''\epsilon_7 \le \epsilon_6/2$. Then (7.5.3) holds also for all $w \notin U'$, so it holds for all $w \in X$. □

The following lemma translates the above lemma into the global setting.

Lemma 7.6. *For all $\epsilon_4 > 0$ there is an $\epsilon_5 > 0$ such that the following is true. Let $\xi \in \boldsymbol{\kappa}(X)^*$, and write $(\xi) = \sum_D n_D D$ in the notation of (7.4). Then for all measurable subsets $T \subseteq X$ such that $\mu_\theta(T) \le \epsilon_5$, we have*

$$\sum_D n_D \int_T g_D(x) \, d\mu_\theta(x) \le \epsilon_4 \deg_{\mathscr{M}} \xi. \tag{7.6.1}$$

Proof. Choose triples (U_i, ϕ_i, U_i'') as in (6.14), and fix for now an index i. Let $c_{6,i}$ and $c_{7,i}$ be as in (6.15).

By Lemma 7.5 there is an $\epsilon_{7,i} > 0$ such that (7.5.1) holds with $\epsilon_6 = \epsilon_4/nc_{7,i}$ for all $T \subseteq \overline{U_i''}$ with $\mu(\phi_i(T)) \le \epsilon_{7,i}$ and all $\xi \in \boldsymbol{\kappa}(X)^*$. By (7.5.1) and (6.15),

$$\sum_D n_D \int_T g_D(x) \, d\mu_\theta(x) \le \frac{\epsilon_4}{n} \deg_{\mathscr{M}} \xi$$

for all such T and ξ.

Now let

$$\epsilon_5 = \min_{1 \le i \le n} c_{6,i}\epsilon_{7,i}.$$

Let T be a measurable subset of X with $\mu_\theta(T) \le \epsilon_5$. By (6.15), we have

$$\mu(\phi_i(T \cap \overline{U_i''})) \le \epsilon_5/c_{6,i} \le \epsilon_{7,i}$$

for all i, and therefore

$$\sum_D n_D \int_{T\cap \overline{U_i''}} g_D(x)\,d\mu_\theta(x) \le \frac{\epsilon_4}{n} \deg_{\mathscr{M}} \xi$$

holds for all ξ and all i. Summing over i then gives (7.6.1). □

The next step in proving Proposition 7.3 is to find an upper bound for c_ξ.

To find this bound, we first find an upper bound for

$$c'_\xi := \frac{1}{\deg_{\mathscr{M}} X} \int_X -\log|\xi|^2 \cdot \theta \tag{7.7}$$

(this is the average value of $-\log|\xi|^2$ over X).

Lemma 7.8. *Let $\xi \in \boldsymbol{\kappa}(X)^*$. Then*

$$c'_\xi \le \frac{2}{\deg_{\mathscr{M}} X} h_X(\xi).$$

Proof. Let $\xi_0 = e^{c'_\xi/2}\xi$, so that $-\log|\xi(x)|^2 = -\log|\xi_0(x)|^2 + c'_\xi$ and therefore

$$\int_X -\log|\xi_0|^2 \cdot \theta = 0.$$

Hence

$$h_X(\xi_0) = \int_X -\log^-|\xi_0|^2 \cdot \theta = \int_X \log^+|\xi_0|^2 \cdot \theta. \tag{7.8.1}$$

Let $\lambda(x) = -\log|\xi_0(x)|^2$ for all $x \in X$ outside of the support of the principal divisor $(\xi_0) = (\xi)$, and for $t \in \mathbb{R}$ let

$$f(t) = \int_X \max\{0, \lambda + t\} \cdot \theta.$$

Then $h_X(\xi) = f(c'_\xi)$, so it suffices to show that

$$f(t) \ge \frac{\deg_{\mathscr{M}} X}{2} t \tag{7.8.2}$$

for all $t \in \mathbb{R}$.

Note that f is continuous, and is differentiable outside a countable set. Also

$$f'(t) = \mu_\theta(\{x \in X : \lambda(x) + t \ge 0\}) \tag{7.8.3}$$

wherever $f'(t)$ is defined. By abuse of notation, we use (7.8.3) to extend f' to a function on all of $\mathbb{R}$. Note that f' is an increasing function of t, so f is concave upward. Also

$$\lim_{t\to-\infty} f(t) = \lim_{t\to-\infty} f'(t) = 0 \quad \text{and} \quad \lim_{t\to\infty} f'(t) = \deg_{\mathscr{M}} X.$$

Let

$$\beta = \sup\left\{t : f'(t) \le \frac{\deg_{\mathscr{M}} X}{2}\right\}.$$

Then, by concavity, it suffices to show that (7.8.2) holds when $t = \beta$.

This is trivial when $\beta \le 0$, so assume that $\beta > 0$.

We have

$$\mu_\theta(\{x : \lambda(x) + \beta > 0\}) = \lim_{n\to\infty} f'\left(\beta - \frac{1}{n}\right) \le \frac{\deg_{\mathscr{M}} X}{2};$$

hence

$$\mu_\theta(\{x : \log|\xi_0(x)|^2 \ge \beta\}) = \mu(\{x : \lambda(x) \le -\beta\}) \ge \frac{\deg_{\mathscr{M}} X}{2}.$$

Then, by (7.8.1) and trivial properties of integration,

$$f(\beta) \ge f(0) = h_X(\xi_0) = \int_X \log^+|\xi_0(x)|^2 \cdot \theta \ge \frac{\deg_{\mathscr{M}} X}{2}\beta.$$

□

To bound c_ξ, it then suffices to compare c_ξ and c'_ξ.

Lemma 7.9. *There is a constant c_8, depending only on X and $\mathscr{M}$, such that*

$$|c'_\xi - c_\xi| \le \frac{c_8 \deg_{\mathscr{M}} \xi}{\deg_{\mathscr{M}} X}$$

for all $\xi \in \boldsymbol{\kappa}(X)^*$.

Proof. Let $V'' \Subset V' \Subset V$, γ, and χ_γ be as in Lemma 6.8. By (6.8.1) and (6.7.1),

$$|\chi_\gamma(\boldsymbol{w}, H, \boldsymbol{z})| \le O\left(\frac{\max\{1, -\log|\boldsymbol{z} - \boldsymbol{w}|\}}{|\boldsymbol{z} - \boldsymbol{w}|^{2d-2}}\right)$$

for all $\boldsymbol{w} \in V''$ and all $\boldsymbol{z} \in V' \setminus \{\boldsymbol{w}\}$, where the implicit constant is independent of $\boldsymbol{z}$ and $\boldsymbol{w}$. Therefore

$$\int_{V'} |\chi_\gamma(\boldsymbol{w}, H, \boldsymbol{z})|\, d\mu(\boldsymbol{z}) \le O(1)$$

for all $\boldsymbol{w} \in V''$ and all H, uniformly in $\boldsymbol{w}$ and H.

Let $U'' \Subset U \subset X$ and $\phi \colon U \to \mathbb{C}^d$ be as in Corollary 6.9, let γ and χ_γ be as in the proof of Corollary 6.9, and let U' be an open subset of X with $U'' \Subset U' \Subset U$. Then, by (6.15),

$$\int_{U'} |\chi_\gamma(w, H, x)| \cdot \theta(x) \le O(1)$$

for all $w \in U''$ and all H, uniformly in w and H. A standard compactness argument on $\tau_1^{-1}(\overline{U''}) \times (X \setminus U')$ gives a similar bound on $\int_X |\chi_\gamma(w, H, x)| \cdot \theta(x)$ for all $w \in U''$ and all H.

Applying this bound to all charts in a finite set of charts as in (6.15) then gives a constant c_8 such that

$$\int_X |\chi_\gamma(w, H, x)| \cdot \theta(x) \le \frac{c_8}{2} \tag{7.9.1}$$

for all $w \in X$ and all H.

By (7.7), (7.4), (6.6.1), Tonelli's theorem, (7.9.1), and (6.12.2), we then have

$$\begin{aligned}
|c'_\xi - c_\xi| &= \frac{1}{\deg_{\mathscr{M}} X}\left|\int_X \sum_D n_D g_D \cdot \theta\right| \\
&= \frac{1}{\deg_{\mathscr{M}} X}\left|\int_X \left(\sum_D n_D \int_{D_{\mathrm{reg}}} \chi_\gamma(\sigma_{D_{\mathrm{reg}},1}(w), x)\cdot\psi(w)\right)\cdot\theta\right| \\
&\le \frac{1}{\deg_{\mathscr{M}} X}\sum_D |n_D| \int_{D_{\mathrm{reg}}}\int_X \left|\chi_\gamma(\sigma_{D_{\mathrm{reg}},1}(w), x)\right|\cdot\theta(x)\cdot\psi(w) \\
&\le \frac{1}{\deg_{\mathscr{M}} X}\sum_D |n_D| \int_{D_{\mathrm{reg}}} \frac{c_8}{2}\cdot\psi(w) \\
&= \frac{c_8 \deg_{\mathscr{M}} \xi}{\deg_{\mathscr{M}} X}.
\end{aligned}$$

□

The proof of Proposition 7.3 is then a matter of combining these lemmas, as follows.

Proof of Proposition 7.3. Let $\xi \in \boldsymbol{\kappa}(X)^*$, and let T be as in the statement of the proposition. Let $T' = \{x \in T : |\xi(x)| < 1\}$. Then $\mu_\theta(T') \le \mu_\theta(T)$ and

$$\int_{T'} -\log|\xi|^2 \cdot \theta = \int_T -\log^-|\xi|^2 \cdot \theta,$$

so instead of (7.3.1) it will suffice to prove

$$\int_T -\log|\xi|^2 \cdot \theta \le \epsilon_4 \deg_{\mathscr{M}} \xi + \frac{\mu_\theta(T)}{\deg_{\mathscr{M}} X}(2h_X(\xi) + c_8 \deg_{\mathscr{M}} \xi) \tag{7.10}$$

for all T as in the proposition.

Given $\epsilon_4 > 0$, let $\epsilon_5 > 0$ be as in Lemma 7.6. Then (7.10) follows from (7.4) and Lemmas 7.6, 7.8 and 7.9. □

8. Reduction to simultaneous approximation: Arithmetic

This section translates Propositions 6.16 and 7.3 into the arithmetic setting, and proves a result on reduction to simultaneous approximation (Proposition 8.12) that will be sufficient to prove Roth's theorem.

Recall from 4.1 that K is an arithmetic function field, that $M = (B; \mathscr{M})$ is a (big) polarization of K, and that $S \subseteq M_K$ is a subset of finite measure. Also recall from 5.1–5.4 that S contains all archimedean places of K, that B is generically smooth, and that $\mathscr{M}$ is ample with positive metric. Finally, let d be the transcendence degree of K over $\mathbb{Q}$.

Let F be the algebraic closure of $\mathbb{Q}$ in K (i.e., the set of all elements of K that are algebraic over $\mathbb{Q}$). It is a number field (by [Lang 2002, Chapter VIII, Exercise 4] it is finitely generated over $\mathbb{Q}$, and by definition it is algebraic over $\mathbb{Q}$; hence $[F : \mathbb{Q}] < \infty$).

Since B is normal and $\mathscr{O}_F$ is integral over $\mathbb{Z}$, the canonical morphism $B \to \operatorname{Spec} \mathbb{Z}$ factors uniquely through a morphism $\pi : B \to \operatorname{Spec} \mathscr{O}_F$. Also, we write $B_F = B \times_{\mathscr{O}_F} F$, and if $\mathscr{L}$ is a continuously metrized line sheaf on B then $\mathscr{L}_F$ will denote the pull-back of $\mathscr{L}_{\text{fin}}$ to B_F.

For any embedding $\sigma : F \to \mathbb{C}$, we let $\mathbb{C}_\sigma$ denote the field $\mathbb{C}$, viewed as an extension field of F via σ, and let $B_\sigma = (B_F \times_F \mathbb{C}_\sigma)^{\text{an}}$. We then have

$$B(\mathbb{C}) = \coprod_{\sigma : F \to \mathbb{C}} B_\sigma .$$

By [EGA IV$_2$ 1965, EGA IV, 4.5.10], B_F is geometrically integral over F. Therefore the schemes $B_F \times_F \mathbb{C}_\sigma$ are integral for all σ, and the B_σ correspond to the irreducible components of $B \times_{\mathbb{Z}} \mathbb{C}$.

Let $\mathscr{L}$ be a continuously metrized line sheaf on B. For all $\sigma : F \hookrightarrow \mathbb{C}$, we let $\mathscr{L}_\sigma$ denote the restriction $\mathscr{L}_{\mathbb{C}}|_{B_\sigma}$. Then, for example, a global section of $\mathscr{L}$ is strictly small if and only if its pull-back to $\mathscr{L}_\sigma$ is strictly small for all σ.

Definition 8.1. If $d \ge 1$ then for all $\xi \in K^*$, we define

$$\deg \xi = \deg_{\mathscr{M}} \xi = \deg_{\mathscr{M}_F} (\xi)_\infty|_{B_F}, \tag{8.1.1}$$

where $\deg_{\mathscr{M}_F} (\xi)_\infty|_{B_F}$ is as in Definition 6.10. (In the latter, note that the intersection degree is taken relative to F.)

For all $d \ge 0$ we also let

$$\deg B = \deg_{\mathscr{M}} B = \mu(M_K^\infty). \tag{8.1.2}$$

For all $\sigma : F \hookrightarrow \mathbb{C}$, let $\mathscr{M}_\sigma$ denote the pull-back of $\mathscr{M}$ to B_σ, and for all $\xi \in K^*$ let ξ_σ denote the pull-back of ξ to an element of $\boldsymbol{\kappa}(B_\sigma)$. Then

$$\deg \xi = \deg_{\mathscr{M}_\sigma} \xi_\sigma \quad \text{for all } \sigma. \tag{8.2}$$

Also, μ_θ in 7.1 coincides with μ on $B_\sigma \subseteq B(\mathbb{C})$ for all σ. Therefore

$$\mu(B_\sigma) = \deg_{\mathscr{M}_\sigma} B_\sigma = c_1(\mathscr{M}_F)^{\cdot d} = \frac{\mu(M_K^\infty)}{[F:\mathbb{Q}]} = \frac{\deg B}{[F:\mathbb{Q}]} \tag{8.3}$$

by (3.4) and (8.1.2).

Next we show that $\deg \xi$ is bounded by a linear function of the height.

Lemma 8.4. *If $d \ge 1$ then*

$$\deg \xi \ll h_K(\xi) \tag{8.4.1}$$

for all $\xi \in K$, where the implicit constant depends only on K and the polarization.

Proof. For all $a \in \mathbb{R}$ let $\mathscr{V}_a$ be the line sheaf on B given by Definition 2.11. By Proposition 2.12, there is an $\epsilon > 0$ such that $\mathscr{N} := \mathscr{M} \otimes \mathscr{V}_{-\epsilon}$ is ample. Let h'_K denote the height on K defined using the polarization $M' := (B; \mathscr{N})$.

As noted below (3.6),

$$h'_K(\xi) \geq 0 \tag{8.4.2}$$

for all $\xi \in K$.

Since $c_1(\|\cdot\|_{\mathscr{V}_{-\epsilon}}) = 0$, the measure μ on M_K^∞ is the same for M' as for the polarization $M = (B, \mathscr{M})$. Now consider $Y \in B^{(1)}$. Since $c_1(\mathscr{V}_{-\epsilon}|_Y) \cdot c_1(\mathscr{V}_{-\epsilon}|_Y) = 0$ by (1.11.4), we have

$$\begin{aligned} h_{M'}(Y) - h_M(Y) &= d\, c_1(\mathscr{M}|_Y)^{\cdot(d-1)} \cdot c_1(\mathscr{V}_{-\epsilon}|_Y) \\ &= d\, c_1(\mathscr{M}|_Y)^{\cdot(d-1)} \cdot (0, -2\epsilon) \\ &= -\epsilon d \int_{Y_{\mathbb{C}}} c_1(\|\cdot\|_{\mathscr{M}})^{\wedge(d-1)} \end{aligned}$$

by (1.7). If Y is vertical then this is zero; otherwise it equals $-\epsilon d[F:\mathbb{Q}] \deg_{\mathscr{M}} Y$ by (6.12.1) and (8.2). By (3.6), (8.1.1), and (8.4.2), we then have

$$h_K(\xi) = h'_K(\xi) + \epsilon d[F:\mathbb{Q}] \deg \xi \gg \deg \xi. \qquad \square$$

Note that μ_θ coincides with μ on $B_\sigma \subseteq B(\mathbb{C})$. Therefore, by (3.6) and the product formula (3.5),

$$\sum_{\sigma : F \hookrightarrow \mathbb{C}} h_{B_\sigma}(\xi) \leq h_K(1/\xi) = h_K(\xi) \tag{8.5}$$

for all $\xi \in K^*$, where h_{B_σ} is as in Definition 7.2.

Also, we note that

$$h_K(\xi \pm \alpha) \leq h_K(\xi) + h_K(\alpha) + (\log 2) \deg B \tag{8.6}$$

for all $\xi, \alpha \in K$. Indeed, this follows from the elementary inequality

$$\max\{1, \|\xi \pm \alpha\|_v\} \leq \max\{1, \|\xi\|_v\} \cdot \max\{1, \|\alpha\|_v\} \cdot \begin{cases} 2 & \text{if } v \text{ is archimedean,} \\ 1 & \text{if } v \text{ is nonarchimedean,} \end{cases}$$

together with (3.6) and (8.1.2).

Finally, we note the closely related inequality

$$\|\alpha_1 + \cdots + \alpha_N\|_v \leq \max\{\|\alpha_1\|_v, \ldots, \|\alpha_N\|_v\} \cdot \begin{cases} N & \text{if } v \text{ is archimedean,} \\ 1 & \text{if } v \text{ is nonarchimedean} \end{cases} \tag{8.7}$$

for all $\alpha_1, \ldots, \alpha_N \in K$ and all $N \in \mathbb{Z}_{>0}$. This inequality is often used in diophantine geometry.

The following lemma adapts Proposition 7.3 to K and its polarization.

Lemma 8.8. *For each $\epsilon_8 > 0$ there is an $\epsilon_5 > 0$ such that*

$$\int_T -\log^- \|\xi\|_v \, d\mu(v) \leq \epsilon_8 h_K(\xi) \tag{8.8.1}$$

holds for all $\xi \in K^*$ and all measurable $T \subseteq M_K^\infty$ for which $\mu(T) \leq \epsilon_5$.

Proof. If $d = 0$ then M_K^∞ is a finite set and μ is the counting measure, so the result is trivial with $\epsilon_5 = \frac{1}{2}$.

Now assume that $d \ge 1$.

For each $\sigma\colon F \hookrightarrow \mathbb{C}$, Proposition 7.3, (8.2), and (8.3) imply that for each $\epsilon_4 > 0$ there is an $\epsilon_5 > 0$ such that

$$\int_{T\cap B_\sigma} -\log^- \|\xi\|_v \, d\mu(v) \le \frac{\epsilon_4}{2}\deg\xi + \frac{\mu(T\cap B_\sigma)}{\mu(B_\sigma)}\left(h_{B_\sigma}(\xi) + \frac{c_8}{2}\deg\xi\right) \tag{8.8.2}$$

holds for all $\xi \in K^*$ and all measurable $T \subseteq B(\mathbb{C})$ with $\mu(T \cap B_\sigma) \le \epsilon_5$.

Let c' be the implicit constant in (8.4.1). Choose $\epsilon_4 > 0$ and shrink ϵ_5 if necessary so that

$$\frac{c'\epsilon_4[F:\mathbb{Q}]}{2} + \frac{\epsilon_5[F:\mathbb{Q}]}{\deg B}\left(1 + \frac{c_8}{2}c'\right) \le \epsilon_8.$$

Summing (8.8.2) over all σ then gives (8.8.1), by (8.3) and (8.5). □

The following proposition gives a similar adaptation of Proposition 6.16.

Proposition 8.9. *For all $\epsilon_9 > 0$ and all $\epsilon_{10} > 0$ there is a cover of S by measurable subsets $C_1, \dots, C_\Lambda$, such that the following condition is true. For all $\xi \in K^*$ there is a measurable subset $T \subseteq M_K^\infty$ such that $\mu(T) \le \epsilon_9$, and such that*

$$\left| -\log^-\|\xi\|_v + \log^-\|\xi\|_{v'} \right| \le \epsilon_{10} h_K(\xi) \tag{8.9.1}$$

for all $v, v' \in C_l \setminus T$ and all $l = 1, \dots, \Lambda$.

Proof. If $d = 0$ then S is a finite set, and we can let $C_1, \dots, C_\Lambda$ be disjoint one-element sets whose union is S. Then the proposition holds trivially with $T = \varnothing$ for all ξ.

Therefore, assume from now on that $d \ge 1$.

Let c' be the implicit constant in (8.4.1). Applying Proposition 6.16 with $X = B_\sigma$ for all σ, and with $\epsilon_2 = \epsilon_9/[F:\mathbb{Q}]$ and $\epsilon_3 = \epsilon_{10}/c'$, gives a cover of $B(\mathbb{C})$ by measurable subsets $C_{0,1}, \dots, C_{0,\Lambda_0}$, such that for all $\xi \in K^*$ there is a measurable subset T_ξ of $B(\mathbb{C})$ such that $\mu(T_\xi) \le \epsilon_9$, and such that

$$\left| -\log^-|\xi(x)| + \log^-|\xi(x')| \right| \le \epsilon_{10} h_K(\xi)$$

for all $x, x' \in C_{0,l} \setminus T_\xi$ and all $l = 1, \dots, \Lambda_0$.

Let $C_l = C_{0,l} \cap S$ for all $l = 1, \dots, \Lambda_0$, and let $C_{\Lambda_0+1}, \dots, C_\Lambda$ be disjoint one-element sets whose union is $S \cap M_K^0$. Then $C_1, \dots, C_\Lambda$ are measurable subsets of S that cover S. Moreover (8.9.1) holds for each $\xi \in K^*$, with $T = T_\xi \cap S$. □

We are now ready to prove the main result of this section. The following lemma carries out the main pigeonhole argument. It is phrased in more general terms in order to use it in later work. Later in this section it will be applied with $\Xi \subseteq K$ and $\lambda_{\xi,j}(v) = -\log^-\|\xi - \alpha_j\|_v$.

Lemma 8.10. *Let Ξ be a set, let (S, Σ, μ) be a measure space of finite measure, let $h\colon \Xi \to [h_0, \infty)$ be an unbounded function with $h_0 > 0$, let $q > 0$ be an integer, let $\lambda_{\xi,1}, \dots, \lambda_{\xi,q}\colon S \to \mathbb{R}_{\ge 0}$ be measurable functions for all $\xi \in \Xi$, let $c_9 \in \mathbb{R}_{\ge 0}$, and let $\epsilon_{10} > 0$. Assume that these satisfy the following hypotheses:*

(i) *For all* $\xi \in \Xi$ *and all* $j = 1, \dots, q$,

$$\int_S \lambda_{\xi,j}\, d\mu \le h(\xi) + c_9. \tag{8.10.1}$$

(ii) *For all* $\epsilon_9 > 0$ *there is a cover of* S *by subsets* $C_1, \dots, C_\Lambda \in \Sigma$ *such that for each* $\xi \in \Xi$ *there is a set* $T_\xi \in \Sigma$ *with* $\mu(T_\xi) \le \epsilon_9$, *such that*

$$|\lambda_{\xi,j}(v) - \lambda_{\xi,j}(v')| \le \epsilon_{10}(h(\xi) + c_9)$$

for all $j = 1, \dots, q$, *all* $v, v' \in C_l \setminus T$, *and all* $l = 1, \dots, \Lambda$.

Then for every $\epsilon_{11} > 0$ *there is a subset* $\Xi' \subseteq \Xi$, *together with subsets* $T_\xi \in \Sigma$ *for all* $\xi \in \Xi'$, *such that* h *is unbounded on* Ξ', *such that* $\mu(T_\xi) \le \epsilon_{11}$ *for all* $\xi \in \Xi'$, *and such that*

$$\left| \frac{\lambda_{\eta,j}(v)}{h(\eta)} - \frac{\lambda_{\zeta,j}(v)}{h(\zeta)} \right| \le \left(4 + \frac{c_9}{h(\eta)} + \frac{c_9}{h(\zeta)}\right)\epsilon_{10} \tag{8.10.2}$$

for all $\eta, \zeta \in \Xi'$, *all* $v \in S \setminus (T_\eta \cup T_\zeta)$, *and all* $j = 1, \dots, q$.

Proof. First, we note that it suffices to prove the special case $q = 1$. Indeed, the general case follows from this case by applying the special case to each of the $\lambda_{\xi,j}$, with ϵ_{11} replaced by ϵ_{11}/q, successively shrinking the set Ξ' for each j.

We now show the special case $q = 1$. Let $\lambda_\xi = \lambda_{\xi,1}$ for all $\xi \in \Xi$.

Let $\epsilon_9 = \epsilon_{11}/2$. By (ii) there is a cover of S by subsets $C_1, \dots, C_\Lambda \in \Sigma$ such that for each $\xi \in \Xi$ there is a subset $T_\xi^0 \in \Sigma$ with $\mu(T_\xi^0) \le \epsilon_9$ such that

$$|\lambda_\xi(v) - \lambda_\xi(v')| \le \epsilon_{10}(h(\xi) + c_9) \tag{8.10.3}$$

for all $v, v' \in C_l \setminus T_\xi^0$ and all $l = 1, \dots, \Lambda$.

We may assume that $C_1, \dots, C_\Lambda$ are mutually disjoint.

For each $\xi \in \Xi$ and each $l = 1, \dots, \Lambda$ for which $\mu(C_l) > 0$, let

$$m_{\xi,l} = \inf\left\{ t \in \mathbb{R} : \mu(\{v \in C_l : \lambda_\xi(v) \ge t h(\xi)\}) \le \frac{\mu(C_l)}{2} \right\}.$$

(One can think of this as "a median value of $\lambda_\xi(v)/h(\xi)$ on C_l.") Note that, for all ξ and l, both sets

$$\{v \in C_l : \lambda_\xi(v) \le m_{\xi,l} h(\xi)\} \quad \text{and} \quad \{v \in C_l : \lambda_\xi(v) \ge m_{\xi,l} h(\xi)\} \tag{8.10.4}$$

have measure at least $\mu(C_l)/2$.

For the next step, we claim that there are constants $c_{10,l}$, $l = 1, \dots, \Lambda$, independent of ξ, such that $m_{\xi,l} \le c_{10,l}$ for all ξ and l that satisfy $\mu(C_l \cap T_\xi^0) < \mu(C_l)/2$. Indeed, for all such ξ and l, the statement about the second set in (8.10.4), together with (8.10.1), gives

$$\frac{\mu(C_l)}{2} m_{\xi,l} h(\xi) \le \int_S \lambda_\xi\, d\mu \le h(\xi) + c_9 \le \left(1 + \frac{c_9}{h_0}\right) h(\xi),$$

and the claim follows. (Note that the condition on ξ and l implies $\mu(C_l) > 0$.)

Next comes a pigeonhole argument.

For each $\xi \in \Xi$ let $\boldsymbol{m}_\xi$ be the vector in $\mathbb{R}^\Lambda$ whose l-th coordinate is

$$m'_{\xi,l} := \begin{cases} m_{\xi,l} & \text{if } \mu(C_l \cap T^0_\xi) < \mu(C_l)/2, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\boldsymbol{m}_\xi \in \prod_{l=1}^{\Lambda}[0, c_{10,l}]$ for all $\xi \in \Xi$. By a pigeonhole argument, there is a vector $\boldsymbol{m}^0 = (m^0_1, \dots, m^0_\Lambda) \in \mathbb{R}^\Lambda$ such that h is unbounded on the set Ξ' of all $\xi \in \Xi$ for which

$$\boldsymbol{m}_\xi \in \prod_{l=1}^{\Lambda}[m^0_l - \epsilon_{10}, m^0_l + \epsilon_{10}]. \tag{8.10.5}$$

For each $\xi \in \Xi'$ let T_ξ be the union of T^0_ξ and all C_l for which $\mu(C_l \cap T^0_\xi) \ge \mu(C_l)/2$. Then $\mu(T_\xi) \le 2\mu(T^0_\xi) \le \epsilon_{11}$.

It remains only to show that (8.10.2) holds.

To show this, let $\eta, \zeta \in \Xi'$ and let $v \in S \setminus (T_\eta \cup T_\zeta)$.

Let l be the (unique) index such that $v \in C_l$. By the definitions of T_η and T_ζ, we have $\mu(C_l \cap T^0_\eta) < \mu(C_l)/2$ and $\mu(C_l \cap T^0_\zeta) < \mu(C_l)/2$. Therefore, by (8.10.4), there are $v' \in C_l \setminus T_\eta$ and $v'' \in C_l \setminus T_\zeta$ such that $\lambda_\eta(v') \le m_{\eta,l} h(\eta)$ and $\lambda_\zeta(v'') \ge m_{\zeta,l} h(\zeta)$.

By (8.10.3), the choice of v', (8.10.5), the choice of v'', and (8.10.3) again, we then have

$$\begin{aligned} \frac{\lambda_\eta(v)}{h(\eta)} &\le \frac{\lambda_\eta(v')}{h(\eta)} + \left(1 + \frac{c_9}{h(\eta)}\right)\epsilon_{10} \\ &\le m_{\eta,l} + \left(1 + \frac{c_9}{h(\eta)}\right)\epsilon_{10} \\ &\le m_{\zeta,l} + \left(3 + \frac{c_9}{h(\eta)}\right)\epsilon_{10} \\ &\le \frac{\lambda_\zeta(v'')}{h(\zeta)} + \left(3 + \frac{c_9}{h(\eta)}\right)\epsilon_{10} \\ &\le \frac{\lambda_\zeta(v)}{h(\zeta)} + \left(4 + \frac{c_9}{h(\eta)} + \frac{c_9}{h(\zeta)}\right)\epsilon_{10}. \end{aligned}$$

A similar inequality holds with η and ζ interchanged, and this gives (8.10.2). □

The next step gives an upper bound for the "cost" of reducing to simultaneous approximation.

Lemma 8.11. *Let Ξ, (S, Σ, μ), h, h_0, $\lambda_{\xi,j}$ ($\xi \in \Xi$, $j = 1, \dots, q$), and c_9 be as in Lemma 8.10, and assume that the conclusion of Lemma 8.10 is true for all $\epsilon_{10} > 0$ (hypotheses (i) and (ii) are not assumed here). Assume also that the following hypothesis is satisfied:*

(iii) *For each $\epsilon_8 > 0$ there is an $\epsilon_5 > 0$ such that*

$$\int_T \lambda_{\xi,j}\, d\mu \le \epsilon_8(h(\xi) + c_9) \tag{8.11.1}$$

for all $j = 1, \dots, q$, all $\xi \in \Xi$, and all $T \in \Sigma$ with $\mu(T) \le \epsilon_5$.

For all $\xi \in \Xi$ define $\lambda_\xi : S \to \mathbb{R}$ by

$$\lambda_\xi(v) = \max\{\lambda_{\xi,1}(v), \dots, \lambda_{\xi,q}(v)\}, \qquad v \in S. \tag{8.11.2}$$

Then, for all $n \in \mathbb{Z}_{>0}$, all $\epsilon'' > 0$, all $c'' \in \mathbb{R}$, and all $r_{\min} \in [1, \infty)$, there exist $\xi_1, \dots, \xi_n \in \Xi$, a subset $T \in \Sigma$, and a measurable function $J : S \setminus T \to \{1, \dots, q\}$ such that

$$\frac{h(\xi_i)}{h(\xi_{i-1})} \ge r_{\min} \quad \textit{for all } i = 2, \dots, n \tag{8.11.3}$$

and

$$\int_S \frac{\lambda_{\xi_i}}{h(\xi_i)}\,d\mu - \int_{S\setminus T} \min_{1\le i'\le n} \frac{\lambda_{\xi_{i'},J(v)}(v)}{h(\xi_{i'})}\,d\mu(v) + \frac{c''}{h(\xi_1)} \le \epsilon'' \quad \textit{for all } i = 1, \dots, n. \tag{8.11.4}$$

Proof. Let n, ϵ'', c'', and $r_{\min}$ be given. We may assume that $c'' \ge 0$.

Choose $\epsilon_8 > 0$, $\epsilon_{10} > 0$, and $h_{\min} > 0$ such that

$$q\epsilon_8\left(1 + \frac{c_9}{h_{\min}}\right) + \left(4 + \frac{2c_9}{h_{\min}}\right)\epsilon_{10}\mu(S) + \frac{c''}{h_{\min}} \le \epsilon''. \tag{8.11.5}$$

Choose $\epsilon_5 > 0$ such that (8.11.1) holds, and let $\epsilon_{11} = \epsilon_5/n$.

Let $\Xi' \subseteq \Xi$ be as in the conclusion to Lemma 8.10. Choose $\xi_1 \in \Xi'$ with $h(\xi_1) \ge h_{\min}$ and choose $\xi_2, \dots, \xi_n \in \Xi'$ to satisfy (8.11.3). Let $T = T_{\xi_1} \cup \dots \cup T_{\xi_n}$; then $\mu(T) \le \epsilon_5$. Since $\lambda_\xi \le \lambda_{\xi,1} + \dots + \lambda_{\xi,q}$, by (8.11.1) we have

$$\int_T \lambda_\xi\,d\mu \le \sum_{j=1}^q \int_T \lambda_{\xi,j}\,d\mu \le q\epsilon_8(h(\xi) + c_9) \le q\epsilon_8\left(1 + \frac{c_9}{h_{\min}}\right)h(\xi) \tag{8.11.6}$$

for all $\xi \in \{\xi_1, \dots, \xi_n\}$.

Now let $v \in S \setminus T$. For conciseness and readability, let $\lambda_{ij} = \lambda_{\xi_i,j}(v)/h(\xi_i)$ for all $1 \le i \le n$ and all $1 \le j \le q$, and let $\epsilon_{12} = (4 + 2c_9/h(\xi_1))\epsilon_{10}$. Then, by (8.10.2),

$$|\lambda_{ij} - \lambda_{i'j}| \le \epsilon_{12} \tag{8.11.7}$$

for all $i, i' \in \{1, \dots, n\}$ and all $j \in \{1, \dots, q\}$. We then claim that there is a $j \in \{1, \dots, q\}$ such that

$$\max_{1\le j'\le q} \lambda_{ij'} - \min_{1\le i'\le n} \lambda_{i'j} \le \epsilon_{12} \quad \text{for all } i = 1, \dots, n. \tag{8.11.8}$$

Indeed, this is equivalent to the existence of j such that

$$\max_{1\le i\le n}\max_{1\le j'\le q} \lambda_{ij'} - \min_{1\le i'\le n} \lambda_{i'j} \le \epsilon_{12}. \tag{8.11.9}$$

The first term is equal to $\max_{1\le j'\le q}\max_{1\le i\le n} \lambda_{ij'}$; pick j such that this equals λ_{ij} for some i. If there are more than one such values of j, choose the smallest one (this ensures that $v \mapsto j$ is a measurable function). Then (8.11.9) reduces to

$$\max_{1\le i\le n} \lambda_{ij} - \min_{1\le i'\le n} \lambda_{i'j} \le \epsilon_{12},$$

and this follows from (8.11.7).

Let $J\colon S\setminus T\to\{1,\dots,q\}$ be the function defined by letting $J(v)$ be the above choice of j for each $v\in S\setminus T$. Then, by (8.11.6), (8.11.8), the choice of ξ_1, and (8.11.5),

$$\begin{aligned}\int_S \frac{\lambda_{\xi_i}}{h(\xi_i)}\,d\mu - \int_{S\setminus T}\min_{1\le i'\le n}\frac{\lambda_{\xi_{i'},J(v)}(v)}{h(\xi_{i'})}\,d\mu(v)+\frac{c''}{h(\xi_1)}&\\ = \int_T \frac{\lambda_{\xi_i}}{h(\xi_i)}\,d\mu + \int_{S\setminus T}\left(\frac{\lambda_{\xi_i}(v)}{h(\xi_i)} - \min_{1\le i'\le n}\frac{\lambda_{\xi_{i'},J(v)}(v)}{h(\xi_{i'})}\right)d\mu(v)+\frac{c''}{h(\xi_1)}&\\ \le q\epsilon_8\left(1+\frac{c_9}{h_{\min}}\right)+\left(4+\frac{2c_9}{h(\xi_1)}\right)\epsilon_{10}\mu(S\setminus T)+\frac{c''}{h(\xi_1)}&\\ \le \epsilon''&\end{aligned}$$

for all $i=1,\dots,n$. This implies (8.11.4). □

The main result of this section now follows easily from Lemmas 8.10 and 8.11.

Proposition 8.12. *Let $\alpha_1,\dots,\alpha_q$ be distinct elements of K, let $\epsilon>0$, and let $c\in\mathbb{R}$. Assume that Theorem 4.5 is false for these values. Let n be a positive integer, let $\epsilon'\in(0,\epsilon)$, let $c'\in\mathbb{R}$, and let $r_{\min}\in[1,\infty)$. Then there exist $\xi_1,\dots,\xi_n\in K$ and mutually disjoint measurable subsets $T_1,\dots,T_q$ of S such that*

$$\frac{h_K(\xi_i)}{h_K(\xi_{i-1})}\ge r_{\min}\quad \text{for all } i=2,\dots,n \tag{8.12.1}$$

and

$$\sum_{j=1}^{q}\int_{T_j}\min_{1\le i\le n}\frac{-\log^-\|\xi_i-\alpha_j\|_v}{h_K(\xi_i)}\,d\mu(v)\ge 2+\epsilon'+\frac{c'}{h_K(\xi_1)}. \tag{8.12.2}$$

Proof. By the assumption that Theorem 4.5 is false, there is an infinite subset

$$\Xi\subseteq K\setminus\{\alpha_1,\dots,\alpha_q\}$$

such that (4.5.1) is false for all $\xi\in\Xi$, using the above choices of $\alpha_1,\dots,\alpha_q$, ϵ, and c. By Northcott's theorem (Theorem 3.16) we may assume that there is some $h_0>0$ such that $h_K(\xi)\ge h_0$ for all $\xi\in\Xi$. Also h_K is unbounded on this set.

We will apply Lemmas 8.10 and 8.11 to this choice of Ξ, with $h=h_K$, with $\lambda_{\xi,j}\colon S\to\mathbb{R}$ given by

$$\lambda_{\xi,j}(v)=-\log^-\|\xi-\alpha_j\|_v,$$

and with

$$c_9=\max_{1\le j\le q}h_K(\alpha_j)+(\log 2)\deg B. \tag{8.12.3}$$

Note that, defining $\lambda_\xi=\max\{\lambda_{\xi,j}:j=1,\dots,q\}$ as in (8.11.2), by the definition of Ξ we have

$$\int_S\frac{\lambda_\xi}{h_K(\xi)}\,d\mu>2+\epsilon+\frac{c}{h_K(\xi)}\quad \text{for all } \xi\in\Xi. \tag{8.12.4}$$

Assumption (i) of Lemma 8.10 holds, since

$$\int_S -\log^- \|\xi - \alpha_j\|_v \, d\mu(v) \le h_K\left(\frac{1}{\xi - \alpha_j}\right) = h_K(\xi - \alpha_j) \le h_K(\xi) + c_9$$

by (3.6), the product formula (3.5), (8.6), and (8.12.3). Assumption (ii) holds by Proposition 8.9, and assumption (iii) of Lemma 8.11 holds by Lemma 8.8.

Therefore Lemma 8.11 applies, and there exist $\xi_1, \dots, \xi_n \in \Xi$, a measurable subset $T \subseteq S$, and a measurable function $J\colon S \setminus T \to \{1, \dots, q\}$ such that (8.11.4) holds with $\epsilon'' = \epsilon - \epsilon'$ and $c'' = c' + \max\{-c, 0\}$, and such that (8.12.1) holds.

Subtracting (8.11.4) from (8.12.4) with $\xi = \xi_i$ then gives

$$\begin{aligned}\int_{S\setminus T} \min_{1\le i'\le n} \frac{\lambda_{\xi_{i'}, J(v)}(v)}{h(\xi_{i'})}\, d\mu(v) &> 2 + \epsilon - \epsilon'' + \frac{c}{h(\xi_i)} + \frac{c''}{h(\xi_1)} \quad \text{for all } i \\ &= 2 + \epsilon' + \frac{c'}{h(\xi_1)} + \frac{c}{h(\xi_i)} + \frac{\max\{-c, 0\}}{h(\xi_1)} \quad \text{for all } i \\ &\ge 2 + \epsilon' + \frac{c'}{h(\xi_1)}.\end{aligned}$$

Upon letting $T_j = J^{-1}(j)$ for all j, this gives (8.12.2). □

9. Siegel's lemma and the auxiliary polynomial

Since $\mathscr{M}$ is ample, work of X. Yuan and (independently) H. Chen allows one to control the number of small global sections of $\mathscr{M}^{\otimes m}$ as $m \to \infty$, providing a counterpart to Axioms 1a and 1b of [Lang 1983, Chapter 7, Section 1].

In more detail, by [Yuan 2009, Section 1.1 and Theorem 2.7] (see also [Chen 2008]), we have

$$\lim_{m\to\infty} \frac{h^0(B, \mathscr{M}^{\otimes m})}{m^{d+1}/(d+1)!} = c_1(\mathscr{M})^{\cdot(d+1)} > 0,$$

since $\mathscr{M}$ is ample by assumption 5.3. (Recall Definition 2.3(a) and that $\dim B = d + 1$.)

Therefore there are constants c_{11} and c_{12}, with $c_{12} > c_{11} > 0$, and an integer m_0 (depending on c_{11}, c_{12}, and $\mathscr{M}$), such that the inequality

$$c_{11} m^{d+1} \le h^0(B, \mathscr{M}^{\otimes m}) \le c_{12} m^{d+1} \tag{9.1}$$

holds for all integers $m \ge m_0$. (Also, c_{11} and c_{12} can be taken so that c_{12}/c_{11} is arbitrarily close to 1, although this fact will not be used here.)

This estimate is sufficient for proving the following Siegel lemma for arithmetic function fields.

Theorem 9.2. *Let c_{11}, c_{12}, and m_0 be as in (9.1). Let h and b be positive integers, and let A be an $M \times N$ matrix with entries in $H^0(B, \mathscr{M}^{\otimes h} \otimes \mathscr{V}_{-\log N})$. Assume that $b \ge m_0$ and that*

$$\left(1 + \frac{h}{b}\right)^{d+1} < \frac{N c_{11}}{M c_{12}}. \tag{9.2.1}$$

Then there is a nonzero vector $\boldsymbol{v} \in H^0(B, \mathscr{M}^{\otimes b} \otimes \mathscr{V}_{\log 2})^N$ such that $A\boldsymbol{v} = \boldsymbol{0}$ (in $H^0(B, \mathscr{M}_{\mathrm{fin}}^{\otimes(h+b)})^M$).

Proof. By (9.1),

$$\log \# H^0(B, \mathscr{M}^{\otimes b})^N \geq N c_{11} b^{d+1}.$$

On the other hand, if $\boldsymbol{v} \in H^0(B, \mathscr{M}^{\otimes b})^N$ then $A\boldsymbol{v}$ lies in $H^0(B, \mathscr{M}^{\otimes(b+h)})^M$ by (8.7), and

$$\log \# H^0(B, \mathscr{M}^{\otimes(b+h)})^M \leq M c_{12}(b+h)^{d+1}.$$

Therefore, by (9.2.1),

$$H^0(B, \mathscr{M}^{\otimes b})^N > H^0(B, \mathscr{M}^{\otimes(b+h)})^M,$$

so there are distinct vectors $\boldsymbol{v}_1, \boldsymbol{v}_2 \in H^0(B, \mathscr{M}^{\otimes b})^N$ such that $A\boldsymbol{v}_1 = A\boldsymbol{v}_2$. Let $\boldsymbol{v} = \boldsymbol{v}_1 - \boldsymbol{v}_2$. Then $\boldsymbol{v} \neq \boldsymbol{0}$, $A\boldsymbol{v} = \boldsymbol{0}$, and

$$\boldsymbol{v} \in H^0(B, \mathscr{M}^{\otimes b} \otimes \mathscr{V}_{\log 2})^N$$

by (8.7), as was to be shown. □

Now we recall the index of a polynomial in $K[x_1, \ldots, x_n]$.

Definition 9.3. Let $n, d_1, \ldots, d_n \in \mathbb{Z}_{>0}$, let $P \in K[x_1, \ldots, x_n]$ be a nonzero polynomial, and let $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$ be a point in K^n. Write P as a polynomial in $x_1 - \xi_1, \ldots, x_n - \xi_n$:

$$P(x_1, \ldots, x_n) = \sum_{\boldsymbol{k} \in \mathbb{N}^n} a_{\boldsymbol{k}} (x_1 - \xi_1)^{k_1} \cdots (x_n - \xi_n)^{k_n}, \tag{9.3.1}$$

with $a_{\boldsymbol{k}} \in K$ for all $\boldsymbol{k}$, where $\boldsymbol{k} = (k_1, \ldots, k_n)$. Then the *index* of P at $\boldsymbol{\xi}$ with respect to $\boldsymbol{d} = (d_1, \ldots, d_n)$ is the number

$$t_{\boldsymbol{d}}(P, \boldsymbol{\xi}) = \min\left\{\frac{k_1}{d_1} + \cdots + \frac{k_n}{d_n} : a_{\boldsymbol{k}} \neq 0\right\}.$$

Following [Lang 1983, Chapter 7, Section 3], we may express the definition of index using (repeated) divided partial derivatives of P, as follows. The expansion (9.3.1) is just the Taylor expansion of P at $\boldsymbol{\xi}$, so $a_{\boldsymbol{k}} = \partial_{\boldsymbol{k}} P(\boldsymbol{\xi})$, where

$$\partial_{\boldsymbol{k}} = \frac{1}{k_1! \cdots k_n!} \left(\frac{\partial}{\partial x_1}\right)^{k_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{k_n}.$$

In particular, the coefficients of $\partial_{\boldsymbol{k}} P$ are integral multiples of the coefficients of P.

Assume from now on that for all i the degree of P with respect to x_i is at most d_i. Then the integral factors in question are nonnegative and bounded by

$$\binom{d_1}{k_1} \cdots \binom{d_n}{k_n} \leq 2^{d_1 + \cdots + d_n}. \tag{9.4}$$

For any $\tau \in \mathbb{R}$, a polynomial P as above has index $\ge \tau$ at $\boldsymbol{\xi}$ if and only if $\partial_{\boldsymbol{k}} P(\boldsymbol{\xi}) = 0$ for all $\boldsymbol{k} = (k_1, \dots, k_n)$ such that $k_i \le d_i$ for all i and $\sum_i k_i/d_i < \tau$. Let $J_{\boldsymbol{d}}(\tau)$ denote the number of such conditions.

Following [Esnault and Viehweg 1984, Section 9], let $\operatorname{Vol}_n(\tau)$ be the Lebesgue measure of the set $\{\boldsymbol{x} = (x_1, \dots, x_n) \in [0, 1]^n : x_1 + \dots + x_n < \tau\}$. Then

$$\frac{J_{\boldsymbol{d}}(\tau)}{d_1 \cdots d_n} = \operatorname{Vol}_n(\tau) + O\biggl(\frac{1}{d_1} + \dots + \frac{1}{d_n}\biggr),$$

where the implicit constant depends only on n.

Now we introduce (as is typical of proofs of Roth's theorem) an auxiliary polynomial. The degree of this polynomial will be taken large, depending on a real number D which in the end will be taken sufficiently large depending on everything else in the proof (except for things defined using D).

Proposition 9.5. *Let n be a positive integer, and let τ be a positive real number such that $\operatorname{Vol}_n(\tau) < 1/q$. Let c_{11} and c_{12} be as in (9.1), let β be a positive real number such that*

$$\biggl(1 + \frac{1}{\beta}\biggr)^{d+1} < \frac{c_{11}}{q c_{12} \operatorname{Vol}_n(\tau)}, \tag{9.5.1}$$

and let $h_1, \dots, h_n$ be positive real numbers. Then there exist a positive integer u, depending only on $\alpha_1, \dots, \alpha_q$ and $\mathscr{M}$, and a real number $D_0 > 0$, depending on all of the foregoing, such that the following is true for all $D \ge D_0$. Let $d_i = \lfloor D/h_i \rfloor$ for all i, let $\boldsymbol{d} = (d_1, \dots, d_n)$, let $h = u(d_1 + \dots + d_n)$, and let $b = \lfloor \beta h \rfloor$. Then there is a nonzero polynomial

$$P \in H^0(B, \mathscr{M}^{\otimes b} \otimes \mathscr{V}_{\log 2})[x_1, \dots, x_n], \tag{9.5.2}$$

of degree at most d_i in x_i for all i, such that

$$t_{\boldsymbol{d}}(P, (\alpha_j, \dots, \alpha_j)) \ge \tau \quad \textit{for all } j = 1 \dots, q. \tag{9.5.3}$$

Proof. Let E be an effective divisor on $B_{\mathbb{Q}}$ such that $E + (\alpha_j)$ is effective for all $j = 1, \dots, q$. Let $\mathscr{F}$ be a smoothly metrized line sheaf on B such that $\mathscr{F}_{\mathbb{Q}} \cong \mathscr{O}(E)$, and such that the canonical section $\mathbf{1}_E$ of $\mathscr{O}(E)$ satisfies

$$\|\mathbf{1}_E\|_v \le \tfrac{1}{2} \quad \text{and} \quad |\alpha_j|_v \|\mathbf{1}_E\|_v \le \tfrac{1}{2}$$

for all $v \in M_K^\infty$ and all j. For all j let $s_j = \alpha_v \mathbf{1}_E$, so that s_j is a global section of $\mathscr{F}$ and $\|s_j\|_v \le \frac{1}{2}$ for all $v \in M_K^\infty$. By [Moriwaki 2014, Proposition 5.43], there is a positive integer u such that $\mathscr{F}^\vee \otimes \mathscr{M}^{\otimes u}$ has a nonzero strictly small global section ρ.

Now let $h_1, \dots, h_n$, D, $d_1, \dots, d_n$, $\boldsymbol{d}$, h, and b be as in the statement of the proposition. We aim to use Theorem 9.2 (Siegel's lemma) to construct P, by letting the coefficients of P be the unknowns in the linear algebra problem and using the conditions $\partial_{\boldsymbol{k}} P(\alpha_j, \dots, \alpha_j) = 0$ ($j = 1, \dots, q$) as the equations.

Let N be the number of terms in P and M be the number of constraints (as mentioned above). Then

$$N = \prod_{i=1}^{n}(d_i + 1) \quad \text{and} \quad M = q\,J_{\boldsymbol{d}}(\tau).$$

Since $N/\prod d_i = 1 + O\big(\sum d_i^{-1}\big)$ and $M/\prod d_i = q\,\mathrm{Vol}_n(\tau) + O\big(\sum d_i^{-1}\big)$, (9.2.1) follows from (9.5.1) for all sufficiently large D.

For all $\boldsymbol{k}$ and all j, $\partial_{\boldsymbol{k}} P(\alpha_j, \dots, \alpha_j)$ is a homogeneous linear form in the coefficients of P. The coefficients of this linear form are elements of K of the form

$$\binom{d_1}{k_1} \cdots \binom{d_n}{k_n} \alpha_j^{l_1+\cdots+l_n}$$

with $0 \le k_i \le d_i$ and $0 \le l_i \le d_i$ for all $i = 1, \dots, n$. By (9.4), multiplying these latter coefficients by $\mathbf{1}_E^{d_1+\cdots+d_n}$ gives small global sections of $\mathscr{F}^{\otimes(d_1+\cdots+d_n)}$; tensoring these with $\rho^{d_1+\cdots+d_n}$ then gives (strictly) small sections of $\mathscr{M}^{\otimes h}$. Since ρ is a strictly small section and since $\log N = o(D) = o(h)$, all of these products lie in $H^0(B, \mathscr{M}^{\otimes h} \otimes \mathscr{V}_{-\log N})$ for sufficiently large D.

Finally, since b grows roughly linearly in D, we have $b \ge m_0$ for all sufficiently large D. Therefore Theorem 9.2 applies, and this gives the polynomial P that satisfies (9.5.2) and (9.5.3). □

10. Conclusion of the proof

The remainder of the proof of Roth's theorem consists of choosing $\xi_1, \dots, \xi_n \in K$, constructing an auxiliary polynomial P, finding a lower bound for the index of P at $(\xi_1, \dots, \xi_n)$, and finally putting everything together to produce a contradiction. The last step in obtaining a contradiction usually involves Roth's lemma. Although Roth's lemma is almost certainly true over arithmetic function fields, here it is more expedient to use Dyson's lemma [Esnault and Viehweg 1984], simply because it is already proved over all fields of characteristic zero.

We start by finding a lower bound for the index. Since this involves a polynomial whose coefficients are global sections of a line sheaf, it involves metrics on that line sheaf at all places, including nonarchimedean places.

Definition 10.1. Let $\mathscr{L}$ be a smoothly metrized line sheaf on B, let s be a nonzero rational section of $\mathscr{L}$, let v be a nonarchimedean place of K, and let Y be the prime divisor on B corresponding to v. Let n_Y be the multiplicity of Y in $\operatorname{div}(s)$. Then we define $\|s\|_v = \exp(-n_Y h_M(Y))$. (Recall also that $\|s\|_v$ at an archimedean place v is defined using the metric of $\mathscr{L}$.)

The following lemma may be regarded as an extension of the product formula (which is the case $d = 0$ here).

Lemma 10.2. *Let $b \in \mathbb{Z}$ and let s be a nonzero rational section of $\mathscr{M}^{\otimes b}$ on B. Then*

$$\int_{M_K} -\log\|s\|_v \, d\mu(v) = b c_1(\mathscr{M})^{\cdot(d+1)}. \tag{10.2.1}$$

Proof. Write $\operatorname{div}(s)_{\text{fin}} = \sum n_Y \cdot Y$. Then, by Lemma 1.11,

$$\begin{aligned} bc_1(\mathscr{M})^{\cdot(d+1)} &= \sum_Y n_Y c_1(\mathscr{M}|_Y)^{\cdot d} - \int_{B(\mathbb{C})} \log\|s\|\, c_1(\|\cdot\|_{\mathscr{M}})^{\wedge d} \\ &= \sum_Y n_Y h_M(Y) - \int_{M_K^\infty} \log\|s\|_v \, d\mu(v) \\ &= \sum_{v \in M_K^0} -\log\|s\|_v - \int_{M_K^\infty} \log\|s\|_v \, d\mu(v) \\ &= \int_{M_K} -\log\|s\|_v \, d\mu(v) \end{aligned}$$

□

We are now ready to prove a lower bound for the index of the polynomial P constructed in Proposition 9.5 at a point $\boldsymbol{\xi}$ satisfying certain conditions. This will involve using the approximation condition (8.12.2) to obtain bounds on $\|P(\boldsymbol{\xi})\|_v$ for all v.

Proposition 10.3. *Let n, τ, u, and D_0 be as in Proposition 9.5. Let $\sigma > 0$ be a real number such that*

$$(2+\epsilon')(\tau - \sigma) > n, \tag{10.3.1}$$

let $r_{\min} \ge 1$ be a real number, and let

$$c' = \frac{n}{\tau - \sigma}\biggl((\log 12) \deg B + \sum_{j=1}^{q} h_K(\alpha_j)\biggr). \tag{10.3.2}$$

Let $\xi_1, \dots, \xi_n$ be elements of K that satisfy (8.12.1) and (8.12.2) for some $\epsilon' > 0$ and some $T_1, \dots, T_q \subseteq S$. Let $h_i = h_K(\xi_i)$ for all $i = 1, \dots, n$. Let $\boldsymbol{d} = (d_1, \dots, d_n)$, b, and P be as in Proposition 9.5. Then the polynomial P also satisfies

$$t_{\boldsymbol{d}}(P, (\xi_1, \dots, \xi_n)) \ge \sigma. \tag{10.3.3}$$

Proof. This proof is adapted from the argument at the end of [Lang 1983, Chapter 7, Section 3].

It will suffice to show that $\partial_{\boldsymbol{k}} P(\xi_1, \dots, \xi_n) = 0$ for all $\boldsymbol{k} = (k_1, \dots, k_n) \in \mathbb{N}^n$ satisfying

$$\frac{k_1}{d_1} + \dots + \frac{k_n}{d_n} < \sigma.$$

Assume by way of contradiction that $\boldsymbol{k}$ is an n-tuple that satisfies the above inequality, but that

$$\partial_{\boldsymbol{k}} P(\xi_1, \dots, \xi_n) \neq 0. \tag{10.3.4}$$

To avoid cluttered notation, let $Q = \partial_{\boldsymbol{k}} P$.

Note that

$$t_{\boldsymbol{d}}(Q, (\alpha_j, \dots, \alpha_j)) \ge \tau - \sigma$$

for all $j = 1, \dots, q$, and therefore $\partial_{\boldsymbol{\ell}} Q(\alpha_j, \dots, \alpha_j) \neq 0$ only if

$$\frac{\ell_1}{d_1} + \cdots + \frac{\ell_n}{d_n} \geq \tau - \sigma. \tag{10.3.5}$$

We start by estimating $\|Q(\xi_1, \dots, \xi_n)\|_v$ for all $v \in M_K$. Here we will think of the coefficients of P as being global sections of $\mathscr{M}^{\otimes b}$, having norms ≤ 2 at all archimedean places (and hence not necessarily small sections). Coefficients of Q will then also be global sections of $\mathscr{M}^{\otimes b}$, and values of Q such as $Q(\xi_1, \dots, \xi_n)$ will be rational sections of $\mathscr{M}^{\otimes b}$.

Let $T_1, \dots, T_q$ be as in Proposition 8.12. By shrinking $T_1, \dots, T_q$ if necessary, we may assume that

$$\min_{1 \leq i \leq n} \frac{-\log^- \|\xi_i - \alpha_j\|_v}{h_K(\xi_i)} > 0 \tag{10.3.6}$$

for all $v \in T_j$, $j = 1, \dots, q$. This does not affect (8.12.2).

First, let v be an archimedean place of K such that $v \in T_j$ for some j.

We consider the Taylor expansion

$$Q(x_1, \dots, x_n) = \sum_{\boldsymbol{\ell} \in \Lambda} \partial_{\boldsymbol{\ell}} Q(\alpha_j, \dots, \alpha_j)(x_1 - \alpha_j)^{\ell_1} \cdots (x_n - \alpha_j)^{\ell_n}, \tag{10.3.7}$$

where Λ is the set of all n-tuples $\boldsymbol{\ell} = (\ell_1, \dots, \ell_n) \in \mathbb{N}^n$ with $\ell_i \leq d_i$ for all i and satisfying (10.3.5). Note that $\partial_{\boldsymbol{\ell}} Q = \partial_{\boldsymbol{\ell}} \partial_{\boldsymbol{k}} P$, that the operator $\partial_{\boldsymbol{\ell}} \partial_{\boldsymbol{k}}$ takes $x_i^{m_i}$ to $\binom{m_i - k_i}{\ell_i}\binom{m_i}{k_i} x_i^{m_i - k_i - \ell_i}$, and that $\binom{m_i - k_i}{\ell_i}\binom{m_i}{k_i}$ is a trinomial coefficient with $m_i \leq d_i$. Hence $\partial_{\boldsymbol{\ell}} Q(\alpha_j, \dots, \alpha_j)$ can be written as a sum of at most $2^{d_1 + \cdots + d_n}$ terms, each with an additional factor of at most $3^{d_1 + \cdots + d_n}$ coming from $\partial_{\boldsymbol{\ell}} \partial_{\boldsymbol{k}}$, so we have

$$\|\partial_{\boldsymbol{\ell}} Q(\alpha_j, \dots, \alpha_j)\|_v \leq 2 \cdot 2^{d_1 + \cdots + d_n} \cdot 3^{d_1 + \cdots + d_n} \cdot \max\{1, \|\alpha_j\|_v\}^{d_1 + \cdots + d_n}.$$

The Taylor expansion (10.3.7) then gives the bound

$$\begin{aligned} -\log \frac{\|Q(\xi_1, \dots, \xi_n)\|_v}{2(12 \max\{1, \|\alpha_j\|_v\})^{d_1 + \cdots + d_n}} &\geq -\log \max_{\boldsymbol{\ell} \in \Lambda} \|(\xi_1 - \alpha_1)^{\ell_1} \cdots (\xi_n - \alpha_j)^{\ell_n}\|_v \\ &\geq (D + o(D)) \min_{\boldsymbol{\ell} \in \Lambda} \left(\sum_{i=1}^n \frac{\ell_i}{d_i} \cdot \frac{-\log^- \|\xi_i - \alpha_j\|_v}{h_K(\xi_i)} \right) \\ &\geq (D + o(D))(\tau - \sigma) \min_{1 \leq i \leq n} \frac{-\log^- \|\xi_i - \alpha_j\|_v}{h_K(\xi_i)}. \end{aligned} \tag{10.3.8}$$

Here we use the bound $|\Lambda| \leq 2^{d_1 + \cdots + d_n}$ in the first step (this changes 6 to 12 in the left-hand side) and (10.3.5) in the last step. Also, the limiting behavior of $o(D)$ can be taken independent of $\boldsymbol{k}$.

For nonarchimedean $v \in T_j$ satisfying (10.3.6), a similar argument gives

$$-\log \frac{\|Q(\xi_1, \dots, \xi_n)\|_v}{(\max\{1, \|\alpha_j\|_v\})^{d_1 + \cdots + d_n}} \geq (D + o(D))(\tau - \sigma) \min_{1 \leq i \leq n} \frac{-\log^- \|\xi_i - \alpha_j\|_v}{h_K(\xi_i)}. \tag{10.3.9}$$

Next consider archimedean v with $v \notin T_1 \cup \cdots \cup T_q$.

By bounds on binomial coefficients arising from applying $\partial_{\boldsymbol{k}}$ to P, the norms of coefficients of Q at archimedean places are bounded by $2^{1+d_1+\cdots+d_n}$. Since Q has at most $2^{d_1+\cdots+d_n}$ terms, we have

$$\|Q(\xi_1,\ldots,\xi_n)\|_v \le 2\cdot 4^{d_1+\cdots+d_n}\cdot \max\{1,\|\xi_1\|_v\}^{d_1}\cdots\max\{1,\|\xi_n\|_v\}^{d_n}. \tag{10.3.10}$$

Finally, for nonarchimedean $v\notin T_1\cup\cdots\cup T_q$, we have simply

$$\|Q(\xi_1,\ldots,\xi_n)\|_v \le \max\{1,\|\xi_1\|_v\}^{d_1}\cdots\max\{1,\|\xi_n\|_v\}^{d_n}. \tag{10.3.11}$$

Combining (10.3.8)–(10.3.11) and (8.12.2), we then have

$$\int_{M_K} -\log\|Q(\xi_1,\ldots,\xi_n)\|_v\,d\mu(v) \ge -\int_{M_K^\infty}\Bigl(\log 2+(\log 12)\sum d_i\Bigr)\,d\mu(v) - \Bigl(\sum h_K(\alpha_j)\Bigr)\sum d_i \\ + (D+o(D))(\tau-\sigma)\Bigl(2+\epsilon'+\frac{c'}{h_K(\xi_1)}\Bigr) - \sum d_i h_K(\xi_i).$$

By (10.3.4) and Lemma 10.2, the left-hand side equals $bc_1(\mathscr{M})^{\cdot(d+1)}$. By (8.1.2) and (10.3.2), this then becomes

$$bc_1(\mathscr{M})^{\cdot(d+1)} + \frac{(\tau-\sigma)c'}{n}\sum d_i + (\log 2)\deg B \ge (D+o(D))(\tau-\sigma)\Bigl(2+\epsilon'+\frac{c'}{h_K(\xi_1)}\Bigr) - \sum d_i h_K(\xi_i).$$

By definition of d_i, we have $d_i h_K(\xi_i) = D+o(D)$ for all i. Furthermore, (8.12.1) and the assumption that $r_{\min}\ge 1$ imply that $h_K(\xi_i)\ge h_K(\xi_1)$ for all i; hence $\sum d_i \le n(D+o(D))/h_K(\xi_1)$. Therefore we have

$$bc_1(\mathscr{M})^{\cdot(d+1)} + (\log 2)\deg B \ge (D+o(D))((\tau-\sigma)(2+\epsilon')-n).$$

By (10.3.1) this gives a contradiction for large enough D; hence (10.3.3) is true. □

The next (and next to last) step in the proof is to choose the main parameters of the proof. It is based on [Esnault and Viehweg 1984, Lemma 9.7].

Lemma 10.4. *Let $q\ge 2$ be an integer and let $\epsilon'>0$ be given. Then there is an integer $n_0=n_0(q,\epsilon')\ge 2$ such that for all $n\ge n_0$ there are real numbers τ and σ such that*

$$q\,\mathrm{Vol}_n(\tau) < 1 < q\,\mathrm{Vol}_n(\tau)+\mathrm{Vol}_n(\sigma) \tag{10.4.1}$$

and

$$(2+\epsilon')(\tau-\sigma) > n. \tag{10.4.2}$$

Proof. We will show that the lemma holds with $\sigma=1$ and with τ chosen such that

$$q\,\mathrm{Vol}_n(\tau) = 1-\frac{1}{2\cdot n!}. \tag{10.4.3}$$

For each n there is such a τ, and since $\mathrm{Vol}_n(1)=1/n!$, these choices satisfy (10.4.1).

Consider the inequality

$$\sqrt{\frac{\log q-\log(1-1/(2\cdot n!))}{6n}} + \frac{1}{n} < \frac12 - \frac{1}{2+\epsilon'}. \tag{10.4.4}$$

Its left-hand side tends to zero as $n \to \infty$, so there is an integer $n_0 \geq 2$ such that this inequality holds for all $n \geq n_0$. It remains only to check that (10.4.2) holds for these values of n, τ, and σ.

Bombieri and Gubler [2006, Lemma 6.3.5] showed that

$$\mathrm{Vol}_n\left(\left(\frac{1}{2} - \eta\right)n\right) \leq e^{-6n\eta^2}$$

for all $\eta \geq 0$. If η satisfies $\left(\frac{1}{2} - \eta\right)n = \tau$ and τ satisfies (10.4.3), then

$$\eta^2 \leq \frac{\log q - \log(1 - 1/(2 \cdot n!))}{6n},$$

and therefore by (10.4.4)

$$\frac{1}{2} - \eta - \frac{1}{n} > \frac{1}{2 + \epsilon'}.$$

The left-hand side equals $(\tau - \sigma)/n$, so (10.4.2) is true. □

We now introduce Dyson's lemma, as extended by Esnault and Viehweg.

Theorem 10.5 [Esnault and Viehweg 1984, Theorem 0.4]. *Let K be a field of characteristic zero. Let $\boldsymbol{\zeta}_j = (\zeta_{j,1}, \dots, \zeta_{j,n})$, $j = 1, \dots, M$, be points in K^n; let $\boldsymbol{d} \in \mathbb{Z}^n$ with $d_1 \geq d_2 \geq \cdots \geq d_n > 0$; let $t_1, \dots, t_M \in [0, \infty)$; and let $P \in K[x_1, \dots, x_n]$. Assume that*

(i) *$\zeta_{j,i} \neq \zeta_{j',i}$ for all $j \neq j'$ and $i = 1, \dots, n$;*

(ii) *P has degree at most d_i in x_i for all i;*

(iii) *$t_{\boldsymbol{d}}(P, \boldsymbol{\zeta}_j) = t_j$ for all $j = 1, \dots, M$.*

Then

$$\sum_{j=1}^{M} \mathrm{Vol}_n(t_j) \leq \prod_{i=1}^{n} \left(1 + \max\{M-2, 0\} \sum_{l=i+i}^{n} \frac{d_l}{d_i}\right). \tag{10.5.1}$$

Remark 10.6. Esnault and Viehweg [1984] stated this theorem only with $K = \mathbb{C}$, but it is true over arbitrary fields of characteristic zero (as above) by the Lefschetz principle in algebraic geometry, or (in the present situation) just by embedding K into $\mathbb{C}$.

More generally, let B be an integral scheme whose function field K has characteristic zero, and let $\mathscr{L}$ be a line sheaf on B. Then Dyson's lemma also holds for polynomials with coefficients in $H^0(B, \mathscr{L})$. Indeed, one can tensor all coefficients with a fixed nonzero element of the stalk of $\mathscr{L}^\vee$ at the generic point of B. The resulting polynomial will have coefficients in K, and it will have the same degree and index properties, so (10.5.1) will then apply to the original polynomial.

Proof of Theorem 4.5. The proof will be by contradiction. Let K, M_K, S, $\alpha_1, \dots, \alpha_q$, $\epsilon > 0$; and $c \in \mathbb{R}$ be as in the statement of Theorem 4.5, and assume that (4.5.1) fails to hold for infinitely many $\xi \in K$. Pick $\epsilon' \in (0, \epsilon)$, and choose n, τ, and σ such that (10.4.1) and (10.4.2) hold.

We shall apply Dyson's lemma with $M = q + 1$, $\boldsymbol{\zeta}_j = (\alpha_j, \dots, \alpha_j)$ $(j = 1, \dots, q)$, and $\boldsymbol{\zeta}_M = \boldsymbol{\xi}$, with $\boldsymbol{\xi}$ yet to be determined.

First, by (10.4.1), we may choose $r_{\min} \geq 1$ such that

$$q \operatorname{Vol}_n(\tau) + \operatorname{Vol}_n(\sigma) > \prod_{i=1}^{n} \left(1 + (q-1) \sum_{l=i+1}^{n} \frac{1}{r_{\min}^{l-i}} \right). \tag{10.7}$$

By Propositions 8.12, 9.5, and 10.3, there are $\xi_1, \ldots, \xi_n \in K$ satisfying (8.12.1), such that the following is true for all sufficiently large D. Let $h_i = h_K(\xi_i)$ and $d_i = \lfloor D/h_i \rfloor$ for all $i = 1, \ldots, n$. Then there is a nonzero polynomial P as in (9.5.2), of degree at most d_i in x_i for all i, such that (9.5.3) and (10.3.3) hold. By (8.12.1) and (10.7), we may also assume that D is sufficiently large so that

$$q \operatorname{Vol}_n(\tau) + \operatorname{Vol}_n(\sigma) > \prod_{i=1}^{n} \left(1 + (q-1) \sum_{l=i+i}^{n} \frac{d_l}{d_i} \right). \tag{10.8}$$

Let $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$. Then, in the notation of Theorem 10.5, we have $t_j \geq \tau$ for all $j = 1, \ldots, q$ and $t_{q+1} \geq \sigma$ by (9.5.3) and (10.3.3), respectively.

Thus (10.8) contradicts (10.5.1) (via Remark 10.6), and Theorem 4.5 is proved. □

Acknowledgements

I thank the Vietnam Institute for Advanced Study in Mathematics (VIASM) and the Institute of Mathematics of the Academia Sinica in Taiwan for their kind hospitality during brief stays while part of the work for this paper took place. Also, I thank the referees of this paper for careful checking and for many helpful suggestions, and Xinyi Yuan for helpful discussions.

References

[Bombieri and Gubler 2006] E. Bombieri and W. Gubler, *Heights in Diophantine geometry*, New Math. Monogr. **4**, Cambridge Univ. Press, 2006. MR Zbl

[Bosch et al. 1984] S. Bosch, U. Güntzer, and R. Remmert, *Non-Archimedean analysis: a systematic approach to rigid analytic geometry*, Grundlehren der Math. Wissenschaften **261**, Springer, 1984. MR Zbl

[Burgos Gil et al. 2016] J. I. Burgos Gil, P. Philippon, and M. Sombra, "Height of varieties over finitely generated fields", *Kyoto J. Math.* **56**:1 (2016), 13–32. MR Zbl

[Chen 2008] H. Chen, "Positive degree and arithmetic bigness", preprint, 2008. arXiv

[Dyson 1947] F. J. Dyson, "The approximation to algebraic numbers by rationals", *Acta Math.* **79** (1947), 225–240. MR Zbl

[EGA IV$_2$ 1965] A. Grothendieck, "Eléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, II", *Inst. Hautes Études Sci. Publ. Math.* **24** (1965), 5–231. MR Zbl

[Esnault and Viehweg 1984] H. Esnault and E. Viehweg, "Dyson's lemma for polynomials in several variables (and the theorem of Roth)", *Invent. Math.* **78**:3 (1984), 445–490. MR Zbl

[Faltings 1992] G. Faltings, *Lectures on the arithmetic Riemann–Roch theorem*, Ann. of Math. Stud. **127**, Princeton Univ. Press, 1992. MR Zbl

[Gillet and Soulé 1990] H. Gillet and C. Soulé, "Arithmetic intersection theory", *Inst. Hautes Études Sci. Publ. Math.* **72** (1990), 93–174. MR Zbl

[Griffiths and Harris 1978] P. Griffiths and J. Harris, *Principles of algebraic geometry*, Wiley, New York, 1978. MR Zbl

[Gubler 1997] W. Gubler, "Heights of subvarieties over M-fields", pp. 190–227 in *Arithmetic geometry* (Cortona, Italy, 1994), edited by F. Catanese, Sympos. Math. **37**, Cambridge Univ. Press, 1997. MR Zbl

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Grad. Texts Math. **52**, Springer, 1977. MR Zbl

[Hindry and Silverman 2000] M. Hindry and J. H. Silverman, *Diophantine geometry: an introduction*, Grad. Texts Math. **201**, Springer, 2000. MR Zbl

[Lang 1960] S. Lang, "Integral points on curves", *Inst. Hautes Études Sci. Publ. Math.* **6** (1960), 27–43. MR Zbl

[Lang 1962] S. Lang, *Diophantine geometry*, Interscience Tracts Pure Appl. Math. **11**, Interscience, New York, 1962. MR Zbl

[Lang 1974] S. Lang, "Higher dimensional Diophantine problems", *Bull. Amer. Math. Soc.* **80** (1974), 779–787. MR Zbl

[Lang 1983] S. Lang, *Fundamentals of Diophantine geometry*, Springer, 1983. MR Zbl

[Lang 1986] S. Lang, "Hyperbolic and Diophantine analysis", *Bull. Amer. Math. Soc.* (*N.S.*) **14**:2 (1986), 159–205. MR Zbl

[Lang 1987] S. Lang, *Introduction to complex hyperbolic spaces*, Springer, 1987. MR Zbl

[Lang 1991] S. Lang, *Number theory, III: Diophantine geometry*, Encycl. Math. Sci. **60**, Springer, 1991. MR Zbl

[Lang 2002] S. Lang, *Algebra*, 3rd ed., Grad. Texts Math. **211**, Springer, 2002. MR Zbl

[LeVeque 1956] W. J. LeVeque, *Topics in number theory*, *II*, Addison-Wesley, Reading, MA, 1956. MR Zbl

[Moriwaki 2000] A. Moriwaki, "Arithmetic height functions over finitely generated fields", *Invent. Math.* **140**:1 (2000), 101–142. MR Zbl

[Moriwaki 2002] A. Moriwaki, "Diophantine geometry viewed from Arakelov geometry", *Sugaku* **54**:2 (2002), 113–129. In Japanese; translated in *Sugaku Expo.* **17**:2 (2004), 219–234. MR Zbl

[Moriwaki 2009] A. Moriwaki, "Continuity of volumes on arithmetic varieties", *J. Algebraic Geom.* **18**:3 (2009), 407–457. MR Zbl

[Moriwaki 2014] A. Moriwaki, *Arakelov geometry*, Transl. Math. Monogr. **244**, Amer. Math. Soc., Providence, RI, 2014. MR Zbl

[Osgood 1985] C. F. Osgood, "Sometimes effective Thue–Siegel–Roth–Schmidt–Nevanlinna bounds, or better", *J. Number Theory* **21**:3 (1985), 347–389. MR Zbl

[Rastegar 2015] A. Rastegar, "A geometric formulation of Siegel's Diophantine theorem", preprint, 2015. arXiv

[Ridout 1958] D. Ridout, "The p-adic generalization of the Thue–Siegel–Roth theorem", *Mathematika* **5** (1958), 40–48. MR Zbl

[Roth 1955] K. F. Roth, "Rational approximations to algebraic numbers", *Mathematika* **2**:1 (1955), 1–20. Correction in **2**:2 (1955), 168. MR Zbl

[Soulé 1992] C. Soulé, *Lectures on Arakelov geometry*, Cambridge Stud. Adv. Math. **33**, Cambridge Univ. Press, 1992. MR Zbl

[Vojta 1987] P. Vojta, *Diophantine approximations and value distribution theory*, Lecture Notes in Math. **1239**, Springer, 1987. MR Zbl

[Vojta 2007] P. Vojta, "Nagata's embedding theorem", preprint, 2007. arXiv

[Vojta 2011] P. Vojta, "Diophantine approximation and Nevanlinna theory", pp. 111–224 in *Arithmetic geometry* (Cetraro, Italy, 2007), edited by P. Corvaja and C. Gasbarri, Lecture Notes in Math. **2009**, Springer, 2011. MR Zbl

[Wang 1996] J. T.-Y. Wang, "An effective Roth's theorem for function fields", *Rocky Mountain J. Math.* **26**:3 (1996), 1225–1234. MR Zbl

[Wirsing 1971] E. A. Wirsing, "On approximations of algebraic numbers by algebraic numbers of bounded degree", pp. 213–247 in 1969 *Number Theory Institute* (New York, 1969), edited by D. J. Lewis, Proc. Sympos. Pure Math. **20**, Amer. Math. Soc., Providence, RI, 1971. MR Zbl

[Yuan 2008] X. Yuan, "Big line bundles over arithmetic varieties", *Invent. Math.* **173**:3 (2008), 603–649. MR Zbl

[Yuan 2009] X. Yuan, "On volumes of arithmetic line bundles", *Compos. Math.* **145**:6 (2009), 1447–1464. MR Zbl

[Zhang 1995] S. Zhang, "Positive line bundles on arithmetic varieties", *J. Amer. Math. Soc.* **8**:1 (1995), 187–221. MR Zbl

Communicated by Joseph H. Silverman
Received 2019-11-24 Revised 2020-12-03 Accepted 2021-01-17

vojta@math.berkeley.edu *Department of Mathematics, University of California Berkeley, Berkeley, CA, United States*

mathematical sciences publishers msp

Algebra & Number Theory

msp.org/ant

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

The subscription price for 2021 is US $415/year for the electronic version, and $620/year (+$60, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

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

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

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

Algebra & Number Theory

Volume 15 No. 8 2021