Diophantine Approximation and Nevanlinna Theory

Paul Vojta

1 Introduction

Beginning with the work of Osgood [65], it has been known that the branch of complex analysis known as *Nevanlinna theory* (also called *value distribution theory*) has many similarities with Roth's theorem on diophantine approximation. This was extended by the author [87] to include an explicit dictionary and to include geometric results as well, such as Picard's theorem and Mordell's conjecture (Faltings' theorem). The latter analogy ties in with Lang's conjecture that a projective variety should have only finitely many rational points over any given number field (i.e., is *Mordellic*) if and only if it is Kobayashi hyperbolic.

This circle of ideas has developed further in the last 20 years: Lang's conjecture on sharpening the error term in Roth's theorem was carried over to a conjecture in Nevanlinna theory which was proved in many cases. In the other direction, Bloch's conjectures on holomorphic curves in abelian varieties (later proved; see Sect. 15 for details) led to proofs of the corresponding results in number theory (again, see Sect. 15). More recently, work in number theory using Schmidt's Subspace Theorem has led to corresponding results in Nevanlinna theory.

This relation with Nevanlinna theory is in some sense similar to the (much older) relation with function fields, in that one often looks to function fields or Nevanlinna theory for ideas that might translate over to the number field case, and that work over function fields or in Nevanlinna theory is often easier than work in the number field case. On the other hand, both function fields and Nevanlinna theory have concepts that (so far) have no counterpart in the number field case. This is especially true of derivatives, which exist in both the function field case and in Nevanlinna theory. In the number field case, however, one would want the "derivative with respect to p," which remains as a major stumbling block, although (separate) work of

P. Vojta (✉)
Department of Mathematics, University of California, 970 Evans Hall #3840, Berkeley, CA 94720-3840, USA
e-mail: vojta@math.berkeley.edu

P. Corvaja and C. Gasbarri (eds.), *Arithmetic Geometry*, Lecture Notes in Mathematics 2009, DOI 10.1007/978-3-642-15945-9_3,

Buium and of Minhyong Kim may ultimately provide some answers. The search for such a derivative is also addressed in these notes, using a potential approach using successive minima.

It is important to note, however, that the relation with Nevanlinna theory does not "go through" the function field case. Although it is possible to look at the function field case over $\mathbb{C}$ and apply Nevanlinna theory to the functions representing the rational points, this is not the analogy being described here. Instead, in the analogy presented here, *one* holomorphic function corresponds to *infinitely many* rational or algebraic points (whether over a number field or over a function field). Thus, the analogy with Nevanlinna theory is less concrete, and may be regarded as a more distant analogy than function fields.

These notes describe some of the work in this area, including much of the necessary background in diophantine geometry. Specifically, Sects. 2–4 recall basic definitions of number theory and the theory of heights of elements of number fields, culminating in the statement of Roth's theorem and some equivalent formulations of that theorem. This part assumes that the reader knows the basics of algebraic number theory and algebraic geometry at the level of Lang [45] and Hartshorne [36], respectively. Some proofs are omitted, however; for those the interested reader may refer to Lang [46].

Sections 5–7 briefly introduce Nevanlinna theory and the analogy between Roth's theorem and the classical work of Nevanlinna. Again, many proofs are omitted; references include Shabat [75], Nevanlinna [60], and Goldberg and Ostrovskii [29] for pure Nevanlinna theory, and Vojta [87] and Ru [69] for the analogy.

Sections 8–16 generalize the content of the earlier sections, in the more geometric context of varieties over the appropriate fields (number fields, function fields, or $\mathbb{C}$). Again, proofs are often omitted; most may be found in the references given above.

Section 15 in particular introduces the main conjectures being discussed here: Conjecture 15.2 in Nevanlinna theory ("Griffiths' conjecture") and its counterpart in number theory, the author's Conjecture 15.6 on rational points on varieties.

Sections 17 and 18 round out the first part of these notes, by discussing the function field case and the subject of the exceptional sets that come up in the study of higher dimensional varieties.

In both Nevanlinna theory and number theory, these conjectures have been proved only in very special cases, mostly involving subvarieties of semiabelian varieties. This includes the case of projective space minus a collection of hyperplanes in general position (Cartan's theorem and Schmidt's Subspace Theorem). Recent work of Corvaja, Zannier, Evertse, Ferretti, and Ru has shown, however, that using geometric constructions one can use Schmidt's Subspace Theorem and Cartan's theorem to derive other weak special cases of the conjectures mentioned above. This is the subject of Sects. 19–23.

Sections 24–28 present generalizations of Conjectures 15.2 and 15.6. Conjecture 15.2, in Nevanlinna theory, can be generalized to involve truncated counting functions (as was done by Nevanlinna in the 1-dimensional case), and can also be posed in the context of finite ramified coverings. In number theory, Conjecture 15.6 can also be generalized to involve truncated counting functions. The simplest nontrivial case of this conjecture, involving the divisor $[0]+[1]+[\infty]$ on $\mathbb{P}^1$, is the celebrated "abc conjecture" of Masser and Oesterlé. Thus, Conjecture 23.5 can be regarded as a generalization of the abc conjecture as well as of Conjecture 15.6. One can also generalize Conjecture 15.6 to treat algebraic points; this corresponds to finite ramified coverings in Nevanlinna theory. This is Conjecture 25.1, which can also be posed using truncated counting functions (Conjecture 25.3).

Sections 29 and 30 briefly discuss the question of derivatives in Nevanlinna theory, and Nevanlinna's "Lemma on the Logarithmic Derivative." A geometric form of this lemma, due to Kobayashi, McQuillan, and Wong, is given, and it is shown how this form leads to an inequality in Nevanlinna theory, due to McQuillan, called the "tautological inequality." This inequality then leads to a conjecture in number theory (Conjecture 30.1), which of course should then be called the "tautological conjecture." This conjecture is discussed briefly; it is of interest since it may shed some light on how one might take "derivatives" in number theory.

The abc conjecture infuses much of this theory, not only because a Nevanlinna-like conjecture with truncated counting functions contains the abc conjecture as a special case, but also because other conjectures also imply the abc conjecture, and therefore are "at least as hard" as abc. Specifically, Conjecture 25.1, on algebraic points, implies the abc conjecture, even if known only in dimension 1, and Conjecture 15.6, on rational points, also implies abc if known in high dimensions. This latter implication is the subject of Sect. 31. Finally, implications in the other direction are explored in Sect. 32.

2 Notation and Basic Results: Number Theory

We assume that the reader has an understanding of the fundamental basic facts of number theory (and algebraic geometry), up through the definitions of (Weil) heights. References for these topics include [46] and [87]. We do, however, recall some of the basic conventions here since they often differ from author to author.

Throughout these notes, k will usually denote a number field; if so, then $\mathscr{O}_k$ will denote its ring of integers and M_k its set of places. This latter set is in one-to-one correspondence with the disjoint union of the set of nonzero prime ideals of $\mathscr{O}_k$, the set of real embeddings $\sigma\colon k \hookrightarrow \mathbb{R}$, and the set of unordered complex conjugate pairs $(\sigma, \overline{\sigma})$ of complex embeddings $\sigma\colon k \hookrightarrow \mathbb{C}$ with image not contained in $\mathbb{R}$. Such elements of M_k are called **non-archimedean** or **finite** places, **real** places, and **complex** places, respectively.

The real and complex places are collectively referred to as **archimedean** or **infinite** places. The set of these places is denoted S_∞. It is a finite set.

To each place $v \in M_k$ we associate a **norm** $\|\cdot\|_v$, defined for $x \in k$ by $\|x\|_v = 0$ if $x = 0$ and

$$\|x\|_v = \begin{cases} (\mathscr{O}_k : \mathfrak{p})^{\operatorname{ord}_\mathfrak{p}(x)} & \text{if } v \text{ corresponds to } \mathfrak{p} \subseteq \mathscr{O}_k; \\ |\sigma(x)| & \text{if } v \text{ corresponds to } \sigma\colon k \hookrightarrow \mathbb{R}; \text{ and} \\ |\sigma(x)|^2 & \text{if } v \text{ is a complex place, corresponding to } \sigma\colon k \hookrightarrow \mathbb{C} \end{cases} \tag{1}$$

if $x \neq 0$. Here $\operatorname{ord}_{\mathfrak{p}}(x)$ means the exponent of $\mathfrak{p}$ in the factorization of the fractional ideal (x). If we use the convention that $\operatorname{ord}_{\mathfrak{p}}(0) = \infty$, then (1) is also valid when $x = 0$.

We refer to $\|\cdot\|_v$ as a norm instead of an absolute value, because $\|\cdot\|_v$ does not satisfy the triangle inequality when v is a complex place. However, let

$$N_v = \begin{cases} 0 & \text{if } v \text{ is non-archimedean;} \\ 1 & \text{if } v \text{ is real; and} \\ 2 & \text{if } v \text{ is complex.} \end{cases} \tag{2}$$

Then the norm associated to a place v of k satisfies the axioms

(3.1) $\|x\|_v \geq 0$, with equality if and only if $x = 0$;
(3.2) $\|xy\|_v = \|x\|_v\|y\|_v$ for all $x, y \in k$; and
(3.3) $\|x_1 + \cdots + x_n\|_v \leq n^{N_v} \max\{\|x_1\|_v, \ldots, \|x_n\|_v\}$ for all $x_1, \ldots, x_n \in k$, $n \in \mathbb{N}$.

(In these notes, $\mathbb{N} = \{0, 1, 2, \ldots\}$.)

Some authors treat complex conjugate embeddings as distinct places. We do not do so here, because they give rise to the same norms.

Note that, if $x \in k$, then x lies in the ring of integers if and only if $\|x\|_v \leq 1$ for all non-archimedean places v. Indeed, if $x \neq 0$ then both conditions are equivalent to the fractional ideal (x) being a genuine ideal.

Let L be a finite extension of a number field k, and let w be a place of L. If w is non-archimedean, corresponding to a nonzero prime ideal $\mathfrak{q} \subseteq \mathscr{O}_L$, then $\mathfrak{p} := \mathfrak{q} \cap \mathscr{O}_k$ is a nonzero prime of $\mathscr{O}_k$, and gives rise to a non-archimedean place $v \in M_k$. If v arises from w in this way, then we say that w **lies over** v, and write $w \mid v$. Likewise, if w is archimedean, then it corresponds to an embedding $\tau \colon L \hookrightarrow \mathbb{C}$, and its restriction $\tau\big|_k \colon k \hookrightarrow \mathbb{C}$ gives rise to a unique archimedean place $v \in M_k$, and again we say that w lies over v and write $w \mid v$.

For each $v \in M_k$, the set of $w \in M_L$ lying over it is nonempty and finite. If $w \mid v$ then we also say that v **lies under** w.

If S is a subset of M_k, then we say $w \mid S$ if w lies over some place in S; otherwise we write $w \nmid S$.

If $w \mid v$, then we have

$$\|x\|_w = \|x\|_v^{[L_w:k_v]} \qquad \text{for all } x \in k, \tag{4}$$

where L_w and k_v denote the completions of L and k at w and v, respectively. We also have

$$\prod_{\substack{w \in M_L \\ w \mid v}} \|y\|_w = \|N_k^L y\|_v \qquad \text{for all } v \in M_k \text{ and all } y \in L\,. \tag{5}$$

This is proved by using the isomorphism $L \otimes_k k_v \cong \prod_{w \mid v} L_w$; see for example [59, Chap. II, Cor. 8.4].

Let $L/K/k$ be a tower of number fields, and let w' and v be places of L and k, respectively. Then $w' \mid v$ if and only if there is a place w of K satisfying $w' \mid w$ and $w \mid v$.

The field $k = \mathbb{Q}$ has no complex places, one real place corresponding to the inclusion $\mathbb{Q} \subseteq \mathbb{R}$, and infinitely many non-archimedean places, corresponding to prime rational integers. Thus, we write

$$M_{\mathbb{Q}} = \{\infty, 2, 3, 5, \dots\} .$$

Places of a number field satisfy a **Product Formula**

$$\prod_{v \in M_k} \|x\|_v = 1 \qquad \text{for all } x \in k^* . \tag{6}$$

This formula plays a key role in number theory: it is used to show that certain expressions for the height are well defined, and it also implies that if $\prod_v \|x\|_v < 1$ then $x = 0$.

The Product Formula is proved first by showing that it is true when $k = \mathbb{Q}$ (by direct verification) and then using (5) to pass to an arbitrary number field.

3 Heights

The height of a number, or of a point on a variety, is a measure of the complexity of that number. For example, $100/201$ and $1/2$ are very close to each other (using the norm at the infinite place, at least), but the latter is a much "simpler" number since it can be written down using fewer digits.

We define the **height** (also called the **Weil height**) of an element $x \in k$ by the formula

$$H_k(x) = \prod_{v \in M_k} \max\{\|x\|_v, 1\} . \tag{7}$$

As an example, consider the special case in which $k = \mathbb{Q}$. Write $x = a/b$ with $a, b \in \mathbb{Z}$ relatively prime. For all (finite) rational primes p, if p^i is the largest power of p dividing a, and p^j is the largest power dividing b, then $\|a\|_p = p^{-i}$ and $\|b\|_p = p^{-j}$, and therefore $\max\{\|x\|_p, 1\} = p^b$. Therefore the product of all terms in (7) over all finite places v is just $|b|$. At the infinite place, we have $\|x\|_\infty = |a/b|$, so this gives

$$H_{\mathbb{Q}}(x) = \max\{|a|, |b|\} . \tag{8}$$

Similarly, if $P \in \mathbb{P}^n(k)$ for some $n \in \mathbb{N}$, we define the Weil height $h_k(P)$ as follows. Let $[x_0 : \dots : x_n]$ be homogeneous coordinates for P (with the x_i always assumed to lie in k). Then we define

$$H_k(P) = \prod_{v \in M_k} \max\{\|x_0\|_v, \dots, \|x_n\|_v\} . \tag{9}$$

By the Product Formula (6), this quantity is independent of the choice of homogeneous coordinates.

If we identify k with $\mathbb{A}^1(k)$ and identify the latter with a subset of $\mathbb{P}^1(k)$ via the standard injection $i\colon \mathbb{A}^1 \hookrightarrow \mathbb{P}^1$, then we note that $H_k(x) = H_k(i(x))$ for all $x \in k$. Similarly, we can identify k^n with $\mathbb{A}^n(k)$, and the standard embedding of $\mathbb{A}^n$ into $\mathbb{P}^n$ gives us a height

$$H_k(x_1,\dots,x_n) = \prod_{v\in M_k} \max\{\|x_1\|_v,\dots,\|x_n\|_v,1\}\ .$$

The height functions defined so far, all using capital "H," are called **multiplicative heights**. Usually it is more convenient to take their logarithms and define **logarithmic heights**:

$$h_k(x) = \log H_k(x) = \sum_{v\in M_k} \log^+ \|x\|_v \tag{10}$$

and

$$h_k([x_0:\dots:x_n]) = \log H_k([x_0:\dots:x_n]) = \sum_{v\in M_k} \log\max\{\|x_0\|_v,\dots,\|x_n\|_v\}.$$

Here

$$\log^+(x) = \max\{\log x, 0\}.$$

Equation (5) tells us how heights change when the number field k is extended to a larger number field L:

$$h_L(x) = [L:k]h_k(x) \tag{11}$$

and

$$h_L([x_0:\dots:x_n]) = [L:k]h_k([x_0:\dots:x_n]) \tag{12}$$

for all $x \in k$ and all $[x_0,\dots,x_n] \in \mathbb{P}^n(k)$, respectively.

Then, given $x \in \overline{\mathbb{Q}}$, we define

$$h_k(x) = \frac{1}{[L:k]} h_L(x)$$

for any number field $L \supseteq k(x)$, and similarly given any $[x_0:\dots:x_n] \in \mathbb{P}^n(\overline{\mathbb{Q}})$, we define

$$h_k([x_0:\dots:x_n]) = \frac{1}{[L:k]} h_L([x_0:\dots:x_n])$$

for any number field $L \supseteq k(x_0,\dots,x_n)$. These expressions are independent of the choice of L by (11) and (12), respectively.

Following EGA, if x is a point on $\mathbb{P}^n_k$, then $\kappa(x)$ will denote the residue field of the local ring at x. If x is a closed point then the homogeneous coordinates can be chosen such that $k(x_0,\dots,x_n) = \kappa(x)$.

With these definitions, (11) and (12) remain valid without the conditions $x \in k$ and $[x_0 : \dots : x_n] \in \mathbb{P}^n(k)$, respectively.

It is common to assume $k = \mathbb{Q}$ and omit the subscript k. The resulting heights are called **absolute heights**.

It is obvious from (7) that $h_k(x) \ge 0$ for all $x \in k$, and that equality holds if $x = 0$ or if x is a root of unity. Conversely, $h_k(x) = 0$ implies $\|x\|_v \le 1$ for all v; if $x \ne 0$ then the Product Formula implies $\|x\|_v = 1$ for all v. Thus x must be a unit, and the known structure of the unit group then leads to the fact that x must be a root of unity.

Therefore, there are infinitely many elements of $\overline{\mathbb{Q}}$ with height 0. If one bounds the degree of such elements over $\mathbb{Q}$, then there are only finitely many; more generally, we have:

Theorem 3.1. (Northcott's finiteness theorem) *For any $r \in \mathbb{Z}_{>0}$ and any $C \in \mathbb{R}$, there are only finitely many $x \in \overline{\mathbb{Q}}$ such that $[\mathbb{Q}(x) : \mathbb{Q}] \le r$ and $h(x) \le C$. Moreover, given any $n \in \mathbb{N}$ there are only finitely many $x \in \mathbb{P}^n(\overline{\mathbb{Q}})$ such that $[\kappa(x) : \mathbb{Q}] \le r$ and $h(x) \le C$.*

The first assertion is proved using the fact that, for any $x \in \overline{\mathbb{Q}}$, if one lets $k = \mathbb{Q}(x)$, then $H_k(x)$ is within a constant factor of the largest absolute value of the largest coefficient of the irreducible polynomial of x over $\mathbb{Q}$, when that polynomial is multiplied by a rational number so that its coefficients are relatively prime integers. The second assertion then follows as a consequence of the first. For details, see [48, Chap. II, Thm. 2.2].

This result plays a central role in number theory, since (for example) proving an upper bound on the heights of rational points is equivalent to proving finiteness.

4 Roth's Theorem

Roth [67] proved a key and much-anticipated theorem on how well an algebraic number can be approximated by rational numbers. Of course rational numbers are dense in the reals, but if one limits the size of the denominator then one can ask meaningful and nontrivial questions.

Theorem 4.1. (Roth) *Fix $\alpha \in \overline{\mathbb{Q}}$, $\varepsilon > 0$, and $C > 0$. Then there are only finitely many $a/b \in \mathbb{Q}$, where a and b are relatively prime integers, such that*

$$\left|\frac{a}{b} - \alpha\right| \le \frac{C}{|b|^{2+\varepsilon}}. \tag{13}$$

Example 4.2. As a diophantine application of Roth's theorem, consider the diophantine equation

$$x^3 - 2y^3 = 11, \qquad x, y \in \mathbb{Z}. \tag{14}$$

If (x, y) is a solution, then x/y must be close to $\sqrt[3]{2}$ (assuming $|x|$ or $|y|$ is large, which would imply both are large):

$$\left|\frac{x}{y}-\sqrt[3]{2}\right|=\left|\frac{11}{y(x^2+xy\sqrt[3]{2}+y^2\sqrt[3]{4})}\right|\ll\frac{1}{|y|^3}.$$

Thus Roth's theorem implies that (14) has only finitely many solutions.

More generally, if $f\in\mathbb{Z}[x,y]$ is homogeneous of degree ≥ 3 and has no repeated factors, then for any $a\in\mathbb{Z}$ $f(x,y)=a$ has only finitely many integral solutions. This is called the **Thue equation** and historically was the driving force behind the development of Roth's theorem (which is sometimes called the Thue-Siegel-Roth theorem, sometimes also mentioning Schneider, Dyson, and Mahler).

The inequality (13) is best possible, in the sense that the 2 in the exponent on the right-hand side cannot be replaced by a smaller number. This can be shown using continued fractions. Of course one can conjecture a sharper error term [49, Intro. to Chap. I].

If a/b is close to α, then after adjusting C one can replace $|b|$ in the right-hand side of (13) with $H_{\mathbb{Q}}(a/b)$ (see (8)). Moreover, the theorem has been generalized to allow a finite set of places (possibly non-archimedean) and to work over a number field:

Theorem 4.3. *Let k be a number field, let S be a finite set of places of k containing all archimedean places, fix $\alpha_v\in\overline{\mathbb{Q}}$ for each $v\in S$, let $\varepsilon>0$, and let $C>0$. Then only finitely many $x\in k$ satisfy the inequality*

$$\prod_{v\in S}\min\{1,\|x-\alpha_v\|_v\}\le\frac{C}{H_k(x)^{2+\varepsilon}}. \tag{15}$$

(Strictly speaking, S can be any finite set of places at this point, but requiring S to contain all archimedean places does not weaken the theorem, and this assumption will be necessary in Sect. 6. See, for example, (29).)

Taking $-\log$ of both sides of (15), dividing by $[k:\mathbb{Q}]$, and rephrasing the logic, the above theorem is equivalent to the assertion that for all $c\in\mathbb{R}$ the inequality

$$\frac{1}{[k:\mathbb{Q}]}\sum_{v\in S}\log^+\left\|\frac{1}{x-\alpha_v}\right\|_v\le(2+\varepsilon)h(x)+c \tag{16}$$

holds for all but finitely many $x\in k$.

In writing (15), we assume that one has chosen an embedding $i_v\colon\bar{k}\hookrightarrow\overline{k_v}$ over k for each $v\in S$. Otherwise the expression $\|x-\alpha_v\|_v$ may not make sense.

This is mostly a moot point, however, since we may restrict to $\alpha_v\in k$ for all v. Clearly this restricted theorem is implied by the theorem without the additional restriction, but in fact it also implies the original theorem. To see this, suppose k, S, ε, and c are as above, and assume that $\alpha_v\in\overline{\mathbb{Q}}$ are given for all $v\in S$. Let L be the Galois closure over k of $k(\alpha_v:v\in S)$, and let T be the set of all places of L lying over places in S. We assume that L is a subfield of $\bar{k}$, so that $\alpha_v\in L$ for all $v\in S$. Then $(i_v)\big|_L\colon L\to\overline{k_v}$ induces a place w_0 of L over v, and all other places w of L over v are conjugates by elements $\sigma_w\in\operatorname{Gal}(L/k)$:

$\|x\|_w = \|\sigma_w^{-1}(x)\|_{w_0}$ for all $x \in L$. Letting $\alpha_w = \sigma_w(\alpha_v)$ for all $w \mid v$, we then have $\|x-\alpha_w\|_w = \|\sigma_w^{-1}(x-\alpha_w)\|_{w_0} = \|x-\alpha_v\|_v^{[L_{w_0}:k_v]}$ for all $x \in k$ by (4), and therefore

$$\sum_{w|v} \log^+ \left\|\frac{1}{x-\alpha_w}\right\|_w = \sum_{w|v} [L_{w_0}:k_v] \log^+ \left\|\frac{1}{x-\alpha_v}\right\|_v = [L:k] \log^+ \left\|\frac{1}{x-\alpha_v}\right\|_v$$

since L/k is Galois. Thus

$$\frac{1}{[k:\mathbb{Q}]} \sum_{v\in S} \log^+ \left\|\frac{1}{x-\alpha_v}\right\|_v = \frac{1}{[L:\mathbb{Q}]} \sum_{w\in T} \log^+ \left\|\frac{1}{x-\alpha_w}\right\|_w$$

for all $x \in k$. Applying Roth's theorem over the field L (where now $\alpha_w \in L$ for all $w \in T$) then gives (16) for almost all $x \in k$.

Finally, we note that Roth's theorem (as now rephrased) is equivalent to the following statement.

Theorem 4.4. *Let k be a number field, let $S \supseteq S_\infty$ be a finite set of places of k, fix distinct $\alpha_1, \dots, \alpha_q \in k$, let $\varepsilon > 0$, and let $c \in \mathbb{R}$. Then the inequality*

$$\frac{1}{[k:\mathbb{Q}]} \sum_{v\in S} \sum_{i=1}^{q} \log^+ \left\|\frac{1}{x-\alpha_i}\right\|_v \le (2+\varepsilon)h(x) + c \tag{17}$$

holds for almost all $x \in k$.

Indeed, given $\alpha_v \in k$ for all $v \in S$, let $\alpha_1, \dots, \alpha_q$ be the distinct elements of the set $\{\alpha_v : v \in S\}$. Then

$$\frac{1}{[k:\mathbb{Q}]} \sum_{v\in S} \log^+ \left\|\frac{1}{x-\alpha_v}\right\|_v \le \frac{1}{[k:\mathbb{Q}]} \sum_{v\in S} \sum_{i=1}^{q} \log^+ \left\|\frac{1}{x-\alpha_i}\right\|_v ,$$

so Theorem 4.4 implies the earlier form of Roth's theorem (as modified).

Conversely, given distinct $\alpha_1, \dots, \alpha_q \in k$, we note that any given $x \in k$ can be close to only one of the α_i at each place v (where the value of i may depend on v).

Therefore, for each v,

$$\sum_{i=1}^{q} \log^+ \left\|\frac{1}{x-\alpha_i}\right\|_v \le \log^+ \left\|\frac{1}{x-\alpha_v}\right\|_v + c_v$$

for some constant c_v independent of x and some $\alpha_v \in \{\alpha_1, \dots, \alpha_q\}$ depending on x and v. Thus, for each $x \in k$, there is a choice of α_v for each $v \in S$ such that

$$\frac{1}{[k:\mathbb{Q}]} \sum_{v\in S} \sum_{i=1}^{q} \log^+ \left\|\frac{1}{x-\alpha_i}\right\|_v \le \frac{1}{[k:\mathbb{Q}]} \sum_{v\in S} \log^+ \left\|\frac{1}{x-\alpha_v}\right\|_v + c',$$

where c' is independent of x. Since there are only finitely many choices of the system $\{\alpha_v : v \in S\}$, finitely many applications of the earlier version of Roth's theorem suffice to imply Theorem 4.4.

5 Basics of Nevanlinna Theory

Nevanlinna theory, developed by R. and F. Nevanlinna in the 1920s, concerns the distribution of values of holomorphic and meromorphic functions, in much the same way that Roth's theorem concerns approximation of elements of a number field.

One can think of it as a generalization of a theorem of Picard, which says that a nonconstant holomorphic function from $\mathbb{C}$ to $\mathbb{P}^1$ can omit at most two points. This, in turn, generalizes Liouville's theorem.

An example relevant to Picard's theorem is the exponential function e^z, which omits the values 0 and ∞. When r is large, the circle $|z| = r$ is mapped to many values close to ∞ (when $\operatorname{Re} z$ is large) and many values close to 0 (when $\operatorname{Re} z$ is highly negative).

So even though e^z omits these two values, it spends a lot of time very close to them. This observation can be made precise, in what is called Nevanlinna's *First Main Theorem.* In order to state this theorem, we need some definitions.

First we recall that $\log^+ x = \max\{\log x, 0\}$, and similarly define

$$\operatorname{ord}_z^+ f = \max\{\operatorname{ord}_z f, 0\}$$

if f is a meromorphic function and $z \in \mathbb{C}$.

Definition 5.1. Let f be a meromorphic function on $\mathbb{C}$. We define the **proximity function** of f by

$$m_f(r) = \int_0^{2\pi} \log^+ \left| f(re^{i\theta}) \right| \frac{d\theta}{2\pi} \tag{18}$$

for all $r > 0$. We also define

$$m_f(\infty, r) = m_f(r) \qquad \text{and} \qquad m_f(a, r) = m_{1/(f-a)}(r)$$

when $a \in \mathbb{C}$.

The integral in (18) converges when f has a zero or pole on the circle $|z| = r$, so it is defined everywhere. The proximity function $m_f(a, r)$ is large to the extent that the values of f on $|z| = r$ are close to a.

Definition 5.2. Let f be a meromorphic function on $\mathbb{C}$. For $r > 0$ let $n_f(r)$ be the number of poles of f in the open disc $|z| < r$ of radius r (counted with multiplicity), and let $n_f(0)$ be the order of the pole (if any) at $z = 0$. We then define the **counting function** of f by

$$N_f(r) = \int_0^r (n_f(s) - n_f(0)) \frac{ds}{s} + n_f(0) \log r. \tag{19}$$

As before, we also define

$$N_f(\infty, r) = N_f(r) \qquad \text{and} \qquad N_f(a, r) = N_{1/(f-a)}(r)$$

when $a \in \mathbb{C}$.

The counting function can also be written

$$N_f(a, r) = \sum_{0<|z|<r} \operatorname{ord}_z^+(f-a) \cdot \log \frac{r}{|z|} + \operatorname{ord}_0^+(f-a) \cdot \log r. \tag{20}$$

Thus, the expression $N_f(a, r)$ is a weighted count, with multiplicity, of the number of times f takes on the value a in the disc $|z| < r$.

Definition 5.3. Let f be as in Definition 5.1. Then the **height function** of f is the function $T_f \colon (0, \infty) \to \mathbb{R}$ given by

$$T_f(r) = m_f(r) + N_f(r). \tag{21}$$

Classically, the above function is called the characteristic function, but here we will use the term height function, since this is more in parallel with terminology in the number field case. The height function T_f does, in fact, measure the complexity of the meromorphic function f.

In particular, if f is constant then so is $T_f(r)$; otherwise,

$$\liminf_{r\to\infty} \frac{T_f(r)}{\log r} > 0. \tag{22}$$

Moreover, it is known that $T_f(r) = O(\log r)$ if and only if f is a rational function. Although this is a well-known fact, I was unable to find a convenient reference, so a proof is sketched here. If f is rational, then direct computation gives $T_f(r) = O(\log r)$. Conversely, if $T_f(r) = O(\log r)$ then f can have only finitely many poles; clearing these by multiplying f by a polynomial changes T_f by at most $O(\log r)$, so we may assume that f is entire. We may also assume that f is nonconstant. By [37, Thm. 1.8], if f is entire and nonconstant and $K > 1$, then

$$\liminf_{r\to\infty} \frac{\log \max_{|z|=r} |f(z)|}{T_f(r)(\log T_f(r))^K} = 0.$$

This implies that $f(z)/z^n$ has a removable singularity at ∞ for sufficiently large n, hence is a polynomial.

The following theorem relates the height function to the proximity and counting functions at points other than ∞.

Theorem 5.4. (First Main Theorem) *Let f be a meromorphic function on $\mathbb{C}$, and let $a \in \mathbb{C}$. Then*

$$T_f(r) = m_f(a,r) + N_f(a,r) + O(1),$$

where the constant in $O(1)$ depends only on f and a.

This theorem is a straightforward consequence of **Jensen's formula**

$$\log|c_f| = \int_0^{2\pi} \log\left|f(re^{i\theta})\right| \frac{d\theta}{2\pi} + N_f(\infty,r) - N_f(0,r),$$

where c_f is the leading coefficient in the Laurent expansion of f at $z=0$. For details, see [60, Chap. VI, (1.2')] or [69, Cor. A1.1.3].

As an example, let $f(z) = e^z$. This function is entire, so $N_f(\infty,r) = 0$ for all r. We also have

$$m_f(\infty,r) = \int_0^{2\pi} \log^+ e^{r\cos\theta} \frac{d\theta}{2\pi} = r\int_{-\pi/2}^{\pi/2} \cos\theta \frac{d\theta}{2\pi} = \frac{r}{\pi}.$$

Thus

$$T_f(r) = \frac{r}{\pi}.$$

Similarly, we have $N_f(0,r) = 0$ and $m_f(0,r) = r/\pi$ for all r, confirming the First Main Theorem in the case $a=0$.

The situation with $a=-1$ is more difficult. The integral in the proximity function seems to be beyond any hope of computing exactly. Since $e^z = -1$ if and only if z is an odd integral multiple of πi, we have

$$N_f(-1,r) = 2\int_0^r \left[\frac{s}{2\pi} + \frac{1}{2}\right] \frac{ds}{s} \approx 2\int_0^r \frac{s}{2\pi}\frac{ds}{s} = \frac{r}{\pi},$$

where $[\cdot]$ denotes the greatest integer function. The error in the above approximation should be $o(r)$, which would give $m(-1,r) = o(r)$. Judging from the general shape of the exponential function, similar estimates should hold for all nonzero $a \in \mathbb{C}$.

In one way of thinking, the First Main Theorem gives an upper bound on the counting function. As the above example illustrates, there is no lower bound for an individual counting function (other than 0), but it is known that there cannot be many values of a for which $N_f(a,r)$ is much smaller than the height. This is what the Second Main Theorem shows.

Theorem 5.5. (Second Main Theorem) *Let f be a meromorphic function on $\mathbb{C}$, and let $a_1,\dots,a_q \in \mathbb{C}$ be distinct numbers. Then*

$$\sum_{j=1}^q m_f(a_j,r) \le_{\text{exc}} 2T_f(r) + O(\log^+ T_f(r)) + o(\log r), \tag{23}$$

where the implicit constants depend only on f and $a_1,\dots,a_q$.

Here the notation $\le_{\text{exc}}$ means that the inequality holds for all $r > 0$ outside of a set of finite Lebesgue measure.

By the First Main Theorem, (23) can be rewritten as a lower bound on the counting functions:

$$\sum_{j=1}^{q} N_f(a_j, r) \ge_{\text{exc}} (q-2)T_f(r) - O(\log^+ T_f(r)) - o(\log r). \tag{24}$$

As another variation, (23) can be written with a weaker error term:

$$\sum_{j=1}^{q} m_f(a_j, r) \le_{\text{exc}} (2+\varepsilon)T_f(r) + c \tag{25}$$

for all $\varepsilon > 0$ and any constant c. The next section will show that this correponds to Roth's theorem.

Corollary 5.6. (Picard's "little" theorem) *If $a_1, a_2, a_3 \in \mathbb{P}^1(\mathbb{C})$ are distinct, then any holomorphic function $f\colon \mathbb{C} \to \mathbb{P}^1(\mathbb{C}) \setminus \{a_1, a_2, a_3\}$ must be constant.*

Proof. Assume that $f\colon \mathbb{C} \to \mathbb{P}^1(\mathbb{C}) \setminus \{a_1, a_2, a_3\}$ is a nonconstant holomorphic function. After applying an automorphism of $\mathbb{P}^1$ if necessary, we may assume that all a_j are finite. We may regard f as a meromorphic function on $\mathbb{C}$.

Since f never takes on the values a_1, a_2, or a_3, the left-hand side of (24) vanishes. Since f is nonconstant, the right-hand side approaches $+\infty$ by (22). This is a contradiction. □

As we have seen, (24) has some advantages over (23). Other advantages include the fact that $q-2$ on the right-hand side is the Euler characteristic of $\mathbb{P}^1$ minus q points, and it will become clear later that the dependence on a metric is restricted to the height term. It is also the preferred form when comparing with the abc conjecture.

6 Roth's Theorem and Nevanlinna Theory

We now claim that Nevanlinna's Second Main Theorem corresponds very closely to Roth's theorem. To see this, we make the following definitions in number theory.

Definition 6.1. Let k be a number field and $S \supseteq S_\infty$ a finite set of places of k. For $x \in k$ we define the **proximity function** to be

$$m_S(x) = \sum_{v \in S} \log^+ \|x\|_v$$

and, for $a \in k$ with $a \neq x$,

$$m_S(a,x) = m_S\left(\frac{1}{x-a}\right) = \sum_{v \in S} \log^+ \left\|\frac{1}{x-a}\right\|_v. \tag{26}$$

Similarly, for distinct $a, x \in k$ the **counting function** is defined as

$$N_S(x) = \sum_{v \notin S} \log^+ \|x\|_v$$

and

$$N_S(a,x) = N_S\left(\frac{1}{x-a}\right) = \sum_{v \notin S} \log^+ \left\|\frac{1}{x-a}\right\|_v. \tag{27}$$

By (10) it then follows that

$$m_S(x) + N_S(x) = \sum_{v \in M_k} \log^+ \|x\|_v = h_k(x) \tag{28}$$

for all $x \in k$. This corresponds to (21).

Note that k does not appear in the notation for the proximity and counting functions, since it is implied by S.

We also note that all places outside of S are non-archimedean, hence correspond to nonzero prime ideals $\mathfrak{p} \subseteq \mathscr{O}_k$. Thus, by (1), (27) can be rewritten

$$N_S(a,x) = \sum_{v \notin S} \operatorname{ord}_{\mathfrak{p}}^+(x-a) \cdot \log(\mathscr{O}_k : \mathfrak{p}), \tag{29}$$

where $\mathfrak{p}$ in the summand is the prime ideal corresponding to v. This corresponds to (20).

The number field case has an analogue of the First Main Theorem, which we prove as follows.

Lemma 6.2. *Let v be a place of a number field k, and let $a, x \in k$. Then*

$$\left|\log^+ \|x\|_v - \log^+ \|x-a\|_v\right| \le \log^+ \|a\|_v + N_v \log 2. \tag{30}$$

Proof. **Case I**: v is archimedean.

We first claim that

$$\log^+(s+t) \le \log^+ s + \log^+ t + \log 2 \tag{31}$$

for all real $s, t \ge 0$. Indeed, let $f(s,t) = \log^+(s+t) - \log^+ s - \log^+ t$. By considering partial derivatives, for each fixed s the function has a global maximum at $t = 1$, and for each fixed t it has a global maximum at $s = 1$. Therefore all s and t satisfy $f(s,t) \le f(1,1) = \log 2$.

Now let $z, b \in \mathbb{C}$. Since $|z| \le |z-b| + |b|$, (31) with $s = |z-b|$ and $t = |b|$ gives

$$\log^+ |z| - \log^+ |z-b| \le \log^+(|z-b| + |b|) - \log^+ |z-b| \le \log^+ |b| + \log 2.$$

Similarly, since $|z-b| \le |z| + |b|$, we have

$$\log^+ |z-b| - \log^+ |z| \le \log^+(|z| + |b|) - \log^+ |z| \le \log^+ |b| + \log 2.$$

These two inequalities together imply (30).

Case II: v is non-archimedean.

In this case $N_v = 0$, so the last term vanishes. Also, since v is non-archimedean, at least two of $\|x\|_v$, $\|x-a\|_v$, and $\|a\|_v$ are equal, and the third (if different) is smaller. If $\|x\|_v = \|x-a\|_v$, then the result is obvious, so we may assume that $\|a\|_v$ is equal to one of the other two. If $\|a\|_v = \|x\|_v$, then

$$\left|\log^+ \|x\|_v - \log^+ \|x-a\|_v\right| = \log^+ \|x\|_v - \log^+ \|x-a\|_v \le \log^+ \|x\|_v = \log^+ \|a\|_v$$

since $0 \le \log^+ \|x-a\|_v \le \log^+ \|x\|_v$. If $\|a\|_v = \|x-a\|_v$ then (30) follows by a similar argument. □

Corresponding to Theorem 5.4, we then have:

Theorem 6.3. *Let k be a number field, let $S \supseteq S_\infty$ be a finite set of places of k, and fix $a \in k$. Then*

$$h_k(x) = m_S(a,x) + N_S(a,x) + O(1),$$

where the constant in $O(1)$ depends only on k and a. In fact, the constant can be taken to be $h_k(a) + [k:\mathbb{Q}]\log 2$.

Proof. First, we note that

$$m_S(a,x) + N_S(a,x) = m_S\left(\frac{1}{x-a}\right) + N_S\left(\frac{1}{x-a}\right) = h_k\left(\frac{1}{x-a}\right).$$

Next, by comparing with the height on $\mathbb{P}^1$, we have

$$h_k\left(\frac{1}{x-a}\right) = h_k([x-a:1]) = h_k([1:x-a]) = h_k(x-a).$$

Therefore, it suffices to show that

$$h_k(x-a) = h_k(x) + O(1),$$

with the constant in $O(1)$ equal to $[k:\mathbb{Q}]\log 2$. This follows immediately by applying Lemma 6.2 termwise to the sums in the two height functions. □

It will later be clear that this theorem is a well-known geometric property of heights.

We now consider the Second Main Theorem. With the notation of Definition 6.1, Roth's theorem can be made to look very similar to Nevanlinna's Second Main Theorem. Indeed, multiplying (17) by $[k:\mathbb{Q}]$ and substituting the definition (26) of the proximity function gives the inequality

$$\sum_{j=1}^{q} m_S(a_j,x) \le (2+\varepsilon)h_k(x)+c,$$

which corresponds to (25). As has been mentioned earlier, it has been conjectured that Roth's theorem should hold with sharper error terms, corresponding to (23). Such conjectures predated the emergence of the correspondence between number theory and Nevanlinna theory, but the latter spurred renewed work in the area. See, for example, [101, 49, 12].

Unfortunately, the correspondence between the statements of Roth's theorem and Nevanlinna's Second Main Theorem does not extend to the proofs of these theorems. Roth's theorem is proved by taking sufficiently many $x \in k$ not satisfying the inequality, using them to construct an auxiliary polynomial, and then deriving a contradiction from the vanishing properties of that polynomial. Nevanlinna's Second Main Theorem has a number of proofs; for example, one proof uses curvature arguments, one follows from Nevanlinna's "lemma on the logarithmic derivative," and one uses Ahlfors' theory of covering spaces. All of these proofs make essential use of the derivative of the meromorphic function, and it is a major unsolved question in the field to find some analogue of this in number theory.

A detailed discussion of these proofs would be beyond the scope of these notes.

Beyond Roth's theorem and the Second Main Theorem, one can define the *defect* of an element of $\mathbb{C}$ or of an element $a \in k$, as follows.

Definition 6.4. Let f be a meromorphic function on $\mathbb{C}$, and let $a \in \mathbb{C} \cup \{\infty\}$. Then the **defect** of a is

$$\delta_f(a) = \liminf_{r\to\infty} \frac{m_f(a,r)}{T_f(r)}.$$

Similarly, let $S \supseteq S_\infty$ be a finite set of places of a number field k, let $a \in k$, and let Σ be an infinite subset of k. Then the defect is defined as

$$\delta_S(a) = \liminf_{x\in\Sigma} \frac{m_S(a,x)}{h_k(x)}.$$

By the First Main Theorem (Theorems 5.4 and 6.3), we then have

$$0 \le \delta_f(a) \le 1 \qquad \text{and} \qquad 0 \le \delta_S(a) \le 1,$$

respectively. The Second Main Theorems (Theorems 5.5 and 4.4) then give

$$\sum_{a\in\mathbb{C}} \delta_f(a) \le 2 \qquad \text{and} \qquad \sum_{j=1}^{q} \delta_S(a_j) \le 2,$$

respectively. This is just an equivalent formulation of the Second Main Theorem, with a weaker error term in the case of Nevanlinna theory, so it is usually better to work directly with the inequality of the Second Main Theorem.

The defect gets it name because it measures the extent to which $N_f(a,r)$ or $N_S(a,x)$ is smaller than the maximum indicated by the First Main Theorem.

We conclude this section by noting that Definition 6.1 can be extended to $x \in \bar{k}$. Indeed, let k and S be as in Definition 6.1, and let $x \in \bar{k}$. Let L be a number field containing $k(x)$, and let T be the set of places of L lying over places in S. If $L' \supseteq L$ is another number field, and if T' is the set of places of L' lying over places of k, then (4) gives

$$m_{T'}(x) = [L' : L]m_T(x) \qquad \text{and} \qquad N_{T'}(x) = [L' : L]N_T(x). \tag{32}$$

This allows us to make the following definition.

Definition 6.5. Let k, S, x, L, and T be as above. Then we define

$$m_S(x) = \frac{1}{[L:k]} m_T(x) \qquad \text{and} \qquad N_S(x) = \frac{1}{[L:k]} N_T(x).$$

These expressions are independent of $L \supseteq k(x)$ by (32). As in (26) and (27), we also let

$$m_S(a,x) = m_S\left(\frac{1}{x-a}\right) \qquad \text{and} \qquad N_S(a,x) = N_S\left(\frac{1}{x-a}\right).$$

Likewise, Theorem 6.3 (the number-theoretic First Main Theorem) extends to $x \in \bar{k}$, by (11), (32), and (6.5). The expression (28) for the height also extends. Roth's theorem, however, does not extend in this manner, and questions of extending Roth's theorem even to algebraic numbers of bounded degree are quite deep and unresolved.

7 The Dictionary (Non-Geometric Case)

The discussion in the preceding section suggests that there should be an analogy between the fields of Nevanlinna theory and number theory. This section describes this dictionary in more detail.

The existence of an analogy between number theory and Nevanlinna theory was first observed by Osgood [65, 66], but he did not provide an explicit dictionary for comparing the two theories. This was provided by Vojta [87]. An updated version of that dictionary is provided here as Table 1.

The first and most important thing to realize about the dictionary is that the analogue of a holomorphic (or meromorphic) function is an *infinite sequence* of rational numbers. While it is tempting to compare number theory with Nevanlinna theory by way of function fields – by viewing a single rational point as being analogous to a rational point over a function field over $\mathbb{C}$ and then applying Nevanlinna theory to

Table 1 The dictionary in the one-dimensional case

Nevanlinna theory	Number theory
$f\colon \mathbb{C} \to \mathbb{C}$, non-constant	$\{x\} \subseteq k$, infinite
r	x
θ	$v \in S$
$\lvert f(re^{i\theta})\rvert$	$\lVert x \rVert_v$, $v \in S$
$\operatorname{ord}_z f$	$\operatorname{ord}_v x$, $v \notin S$
$\log \dfrac{r}{\lvert z\rvert}$	$\log(\mathscr{O}_k : \mathfrak{p})$
Height function	Logarithmic height
$T_f(r) = \int_0^{2\pi} \log^+ \lvert f(re^{i\theta})\rvert \dfrac{d\theta}{2\pi} + N_f(\infty, r)$	$h_k(x) = \sum_{v \in M_k} \log^+ \lVert x \rVert_v$
Proximity function	
$m_f(a,r) = \int_0^{2\pi} \log^+ \left\lvert \dfrac{1}{f(re^{i\theta}) - a} \right\rvert \dfrac{d\theta}{2\pi}$	$m_S(a,x) = \sum_{v \in S} \log^+ \left\lVert \dfrac{1}{x-a} \right\rVert_v$
Counting function	
$N_f(a,r) = \sum_{\lvert z\rvert < r} \operatorname{ord}_z^+(f-a) \log \dfrac{r}{\lvert z\rvert}$	$N_S(a,x) = \sum_{v \notin S} \operatorname{ord}_{\mathfrak{p}}^+(x-a) \log(\mathscr{O}_k : \mathfrak{p})$
First main theorem	Property of heights
$N_f(a,r) + m_f(a,r) = T_f(r) + O(1)$	$N_S(a,x) + m_S(a,x) = h_k(x) + O(1)$
Second main theorem	Conjectured refinement of Roth
$\sum_{i=1}^m m_f(a_i, r) \le_{\text{exc}} 2T_f(r) - N_{1,f}(r) + O(r \log T_f(r))$	$\sum_{i=1}^m m_S(a_i, x) \le 2h_k(x) + O(\log h_k(x))$
Defect	
$\delta(a) = \liminf_{r \to \infty} \dfrac{m_f(a,r)}{T_f(r)}$	$\delta(a) = \liminf_{x} \dfrac{m_S(a,x)}{h_k(x)}$
Defect relation	Roth's theorem
$\sum_{a \in \mathbb{C}} \delta(a) \le 2$	$\sum_{a \in k} \delta(a) \le 2$
Jensen's formula	Artin-Whaples product formula
$\log \lvert c_f\rvert = \int_0^{2\pi} \log \lvert f(re^{i\theta})\rvert \dfrac{d\theta}{2\pi} + N_f(\infty, r) - N_f(0, r)$	$\sum_{v \in M_k} \log \lVert x \rVert_v = 0$

the corresponding section map – this is not what is being compared here. Note that the Second Main Theorem posits the non-existence of a meromorphic function violating the inequality for too many r, and Roth's theorem claims the non-existence of an infinite sequence of rational numbers not satisfying its main inequality.

We shall now describe Table 1 in more detail. Much of it (below the top six rows) has already been described in Sect. 6, with the exception of the last line. This is left to the reader.

The bottom two-thirds of the table can be broken down further, leading to the top six rows. The first row has been described above. One can say more, though. The analogue of a single rational number can be viewed as the restriction of f to the closed disc $\overline{\mathbb{D}}_r$ of radius r. Of course $f\big|_{\overline{\mathbb{D}}_r}$ for varying r are strongly related, in the sense that if one knows one of them then all of them are uniquely determined. This is not true of the number field case (as far as is known); thus the analogy is not perfect.

However, when comparing $f\big|_{\overline{\mathbb{D}}_r}$ to a given element of k, there are further similarities between the respective proximity functions and counting functions. As far as the proximity functions are concerned, in Nevanlinna theory $m_f(a,r)$ depends only on the values of f on the circle $|z| = r$, whereas in number theory $m_S(a,r)$ involves only the places in S. So places in S correspond to $\partial\mathbb{D}_r$, and both types of proximity functions involve the absolute values at those places. Moreover, in Nevanlinna theory the proximity function is an integral over a set of finite measure, while in number theory the proximity function is a finite sum.

As for counting functions, they involve the open disc $\mathbb{D}_r$ in Nevanlinna theory, and places outside of S (all of which are non-archimedean) in number theory. Both types of counting functions involve an infinite weighted sum of orders of vanishing at those places, and the sixth line of Table 1 compares these weights.

It should also be mentioned that many of these theorems in Nevanlinna theory have been extended to holomorphic functions with domains other than $\mathbb{C}$. In one direction, one can replace the domain with $\mathbb{C}^m$ for some $m > 0$. While this is useful from the point of view of pure Nevanlinna theory, it is less interesting from the point of view of the analogy with number theory, since number rings are one-dimensional. Moreover, in Nevanlinna theory, the proofs that correspond most closely to proofs in number theory concern maps with domain $\mathbb{C}$.

There is one other way to change the domain of a holomorphic function, though, which is highly relevant to comparisons with number theory. Namely, one can replace the domain $\mathbb{C}$ with a ramified cover. Let B be a connected Riemann surface, let $\pi\colon B \to \mathbb{C}$ be a proper surjective holomorphic map, and let $f\colon B \to \mathbb{C}$ be a meromorphic function. In place of $\overline{\mathbb{D}}_r$ in the above discussion, one can work with $\pi^{-1}\left(\overline{\mathbb{D}}_r\right)$ and define the proximity, counting, and height functions accordingly. For detailed definitions, see Sect. 27.

When working with a finite ramified covering, though, the Second Main Theorem requires an additional term $N_{\mathrm{Ram}(\pi)}(r)$, which is a counting function for ramification points of π (Definition 27.3c). The main inequality (23) of the Second Main Theorem then becomes

$$\sum_{j=1}^{q} m_f(a_j,r) \le_{\text{exc}} 2T_f(r) + N_{\text{Ram}(\pi)}(r) + O(\log^+ T_f(r)) + o(\log r)$$

in this context.

In number theory, the corresponding situation involves algebraic numbers of bounded degree over k instead of elements of k itself. Again, the inequality in the Second Main Theorem becomes weaker in this case, conjecturally by adding the following term.

Definition 7.1. Let D_k denote the discriminant of a number field k, and for number fields $L \supseteq k$ define

$$d_k(L) = \frac{1}{[L:k]} \log|D_L| - \log|D_k|.$$

For $x \in \bar{k}$ we then define

$$d_k(x) = d_k(k(x)).$$

It is then conjectured that Roth's theorem for $x \in \bar{k}$ of bounded degree over k still holds, with inequality

$$\sum_{j=1}^{q} m_S(a_j,x) \le (2+\varepsilon)h_k(x) + d_k(x) + C. \tag{33}$$

For further discussion of this situation, including its relation to the abc conjecture, see Sects. 25–26.

8 Cartan's Theorem and Schmidt's Subspace Theorem

In both Nevanlinna theory and number theory, the first extensions of the Second Main Theorem and its counterpart to higher dimensions were theorems involving approximation to hyperplanes in projective space.

We start with a definition needed for both theorems.

Definition 8.1. A collection of hyperplanes in $\mathbb{P}^n$ is in **general position** if for all $j \le n$ the intersection of any j of them has dimension $n-j$, and if the intersection of any $n+1$ of them is empty.

The Second Main Theorem for approximation to hyperplanes in $\mathbb{P}^n$ was first proved by Cartan [11]. Before stating it, we need to define the proximity, counting, and height functions.

Definition 8.2. Let H be a hyperplane in $\mathbb{P}^n(\mathbb{C})$ $(n > 0)$, and let $a_0x_0 + \cdots + a_nx_n$ be a linear form defining it. Let $P \in \mathbb{P}^n \setminus H$ be a point, and let $[x_0 : \ldots : x_n]$ be homogeneous coordinates for P. We then define

$$\lambda_H(P) = -\frac{1}{2}\log\frac{|a_0x_0+\cdots+a_nx_n|^2}{|x_0|^2+\cdots+|x_n|^2} \tag{34}$$

(this depends on $a_0,\dots,a_n$, but only up to an additive constant). It is independent of the choice of homogeneous coordinates for P.

If $n = 1$ and H is a finite number $a \in \mathbb{C}$ (via the usual identification of $\mathbb{C}$ as a subset of $\mathbb{P}^1(\mathbb{C})$), then

$$\lambda_H(x) = \log^+\left|\frac{1}{x-a}\right| + O(1). \tag{35}$$

Recall that a **holomorphic curve** in a complex variety X is a holomorphic function from $\mathbb{C}$ to $X(\mathbb{C})$.

Definition 8.3. Let n, H, and λ_H be as in Definition 8.2, and let $f\colon \mathbb{C} \to \mathbb{P}^n$ be a holomorphic curve whose image is not contained in H. Then the **proximity function** for H is

$$m_f(H,r) = \int_0^{2\pi} \lambda_H(f(re^{i\theta}))\,\frac{d\theta}{2\pi}. \tag{36}$$

For the following, recall that an **analytic divisor** on $\mathbb{C}$ is a formal sum

$$\sum_{z\in\mathbb{C}} n_z \cdot z,$$

where $n_z \in \mathbb{Z}$ for all z and the set $\{z \in \mathbb{C} : n_z \neq 0\}$ is a discrete set (which may be infinite).

Definition 8.4. Let n, H, and f be as above. Then f^*H is an analytic divisor on $\mathbb{C}$, and for $z \in \mathbb{C}$ we let $\operatorname{ord}_z f^*H$ denote its multiplicity at the point z. Then the **counting function** for H is defined to be

$$N_f(H,r) = \sum_{0<|z|<r} \operatorname{ord}_z f^*H \cdot \log\frac{r}{|z|} + \operatorname{ord}_0 f^*H \cdot \log r \tag{37}$$

Definition 8.5. Let $f\colon \mathbb{C} \to \mathbb{P}^n(\mathbb{C})$ be a holomorphic curve ($n > 0$). We then define the **height** of f to be

$$T_f(r) = m_f(H,r) + N_f(H,r)$$

for any hyperplane H not containing the image of f. The First Main Theorem can be shown to hold in the context of hyperplanes in projective space, so the height depends on H only up to $O(1)$.

We may now state Cartan's theorem.

Theorem 8.6. (Cartan) *Let $n > 0$ and let $H_1,\dots,H_q$ be hyperplanes in $\mathbb{P}^n$ in general position. Let $f\colon \mathbb{C} \to \mathbb{P}^n(\mathbb{C})$ be a holomorphic curve whose image is not contained in any hyperplane. Then*

$$\sum_{j=1}^{q} m_f(H_j, r) \le_{\text{exc}} (n+1)T_f(r) + O(\log^+ T_f(r)) + o(\log r). \tag{38}$$

If $n = 1$ then by (35) this reduces to the classical Second Main Theorem (Theorem 5.5).

Inequality (38) can also be expressed using counting functions as

$$\sum_{j=1}^{q} N_f(H_j, r) \ge_{\text{exc}} (q-n-1)T_f(r) - O(\log^+ T_f(r)) - o(\log r) \tag{39}$$

(cf. (24)).

The corresponding definitions and theorem in number theory are as follows. These will all assume that k is a number field, that $S \supseteq S_\infty$ is a finite set of places of k, and that $n > 0$.

Definition 8.7. Let H be a hyperplane in $\mathbb{P}^n_k$ and let $a_0x_0 + \cdots + a_nx_n = 0$ be a linear form defining it. (Since $\mathbb{P}^n_k$ is a scheme over k, this implies that $a_0, \dots, a_n \in k$.) For all places v of k and all $P \in \mathbb{P}^n(k)$ not lying on H we then define

$$\lambda_{H,v}(P) = -\log \frac{\|a_0x_0 + \cdots + a_nx_n\|_v}{\max\{\|x_0\|_v, \dots, \|x_n\|_v\}}, \tag{40}$$

where $[x_0 : \dots : x_n]$ are homogeneous coordinates for P. Again, this is independent of the choice of homogeneous coordinates $[x_0 : \dots : x_n]$ and depends on the choice of $a_0, \dots, a_n$ only up to a bounded function which is zero for almost all v.

These functions are special cases of *Weil functions* (Definition 9.6), with domain restricted to $\mathbb{P}^n(k) \setminus H$.

Definition 8.8. For H and P as above, the **proximity function** for H is defined to be

$$m_S(H, P) = \sum_{v \in S} \lambda_{H,v}(P), \tag{41}$$

and the **counting function** is defined by

$$N_S(H, P) = \sum_{v \notin S} \lambda_{H,v}(P). \tag{42}$$

We then note that

$$\begin{aligned} m_S(H,P) + N_S(H,P) &= \sum_{v \in M_k} -\log \frac{\|a_0x_0 + \cdots + a_nx_n\|_v}{\max\{\|x_0\|_v, \dots, \|x_n\|_v\}} \\ &= \sum_{v \in M_k} \log \max\{\|x_0\|_v, \dots, \|x_n\|_v\} \\ &= h_k(P) \end{aligned}$$

by the Product Formula.

Although this equality holds exactly, the proximity and counting functions depend (up to $O(1)$) on the choice of linear form $a_0x_0 + \cdots + a_nx_n$ describing D, so we regard them as being defined only up to $O(1)$.

The counterpart to Theorem 8.6 (with, of course, a weaker error term) is a slightly weaker form of Schmidt's Subspace Theorem.

Theorem 8.9. (Schmidt) *Let k, S, and n be as above, let $H_1,\dots,H_q$ be hyperplanes in $\mathbb{P}^n_k$ in general position, let $\varepsilon > 0$, and let $c \in \mathbb{R}$. Then*

$$\sum_{j=1}^{q} m_S(H_j,x) \le (n+1+\varepsilon)h_k(x) + c \tag{43}$$

for all $x \in \mathbb{P}^n(k)$ outside of a finite union of proper linear subspaces of $\mathbb{P}^n_k$. This latter set depends on k, S, $H_1,\dots,H_q$, ε, c, and the choices used in defining the $m_S(H_j,x)$, but not on x.

When $n = 1$ this reduces to Roth's theorem (in the form of Theorem 4.4).

Note, in particular, that the H_i are hyperplanes in the k-scheme $\mathbb{P}^n_k$. This automatically implies that they can be defined by linear forms with coefficients in k. This corresponds to requiring the α_j to lie in k in the case of Roth's theorem. Schmidt's original formulation of his theorem allowed hyperplanes with algebraic coefficients; the reduction to hyperplanes in $\mathbb{P}^n_k$ is similar to the reduction for Roth's theorem and is omitted here. Also, Schmidt's original formulation was stated in terms of hyperplanes in $\mathbb{A}^{n+1}_k$ passing through the origin and points in $\mathbb{A}^{n+1}_k$ with integral coefficients. He also used the *size* instead of the height. For details on the equivalence of his original formulation and the form given here, see [87, Chap. 2, Sect. 2].

Theorem 8.9 is described as a slight weakening of Schmidt's Subspace Theorem because Schmidt actually allowed the set of hyperplanes to vary with v. Thus, to get a statement that was fully equivalent to Schmidt's original theorem, (43) would need to be replaced by

$$\sum_{v\in S}\sum_{j=1}^{q_v} m_S(H_{v,j},x) \le (n+1+\varepsilon)h_k(x) + c,$$

where for each $v \in S$, $H_{v,1},\dots,H_{v,q_v}$ are hyperplanes in general position (but in totality the set $\{H_{v,j} : v \in S,\ 1 \le j \le q_v\}$ need not be in general position, even after eliminating duplicates). Of course, at a given place v a point can be close to at most n of the $H_{v,j}$, so we may assume $q_v = n$ for all v (or actually $n+1$ is somewhat easier to work with).

Thus, a full statement of Schmidt's Subspace Theorem, rendered using the notation of Sect. 6, is as follows. It has been stated in a form that most readily carries over to Nevanlinna theory.

Theorem 8.10. (Schmidt's Subspace Theorem [71, Chap. VIII, Thm. 7A]) *Let k, S, and n be as above, and let $H_1,\dots,H_q$ be distinct hyperplanes in $\mathbb{P}^n_k$. Then for all $\varepsilon > 0$ and all $c \in \mathbb{R}$ the inequality*

$$\sum_{v\in S}\max_J \sum_{j\in J}\lambda_{H_j,v}(x) \le (n+1+\varepsilon)h_k(x)+c \tag{44}$$

holds for all $x \in \mathbb{P}^n(k)$ outside of a finite union of proper linear subspaces depending only on k, S, $H_1,\dots,H_q$, ε, c, and the choices used in defining the $\lambda_{H_j,v}$. The max in this inequality is taken over all subsets J of $\{1,\dots,q\}$ corresponding to subsets of $\{H_1,\dots,H_q\}$ in general position.

In Nevanlinna theory there are infinitely many angles θ, so if one allowed the collection of hyperplanes to vary with θ without additional restriction, then the resulting statement could involve infinitely many hyperplanes, and would therefore likely be false (although this has not been proved). Therefore an overall restriction on the set of hyperplanes is needed in the case of Cartan's theorem, and is why Theorem 8.10 was stated in the way that it was.

Cartan's theorem itself can be generalized as follows.

Theorem 8.11. [90] *Let $n \in \mathbb{Z}_{>0}$, let $H_1,\dots,H_q$ be hyperplanes in $\mathbb{P}^n_{\mathbb{C}}$, and let $f\colon \mathbb{C} \to \mathbb{P}^n(\mathbb{C})$ be a holomorphic curve whose image is not contained in a hyperplane. Then*

$$\int_0^{2\pi} \max_J \sum_{j\in J}\lambda_{H_j}(f(re^{i\theta}))\,\frac{d\theta}{2\pi} \le_{\text{exc}} (n+1)T_f(r)+O(\log^+ T_f(r))+o(\log r),$$

where J varies over the same collection of sets as in Theorem 8.10.

This has proved to be a useful formulation for applications; see [90] and [68]. The latter reference also improves the error term in Theorem 8.11.

Remark 8.12. It has been further shown that in Theorem 8.10, the finite set of linear subspaces can be taken to be the union of a finite number of points (depending on the same data as given in the theorem), together with a finite union of linear subspaces (of higher dimension) depending only on the collection of hyperplanes [88]. In other words, the higher-dimensional part of the exceptional set depends only on the geometric data. Correspondingly, Theorem 8.11 holds for all nonconstant holomorphic curves whose image is not contained in the union of this latter set [90]. For an example of the collection of higher dimensional subspaces for a specific set of lines in $\mathbb{P}^2$, see Example 14.3.

9 Varieties and Weil Functions

The goal of this section and the next is to carry over the definitions of the proximity, counting, and height functions to the context of varieties.

First it is necessary to define variety. Generally speaking, varieties and other algebro-geometric objects are as defined in [36], except that varieties (when discussing number theory at least) may be defined over a field that is not necessarily algebraically closed.

Definition 9.1. A **variety** over a field k, or a k-variety, is an integral separated scheme of finite type over k (i.e., over $\operatorname{Spec} k$). A **curve** over k is a variety over k of dimension 1. A **morphism** of varieties over k is a morphism of k-schemes. Finally, a **subvariety** of a variety (resp. closed subvariety, open subvariety) is an integral subscheme (resp. closed integral subscheme, open integral subscheme) of that variety (with induced map to $\operatorname{Spec} k$).

As an example, $X := \operatorname{Spec} \mathbb{Q}[x,y]/(y^2 - 2x^2)$ is a variety over $\mathbb{Q}$. Indeed, it is an integral scheme because the ring $\mathbb{Q}[x,y]/(y^2 - 2x^2)$ is entire. However, $X \times_{\mathbb{Q}} \mathbb{Q}(\sqrt{2})$ is not a variety over $\mathbb{Q}(\sqrt{2})$, since $\mathbb{Q}[x,y]/(y^2 - 2x^2) \otimes_{\mathbb{Q}} \mathbb{Q}(\sqrt{2})$ is not entire (the polynomial $y^2 - 2x^2$ is not irreducible over $\mathbb{Q}(\sqrt{2})$). Therefore, some authors require a variety to be *geometrically integral,* but we do not do so here. The advantage of not requiring geometric integrality is that every reduced closed subset is a finite union of closed subvarieties, without requiring base change to a larger field.

Many people would be tempted to say that the variety $X := \operatorname{Spec} \mathbb{Q}[x,y]/(y^2 - 2x^2)$ is not *defined over* $\mathbb{Q}$. Such wording does not make sense in this context (the variety is, after all, a $\mathbb{Q}$-variety). This wording usually comes about because the variety (in this instance) is associated to the line $y = \sqrt{2}x$ in $\mathbb{A}^2_{\overline{\mathbb{Q}}}$, which does not come from any subvariety of $\mathbb{A}^2_{\mathbb{Q}}$ (without also obtaining the conjugate $y = -\sqrt{2}x$). The correct way to express this situation is to say that X is not *geometrically irreducible* (or not geometrically integral).

Strictly speaking, if $k \subseteq L$ are distinct fields, then $X(k)$ and $X(L)$ are disjoint sets. However, we will at times identify $X(k)$ with a subset of $X(L)$ in the obvious way. Following EGA, if $x \in X$ is a point, then $\kappa(x)$ will denote the residue field of the local ring at x. If $x \in X(L)$, then it is technically a morphism, but by abuse of notation $\kappa(x)$ will refer to the corresponding point on X (so $\kappa(x)$ may be smaller than L).

We also recall that the function field of a variety X is denoted $K(X)$. If ξ is the generic point of X, then $K(X) = \kappa(\xi)$.

The next goal of this section is to introduce Weil functions. These functions were introduced in Weil's thesis [98] and further developed in a later paper [99]. Weil functions give a way to write the height as a sum over places of a number field, and are exactly what is needed in order to generalize the proximity and counting functions to the geometric setting.

The description provided here will be somewhat brief; for a fuller treatment, see [46, Chap. 10].

We start with the very easy setting used in Nevanlinna theory.

Weil functions are best described using Cartier divisors.

Definition 9.2. Let D be a Cartier divisor on a complex variety X. Then a **Weil function** for D is a function $\lambda_D\colon (X \setminus \operatorname{Supp} D)(\mathbb{C}) \to \mathbb{R}$ such that for all $x \in X$

there is an open neighborhood U of x in X, a nonzero function $f \in K(X)$ such that $D\big|_U = (f)$, and a continuous function $\alpha\colon U(\mathbb{C}) \to \mathbb{R}$ such that

$$\lambda_D(x) = -\log|f(x)| + \alpha(x) \tag{45}$$

for all $x \in (U \setminus \operatorname{Supp} D)(\mathbb{C})$. Here the topology on $U(\mathbb{C})$ is the complex topology.

It is fairly easy to show that if λ_D is a Weil function, then the above condition is satisfied for any open set U and any nonzero $f \in \mathscr{O}_U$ satisfying $D\big|_U = (f)$.

Recall that linear equivalence classes of Cartier divisors on a variety are in natural one-to-one correspondence with isomorphism classes of line sheaves (invertible sheaves) on that variety. Moreover, for each divisor D on a variety X, if $\mathscr{L}$ is the corresponding line sheaf, then there is a nonzero rational section s of $\mathscr{L}$ whose vanishing describes D: $D = (s)$. As was noted by Néron, Weil functions on D correspond to metrics on $\mathscr{L}$.

Recall that if X is a complex variety and $\mathscr{L}$ is a line sheaf on X, then a **metric** on $\mathscr{L}$ is a collection of norms on the fibers of the complex line bundle corresponding to the sheaf $\mathscr{L}$, varying smoothly or continuously with the point on X. Such a metric is called a **smooth metric** or **continuous metric**, respectively. In these notes, **smooth** means C^∞. If X is singular, then we say that a function $f\colon X(\mathbb{C}) \to \mathbb{C}$ is C^∞ at a point $P \in X(\mathbb{C})$ if there is an open neighborhood U of P in $X(\mathbb{C})$ in the complex topology, a holomorphic function $\phi\colon U \to \mathbb{C}^n$ for some n, and a C^∞ function $g\colon \mathbb{C}^n \to \mathbb{C}$ such that $f = g \circ \phi$. This reduces to the usual concept of C^∞ function at smooth points of X.

To describe a metric on $\mathscr{L}$ in concrete terms, let U be an open subset of X and let $\phi_U\colon \mathscr{O}_U \xrightarrow{\sim} \mathscr{L}\big|_U$ be a local trivialization. Then the function $\rho_U\colon U(\mathbb{C}) \to \mathbb{R}_{>0}$ given by $\rho_U(x) = |\phi_U(1)(x)|$ is smooth (resp. continuous), and for any section $s \in \mathscr{L}(U)$ and any $x \in U(\mathbb{C})$, we have $|s(x)| = \rho_U(x) \cdot |\phi_U^{-1}(s)(x)|$. Moreover, if V is another open set in X and $\phi_V\colon \mathscr{O}_V \xrightarrow{\sim} \mathscr{L}\big|_V$ is a local trivialization on V, then $\phi_U^{-1} \circ \phi_V$ (appropriately restricted) is an automorphism of $\mathscr{O}_{U\cap V}$ corresponding to multiplication by a function $\alpha_{UV} \in \mathscr{O}_{U\cap V}^*$. Again letting $\rho_V(x) = |\phi_V(1)(x)|$, we see that ρ_U and ρ_V are related by $\rho_V(x) = |\alpha_{UV}(x)|\rho_U(x)$ for all $x \in (U \cap V)(\mathbb{C})$.

Conversely, an isomorphism class of line sheaves on X can be uniquely specified by giving an open cover $\mathscr{U}$ of X and $\alpha_{UV} \in \mathscr{O}_{U\cap V}^*$ for all $U, V \in \mathscr{U}$ satisfying $\alpha_{UU} = 1$ and $\alpha_{UW} = \alpha_{UV}\alpha_{VW}$ on $U \cap V \cap W$ for all $U, V, W \in \mathscr{U}$. Moreover, with these data, one can specify a metric on the associated line sheaf by giving smooth or continuous functions $\rho_U\colon U(\mathbb{C}) \to \mathbb{R}_{>0}$ for each $U \in \mathscr{U}$ that satisfy $\rho_V = |\alpha_{UV}|\rho_U$ on $U \cap V$ for all $U, V \in \mathscr{U}$.

A continuous metric on a line sheaf $\mathscr{L}$ determines a Weil function for any associated Cartier divisor D. Indeed, if s is a nonzero rational section of $\mathscr{L}$ such that $D = (s)$, then $\lambda_D(x) = -\log|s(x)|$ is a Weil function for D. Conversely, a Weil function for D determines a continuous metric on $\mathscr{L}$.

In Nevanlinna theory it is customary to work only with smooth metrics, and hence it is often better to work with Weil functions associated to smooth metrics (equivalently, to Weil functions for which the functions α in (45) are all smooth).

An example of a Weil function in Nevanlinna theory (and perhaps the primary example) is the function λ_H of Definition 8.2 used in Cartan's theorem.

Likewise, the function $\lambda_{H,v}$ of Definition 8.7 is an example of a Weil function in number theory. In this case, it is no longer sufficient to say that two Weil functions agree up to $O(1)$: the implied constant also has to vanish for almost all v. For example, Lemma 6.2 compares the difference of two Weil functions, and shows that the difference is bounded by a bound that vanishes for almost all v. A plain bound of $O(1)$ would not suffice to give a finite bound in Theorem 6.3.

Before defining Weil functions in the number theory case, we first give some definitions relevant to the domains of Weil functions.

Definition 9.3. Let v be a place of a number field k. Then $\mathbb{C}_v$ is the completion of the algebraic closure $\bar{k}_v$ of the completion k_v of k at v.

Recall [42, Chap. III, Sects. 3 and 4] that if v is non-archimedean then $\bar{k}_v$ is not complete, but its completion $\mathbb{C}_v$ is algebraically closed. If v is archimedean, then $\mathbb{C}_v$ is isomorphic to the field of complex numbers (as is $\bar{k}_v$). The norm $\|\cdot\|_v$ on k extends uniquely to norms on k_v, on $\bar{k}_v$, and on $\mathbb{C}_v$. If X is a variety, then the norm on $\mathbb{C}_v$ defines a topology on $X(\mathbb{C}_v)$, called the v-topology. It is defined to be the coarsest topology such that for all open $U \subseteq X$ and all $f \in \mathscr{O}(U)$, $U(\mathbb{C}_v)$ is open and $f \colon U(\mathbb{C}_v) \to \mathbb{C}_v$ is continuous.

One can also work just with the algebraic closure $\bar{k}_v$ when defining Weil functions, without any essential difference.

Definition 9.4. Let X be a variety over a number field k. Then $X(M_k)$ is the disjoint union

$$X(M_k) = \coprod_{v \in M_k} X(\mathbb{C}_v).$$

This set is given a topology defined by the condition that $A \subseteq X(M_k)$ is open if and only if $A \cap X(\mathbb{C}_v)$ is open in the v-topology for all v.

Definition 9.5. Let k be a number field. Then an **M_k-constant** is a collection (c_v) of constants $c_v \in \mathbb{R}$ for each $v \in M_k$, such that $c_v = 0$ for almost all v. If X is a variety over k, then a function $\alpha \colon X(M_k) \to \mathbb{R}$ is said to be $O_{M_k}(1)$ if there is an M_k-constant (c_v) such that $|\alpha(x)| \le c_v$ for all $x \in X(\mathbb{C}_v)$ and all $v \in M_k$.

We may then define Weil functions as follows.

Definition 9.6. Let X be a variety over a number field k, and let D be a Cartier divisor on X. Then a **Weil function** for D is a function $\lambda_D \colon (X \setminus \operatorname{Supp} D)(M_k) \to \mathbb{R}$ that satisfies the following condition. For each $x \in X$ there is an open neighborhood U of x, a nonzero function $f \in \mathscr{O}(U)$ such that $D\big|_U = (f)$, and a continuous locally M_k-bounded function $\alpha \colon U(M_k) \to \mathbb{R}$ satisfying

$$\lambda_D(x) = -\log \|f(x)\|_v + \alpha(x) \tag{46}$$

for all $v \in M_k$ and all $x \in (U \setminus \operatorname{Supp} D)(\mathbb{C}_v)$.

For the definition of locally M_k-bounded function, see [46, Chap. 10, Sect. 1]. The definition is more complicated than one would naively expect, stemming from the fact that $\mathbb{C}_v$ is totally disconnected, and not locally compact. For our purposes, though, it suffices to note that if X is a complete variety then such a function is $O_{M_k}(1)$. (In other contexts, these problems are dealt with by using Berkovich spaces, but Weil's work does not use them, not least because it came much earlier.)

As with Definition 9.2, if λ is a Weil function for D, then it can be shown that the above condition is true for all open $U \subseteq X$ and all $f \in \mathscr{O}(U)$ for which $D\big|_U = (f)$.

If λ_D is a Weil function for D, then we write

$$\lambda_{D,v} = \lambda_D\big|_{(X\setminus\operatorname{Supp} D)(\mathbb{C}_v)}$$

for all places v of k. If v is an archimedean place, then $\mathbb{C}_v \cong \mathbb{C}$, and $\lambda_{D,v}$ is a Weil function for D in the sense of Definition 9.2(up to a factor $1/2$ if v is a complex place).

In the future, if $x \in X(M_k)$ and f is a function on X, then $\|f(x)\|$ will mean $\|f(x)\|_v$ for the (unique) place v such that $x \in X(\mathbb{C}_v)$. Thus, (46) could be shortened to $\lambda(x) = -\log\|f(x)\| + \alpha(x)$ for all $x \in (U \setminus \operatorname{Supp} D)(M_k)$.

Of course, this discussion would be academic without the following theorem.

Theorem 9.7. *Let k be a number field, let X be a projective variety over k, and let D be a Cartier divisor on X. Then there exists a Weil function for D.*

For the proof, see [46, Chap. 10]. This is also true for complete varieties, using Nagata's embedding theorem to construct a model for X and then using Arakelov theory to define the Weil function. But, again, the details are beyond the scope of these notes.

Weil functions have the following properties.

Theorem 9.8. *Let X be a complete variety over a number field k. Then*

(a) ***Additivity:*** *If λ_1 and λ_2 are Weil functions for Cartier divisors D_1 and D_2 on X, respectively, then $\lambda_1 + \lambda_2$ 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 X, and if $f\colon X' \to X$ is a morphism of k-varieties such that $f(X') \not\subseteq \operatorname{Supp} D$, then $x \mapsto \lambda(f(x))$ is a Weil function for the Cartier divisor f^*D on X'.*

(c) ***Normalization:*** *If $X = \mathbb{P}^n_k$, and if D is the hyperplane at infinity, then the function*

$$\lambda_{D,v}([x_0 : \dots : x_n]) := -\log \frac{\|x_0\|_v}{\max\{\|x_0\|_v, \dots, \|x_n\|_v\}} \tag{47}$$

is a Weil function for D.

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

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

(f) ***Principal divisors:*** *If D is a principal divisor (f), then $-\log\|f\|$ is a Weil function for D.*

The proofs of these properties are left to the reader (modulo the properties of locally M_k-bounded functions).

Parts (b) and (c) of the above theorem combine to give a way of computing Weil functions for effective very ample divisors. This, in turn, gives rise to the "max-min" method for computing Weil functions for arbitrary Cartier divisors on projective varieties.

Lemma 9.9. *Let $\lambda_1,\dots,\lambda_n$ be Weil functions for Cartier divisors $D_1,\dots,D_n$, respectively, on a complete variety X over a number field k. Assume that the divisors D_i are of the form $D_i = D_0 + E_i$, where D_0 is a fixed Cartier divisor and E_i are effective for all i. Assume also that* $\operatorname{Supp} E_1 \cap \cdots \cap \operatorname{Supp} E_n = \emptyset$. *Then the function*

$$\lambda(x) = \min\{\lambda_i(x) : x \notin \operatorname{Supp} E_i\}$$

is defined everywhere on $(X \setminus \operatorname{Supp} D_0)(M_k)$, and is a Weil function for D_0.

Proof. See [46, Chap. 10, Prop. 3.2]. □

Theorem 9.10. (Max-min) *Let X be a projective variety over a number field k, and let D be a Cartier divisor on X. Then there are positive integers m and n, and nonzero rational functions f_{ij} on X, $1 = 1,\dots,n$, $j = 1,\dots,m$, such that*

$$\lambda(x) := \max_{1\le i\le n} \min_{1\le j\le m} \bigl(-\log \|f_{ij}\|\bigr)$$

defines a Weil function for D.

Proof. We may write D as a difference $E - F$ of very ample divisors. Let $E_1,\dots,E_n$ be effective Cartier divisors linearly equivalent to E such that $\bigcap \operatorname{Supp} E_i = \emptyset$ (for example, pull-backs of hyperplane sections via a projective embedding associated to E). Likewise, let $F_1,\dots,F_m$ be effective Cartier divisors linearly equivalent to F with $\bigcap \operatorname{Supp} F_j = \emptyset$. Then $D - E_i + F_j$ is a principal divisor for all i and j; hence

$$D - E_i + F_j = (f_{ij})$$

for some $f_{ij} \in K(X)^*$ and all i and j. Applying Lemma 9.9 to $-\log\|f_{ij}\|$ then implies that $\min_{1\le j\le m}\bigl(-\log\|f_{ij}\|\bigr)$ is a Weil function for $D - E_i$ for all i. Applying Lemma 9.9 again to the negatives of these Weil functions then gives the theorem. □

To conclude the section, we give some notation that will be useful for working with rational and algebraic points.

Definition 9.11. Let X be a variety over a number field k, let D be a Cartier divisor on X, and let λ_D be a Weil function for D. If L is a number field containing k, and if w is a place of L lying over a place v of k, then we identify $\mathbb{C}_w$ with $\mathbb{C}_v$ in the obvious manner, and write

$$\lambda_{D,w} = [L_w : k_v]\lambda_{D,v}. \tag{48}$$

(Recall that $\|x\|_w = \|x\|_v^{[L_w:k_v]}$ for all $x \in \mathbb{C}_v$, by (4).) Finally, each point $x \in X(L)$ gives rise to points $x_w \in X(\mathbb{C}_w)$ for all $w \in M_L$, and we define

$$\lambda_{D,w}(x) = \lambda_{D,w}(x_w) \tag{49}$$

if $x \notin \operatorname{Supp} D$.

Note that, if $x \in (X \setminus \operatorname{Supp} D)(L)$, if L' is a number field containing L, if w is a place of L, and if w' is a place of L' lying over w, then

$$\lambda_{D,w'}(x) = [L'_{w'} : L_w]\lambda_{D,w}(x), \tag{50}$$

regardless of whether the left-hand side is defined using (48) or (49) (by regarding $X(L)$ as a subset of $X(L')$ for the latter).

If (c_v) is an M_k-constant, if w is a place of a number field L containing k, and if v is the place of k lying under w, then we write

$$c_w = [L_w : k_v]c_v, \tag{51}$$

so that the condition $\lambda_{D,w} \le c_w$ will be equivalent to $\lambda_{D,v} \le c_v$, by (48).

10 Height Functions on Varieties in Number Theory

Weil functions can be used to generalize the height h_k (defined in Sect. 3) to arbitrary complete varieties over k. This can also be done by working directly with heights; see [46, Chap. 3] or [77].

Throughout this section, k is a number field, X is a complete variety over k, and D is a Cartier divisor on X, unless otherwise specified.

Definition 10.1. Let λ be a Weil function for D, and let $x \in X(\bar{k})$ be an algebraic point with $x \notin \operatorname{Supp} D$. Then the **height** of x relative to λ and k is defined as

$$h_{\lambda,k}(x) = \frac{1}{[L:k]} \sum_{w \in M_L} \lambda_w(x) \tag{52}$$

for any number field $L \supseteq \kappa(x)$. It is independent of the choice of L by (50).

In particular, if $x \in X(k)$, then

$$h_{\lambda,k}(x) = \sum_{v \in M_k} \lambda_v(x).$$

Specializing in a different direction, if $X = \mathbb{P}^n_k$, if D is the hyperplane at infinity, and if λ is the Weil function (47), then

$$\begin{aligned} h_{\lambda,k}([x_0 : \ldots : x_n]) &= -\frac{1}{[L:k]} \sum_{w \in M_L} \log \frac{\|x_0\|_v}{\max\{\|x_0\|_v, \ldots, \|x_n\|_v\}} \\ &= \frac{1}{[L:k]} \sum_{w \in M_L} \log \max\{\|x_0\|_v, \ldots, \|x_n\|_v\} \\ &= h_k([x_0 : \ldots : x_n]) \end{aligned} \tag{53}$$

for all $[x_0 : \ldots : x_n] \in \mathbb{P}^n(\bar{k})$ with $x_0 \neq 0$, where L is any number field containing the field of definition of this point.

The restriction $x \notin \operatorname{Supp} D$ can be eliminated as follows.

Let D' be another Cartier divisor on X linearly equivalent to D, say $D' = D + (f)$; then $\lambda' := \lambda - \log\|f\|$ is a Weil function for D'. If $x \in X(\bar{k})$ does not lie on $\operatorname{Supp} D \cup \operatorname{Supp} D'$, and if L is a number field containing $\kappa(x)$, then

$$\begin{aligned} h_{\lambda',k}(x) &= \frac{1}{[L:k]} \sum_{w \in M_L} \lambda'_v(x) \\ &= \frac{1}{[L:k]} \sum_{w \in M_L} \lambda_v(x) - \frac{1}{[L:k]} \sum_{w \in M_L} \log\|f(x)\|_w \\ &= h_{\lambda,k}(x) \end{aligned} \tag{54}$$

by the Product Formula (6). Thus we have:

Definition 10.2. Let λ be a Weil function for D, and let $x \in X(\bar{k})$. Then, for any $f \in K(X)^*$ such that the support of $D + (f)$ does not contain x, we define

$$h_{\lambda,k}(x) = h_{\lambda - \log\|f\|,k}(x),$$

where $h_{\lambda - \log\|f\|,k}$ on the right-hand side is defined using Definition 10.1. This expression is independent of the choice of f by (54), and agrees with Definition 10.1 when $x \notin \operatorname{Supp} D$ since we can take $f = 1$ in that case.

With this definition, (53) holds without the restriction $x_0 \neq 0$.

Proposition 10.3. *If both λ and λ' are Weil functions for D, then*

$$h_{\lambda',k} = h_{\lambda,k} + O(1).$$

Proof. Indeed, this is immediate from Theorem 9.8d. □

Thus, the height function defined above depends only on the divisor; moreover, by (54) it depends only on the linear equivalence class of the divisor.

Definition 10.4. The height $h_{D,k}(x)$ for points $x \in X(\bar{k})$ is defined, up to $O(1)$, as

$$h_{D,k}(x) = h_{\lambda,k}(x)$$

for any Weil function λ for D. If $\mathscr{L}$ is a line sheaf on X, then we define

$$h_{\mathscr{L},k}(x) = h_{D,k}(x)$$

for points $x \in X(\bar{k})$, where D is any Cartier divisor for which $\mathscr{O}(D) \cong \mathscr{L}$. Again, it is only defined up to $O(1)$.

By (53), we then have

$$h_{\mathscr{O}(1),k} = h_k + O(1)$$

on $\mathbb{P}^n_k$ for all $n > 0$. Since the automorphism group of $\mathbb{P}^n_k$ is transitive on the set of rational points, and since automorphisms preserve the line sheaf $\mathscr{O}(1)$, the term $O(1)$ in the above formula cannot be eliminated without additional structure. Thus, Definition 10.4 cannot give an exact definition for the height without additional structure. (This additional structure can be given using Arakelov theory.)

Theorem 9.8 and (50) also immediately imply the following properties of heights:

Theorem 10.5. *(a)* ***Functoriality:*** *If $f\colon X' \to X$ is a morphism of k-varieties, and if $\mathscr{L}$ is a line sheaf on X, then*

$$h_{f^*\mathscr{L},k}(x) = h_{\mathscr{L},k}(f(x)) + O(1)$$

for all $x \in X'(\bar{k})$, where the implied constant depends only on f, $\mathscr{L}$, and the choices of the height functions.

(b) ***Additivity:*** *If $\mathscr{L}_1$ and $\mathscr{L}_2$ are line sheaves on X, then*

$$h_{\mathscr{L}_1\otimes\mathscr{L}_2,k}(x) = h_{\mathscr{L}_1,k}(x) + h_{\mathscr{L}_2,k}(x) + O(1)$$

for all $x \in X(\bar{k})$, where the implied constant depends only on $\mathscr{L}_1$, $\mathscr{L}_2$, and the choices of the height functions.

(c) ***Base locus:*** *If $h_{D,k}$ is a height function for D, then it is bounded from below outside of the base locus of the complete linear system $|D|$.*

(d) ***Globally generated line sheaves:*** *If $\mathscr{L}$ is a line sheaf on X, and is generated by its global sections, then $h_{\mathscr{L},k}(x)$ is bounded from below for all $x \in X(\bar{k})$, by a bound depending only on $\mathscr{L}$ and the choice of height function.*

(e) ***Change of number field:*** *If $L \supseteq k$ then*

$$h_{\mathscr{L},L}(x) = [L:k]h_{\mathscr{L},k}(x)$$

for all line sheaves $\mathscr{L}$ on X and all $x \in X(\bar{k})$. (Strictly speaking, the left-hand side should be $h_{\mathscr{L}',L}(x')$, where $\mathscr{L}'$ is the pull-back of $\mathscr{L}$ to $X_L := X \times_k L$ and x' is the point in $X_L(\bar{k})$ corresponding to $x \in X(\bar{k})$.)

Corollary 10.6. *If $\mathscr{L}$ is an ample line sheaf on X, then $h_{\mathscr{L},k}(x)$ is bounded from below for all $x \in X(\bar{k})$, by a bound depending only on $\mathscr{L}$ and the choice of height function.*

Proof. By Theorem 10.5b, we may replace $\mathscr{L}$ with $\mathscr{L}^{\otimes n}$ for any positive integer n, and therefore may assume that $\mathscr{L}$ is very ample. Then the result follows from Theorem 10.5d. □

The following result shows that heights relative to ample line sheaves are the largest possible heights, up to a constant multiple.

Proposition 10.7. *Let $\mathscr{L}$ and $\mathscr{M}$ be line sheaves on X, with $\mathscr{L}$ ample. Then there is a constant C, depending only on $\mathscr{L}$ and $\mathscr{M}$, such that*

$$h_{\mathscr{M},k}(x) \le C h_{\mathscr{L},k}(x) + O(1)$$

for all $x \in X(\bar{k})$, where the implied constant depends only on $\mathscr{L}$, $\mathscr{M}$, and the choices of height functions.

Proof. By the definition of ampleness, there is an integer n such that the line sheaf $\mathscr{L}^{\otimes n} \otimes \mathscr{M}^{\vee}$ is generated by global sections. Therefore an associated height function

$$h_{\mathscr{L}^{\otimes n} \otimes \mathscr{M}^{\vee},k} = n h_{\mathscr{L},k} - h_{\mathscr{M},k} + O(1)$$

is bounded from below, giving the result with $C = n$. □

For projective varieties, Northcott's finiteness theorem can be carried over.

Theorem 10.8. (Northcott) *Assume that X is projective, and let $\mathscr{L}$ be an ample line sheaf on X. Then, for all integers d and all $c \in \mathbb{R}$, the set*

$$\{x \in X(\bar{k}) : [\kappa(x) : k] \le d \text{ and } h_{\mathscr{L},k}(x) \le c\} \tag{55}$$

is finite.

Proof. First, if $X = \mathbb{P}^n_k$ and $\mathscr{L} = \mathscr{O}(1)$, then the result follows by bounding the heights of $[x_i : x_0]$ (if $x_0 \ne 0$, which we assume without loss of generality), and applying Theorem 3.1 to these points for each i. The general case then follows by replacing $\mathscr{L}$ with a very ample positive multiple and using an associated projective embedding and functoriality of heights. □

Of course, if X is not projective then it has no ample divisors, making the above two statements vacuous. Complete varieties have a notion that is almost as good, though.

Definition 10.9. Let X be a complete variety over an arbitrary field. A line sheaf $\mathscr{L}$ on X is **big** if there is a constant $c > 0$ such that

$$h^0(X, \mathscr{L}^{\otimes n}) \geq cn^{\dim X}$$

for all sufficiently large and divisible n. A Cartier divisor D on X is big if $\mathscr{O}(D)$ is big.

If X is a complete variety over an arbitrary field, then by Chow's lemma there is a projective variety X' and a proper birational morphism $\pi\colon X' \to X$. If $\mathscr{L}$ is a big line sheaf on X, then $\pi^*\mathscr{L}$ is big on X'. Therefore, it makes some sense to compare big line sheaves with ample ones.

Proposition 10.10. (Kodaira's lemma) *Let X be a projective variety over an arbitrary field, and let $\mathscr{L}$ and $\mathscr{A}$ be line sheaves on X, with $\mathscr{A}$ ample. Then $\mathscr{L}$ is big if and only if there is a positive integer n such that $H^0(X, \mathscr{L}^{\otimes n} \otimes \mathscr{A}^\vee) \neq 0$. Equivalently, if D and A are Cartier divisors on X, with A ample, then D is big if and only if some positive multiple of it is linearly equivalent to the sum of A and an effective divisor.*

Proof. See [87, Prop. 1.2.7]. □

The above allows us to show that heights relative to big line sheaves are also, well, big.

Proposition 10.11. *Let X be a complete variety over a number field. Let $\mathscr{L}$ and $\mathscr{M}$ be line sheaves on X, with $\mathscr{L}$ big. Then there is a constant C and a proper Zariski-closed subset Z of X, depending only on $\mathscr{L}$ and $\mathscr{M}$, such that*

$$h_{\mathscr{M},k}(x) \leq C h_{\mathscr{L},k}(x) + O(1)$$

for all $x \in X(\bar{k})$ outside of Z, where the implied constant depends only on $\mathscr{L}$, $\mathscr{M}$, and the choices of height functions.

Proof. After applying Chow's lemma and pulling back $\mathscr{L}$ and $\mathscr{M}$, we may assume that X is projective. We may also replace $\mathscr{L}$ with a positive multiple, and hence may assume that $\mathscr{L}$ is isomorphic to $\mathscr{A} \otimes \mathscr{O}(D)$, where $\mathscr{A}$ is an ample line sheaf and D is an effective Cartier divisor. Then the result follows from Proposition 10.7, with $Z = \operatorname{Supp} D$, by Theorem 10.5. □

Unfortunately, it is still not true that an arbitrary complete variety must have a big line sheaf. But it is true if the variety is nonsingular, since one can then take the complement of any open affine subset.

For general complete varieties, we can do the following.

Remark 10.12. For a general complete variety X over k, we can define a **big height** to be a function $h\colon X(\bar{k}) \to \mathbb{R}$ for which there exist disjoint subvarieties $U_1, \dots, U_n$ of X (not necessarily open or closed), with $\bigcup U_i = X$; and for each $i = 1, \dots, n$ a projective embedding $U_i \hookrightarrow \overline{U}_i$, an ample line sheaf $\mathscr{L}_i$ on $\overline{U}_i$, and real constants $c_i > 0$ and C_i such that $h(x) \geq c_i h_{\mathscr{L}_i,k}(x) + C_i$ for all $x \in U_i(\bar{k})$. One can then show:

- Every complete variety over k has a big height;
- Any two big heights on a given complete variety are bounded from above by linear functions of each other;
- If X is a projective variety and $\mathscr{L}$ is an ample line sheaf on X then $h_{\mathscr{L},k}$ is a big height on X;
- If $\mathscr{L}$ is a line sheaf on X then there are real constants c and C such that $h_{\mathscr{L},k}(x) \le c\,h(x) + C$ for all $x \in X(\bar{k})$;
- The restriction of a big height to a closed subvariety is a big height on that subvariety; and
- A counterpart to Proposition 10.13 (below) holds for big heights on complete varieties.

Details of these assertions are left to the reader.

Big heights are useful for error terms: the conjectures and theorems that follow are generally stated for projective varieties, with an error term involving a height relative to an ample divisor. However, they can also be stated more generally for complete varieties if the height is changed to a big height. For concreteness, though, the more restricted setting of projective varieties is used.

Finally, we note a case in which Z can be bounded explicitly. This will be used in the proof of Proposition 30.3.

Proposition 10.13. *Let $f\colon X_1 \to X_2$ be a morphism of projective varieties over a number field, and let $\mathscr{A}_1$ and $\mathscr{A}_2$ be ample line sheaves on X_1 and X_2, respectively. Then there is a constant C, depending only on f, $\mathscr{A}_1$, and $\mathscr{A}_2$ such that*

$$h_{\mathscr{A}_1,k}(x) \le C\,h_{\mathscr{A}_2,k}(f(x)) + O(1) \tag{56}$$

for all points $x \in X_1(\bar{k})$ that are isolated in their fibers of f, where the implied constant depends only on f, $\mathscr{A}_1$, $\mathscr{A}_2$, and the choices of height functions.

Proof. If no closed points x of X_1 are isolated in their fibers of f, then there is nothing to prove. If there is at least one such point x, then $\dim f(X_1) = \dim X_1$, so $f^*\mathscr{A}_2$ is big. The result then follows by Proposition 10.11 and noetherian induction applied to the irreducible components of the set Z in that proposition □

Note that if any fiber component of f has dimension >0, then it contains algebraic points of arbitrarily large height, so (56) cannot possibly hold for all such points.

11 Proximity and Counting Functions on Varieties in Number Theory

The definitions of proximity and counting functions given in Sects. 6 and 8 also generalize readily to points on varieties.

Throughout this section, k is a number field, S is a finite set of places of k containing S_∞, and X is a complete variety over k.

Definition 11.1. Let D be a Cartier divisor on X, let λ_D be a Weil function for D, let $x \in X(\bar{k})$ with $x \notin \operatorname{Supp} D$, let $L \supseteq k$ be a number field such that $x \in X(L)$, and let T be the set of places of L lying over places in S. Then the **proximity function** and **counting function** in this setting are defined up to $O(1)$ by

$$m_S(D,x) = \frac{1}{[L:k]} \sum_{w \in T} \lambda_{D,w}(x) \qquad \text{and} \qquad N_S(D,x) = \frac{1}{[L:k]} \sum_{w \notin T} \lambda_{D,w}(x).$$

These expressions are independent of the choice of L, by (48). They depend on the choice of λ_D only up to $O(1)$.

Unlike the height, the proximity and counting functions depend on D, even within a linear equivalence class. Therefore the restriction $x \notin \operatorname{Supp} D$ cannot be eliminated.

By (52) and Definition 10.4, we have

$$h_{D,k}(x) = m_S(D,x) + N_S(D,x)$$

for all $x \in X(\bar{k})$ outside of the support of D. This is, basically, the First Main Theorem. The Second Main Theorem in this context is still a conjecture (Conjecture 15.6).

Theorem 9.8 immediately implies the following properties of proximity and counting functions.

Proposition 11.2. *In number theory, proximity and counting functions have the following properties.*

(a) ***Additivity:*** *If D_1 and D_2 are Cartier divisors on X, then*

$$m_S(D_1 + D_2, x) = m_S(D_1, x) + m_S(D_2, x) + O(1)$$

and

$$N_S(D_1 + D_2, x) = N_S(D_1, x) + N_S(D_2, x) + O(1)$$

for all $x \in X(\bar{k})$ outside of the supports of D_1 and D_2.

(b) ***Functoriality:*** *If $f : X' \to X$ is a morphism of complete k-varieties and D is a divisor on X whose support does not contain the image of f, then*

$$m_S(f^*D, x) = m_S(D, f(x)) + O(1)$$

and

$$N_S(f^*D,x) = N_S(D,f(x)) + O(1)$$

for all $x \in X'(\bar{k})$ outside of the support of f^*D.

(c) ***Effective divisors:*** *If D is an effective Cartier divisor on X, then $m_S(D,x)$ and $N_S(D,x)$ are bounded from below for all $x \in X(\bar{k})$ outside of the support of D.*

(d) ***Change of number field:*** *If L is a number field containing k and if T is the set of places of L lying over places in S, then*

$$m_T(D,x) = [L:k]m_S(D,x) + O(1)$$

and

$$N_T(D,x) = [L:k]N_S(D,x) + O(1)$$

for all $x \in X(\bar{k})$ outside of the support of D (with the same abuse of notation as in Theorem 10.5e).

In each of the above cases, the implied constant in $O(1)$ depends on the varieties, divisors, and morphisms, but not on x.

When working with proximity and height functions, the divisor D is almost always assumed to be effective.

12 Height, Proximity, and Counting Functions in Nevanlinna Theory

The height, proximity, and counting functions of Nevanlinna theory can also be generalized to the context of a divisor on a complete complex variety.

In this section, X is a complete complex variety, D is a Cartier divisor on X, and $f : \mathbb{C} \to X$ is a holomorphic curve whose image is not contained in the support of D. Throughout these notes, we will often implicitly think of a complex variety X as a complex analytic space [36, App. B].

We begin with the proximity and counting functions.

Definition 12.1. Let λ be a Weil function for D. Then the **proximity function** for f relative to D is the function

$$m_f(D,r) = \int_0^{2\pi} \lambda(f(re^{i\theta}))\,\frac{d\theta}{2\pi}.$$

It is defined only up to $O(1)$.

If λ' is another Weil function for D, then $|\lambda - \lambda'|$ is bounded, so the proximity function is independent of D (up to $O(1)$).

Definition 12.2. The **counting function** for f relative to D is the function

$$N_f(D,r) = \sum_{0<|z|<r} \operatorname{ord}_z f^*D \cdot \log \frac{r}{|z|} + \operatorname{ord}_0 f^*D \cdot \log r.$$

Unlike the proximity function and the counting function in Nevanlinna theory, this function is defined exactly.

Corresponding to Proposition 11.2, we then have

Proposition 12.3. *In Nevanlinna theory, proximity and counting functions have the following properties.*

(a) ***Additivity:*** *If D_1 and D_2 are Cartier divisors on X, then*

$$m_f(D_1 + D_2, r) = m_f(D_1, r) + m_f(D_2, r) + O(1)$$

and

$$N_f(D_1 + D_2, r) = N_f(D_1, r) + N_f(D_2, r).$$

(b) ***Functoriality:*** *If $\phi : X \to X'$ is a morphism of complete complex varieties and D' is a Cartier divisor on X' whose support does not contain the image of $\phi \circ f$, then*

$$m_f(\phi^*D', r) = m_{\phi\circ f}(D', r) + O(1) \qquad \text{and} \qquad N_f(\phi^*D', r) = N_{\phi\circ f}(D', r).$$

(c) ***Effective divisors:*** *If D is effective, then $m_f(D,r)$ is bounded from below and $N_f(D,r)$ is nonnegative.*

In each of the above cases, the implied constant in $O(1)$ depends on the varieties, divisors, and morphisms, but not on f or r.

We can now define the height.

Definition 12.4. The **height function** relative to D is defined, up to $O(1)$, as

$$T_{D,f}(r) = m_f(D,r) + N_f(D,r).$$

Proposition 12.5. *The height function $T_{D,f}$ is additive in D, is functorial, and is bounded from below if D is effective, as in Proposition 12.3.*

Proof. Immediate from Proposition 12.3. □

Proposition 12.6. (First Main Theorem) *Let D' be another Cartier divisor on X whose support does not contain the image of f, and assume that D' is linearly equivalent to D. Then*

$$T_{D',f}(r) = T_{D,f}(r) + O(1). \tag{57}$$

Proof. We first consider the special case $X = \mathbb{P}^1_{\mathbb{C}}$, $D = [0]$ (the image of the point 0 under the injection $\mathbb{A}^1 \hookrightarrow \mathbb{P}^1$), and $D' = [\infty]$ (the point at infinity, with multiplicity one). Then $T_{D',f}(r) = T_f(r) + O(1)$, $m_f(D,r) = m_f(0,r) + O(1)$, and $N_f(D,r) = N_f(0,r)$ (where $T_f(r)$, $m_f(0,r)$, and $N_f(0,r)$ are as defined in Sect. 5). The result then follows by Theorem 5.4 (the First Main Theorem for meromorphic functions).

In the general case, write $D - D' = (g)$ for some $g \in K(X)^*$. Then g defines a rational map $X \dashrightarrow \mathbb{P}^1_{\mathbb{C}}$. Let X' be the closure of the graph, with projections $p\colon X' \to X$ and $q\colon X' \to \mathbb{P}^1_{\mathbb{C}}$. By the additivity property of heights, (57) is equivalent to $T_{D-D',f}(r)$ being bounded. By the special case proved already, $T_{[0]-[\infty],g\circ f}(r)$ is bounded. The holomorphic curve $f\colon \mathbb{C} \to X$ lifts to a function $f'\colon \mathbb{C} \to X'$ that satisfies $p \circ f' = f$ and $q \circ f' = g \circ f$. By functoriality, we then have

$$\begin{aligned} T_{D-D',f}(r) &= T_{p^*(D-D'),f'}(r) + O(1) \\ &= T_{q^*([0]-[\infty]),f'}(r) + O(1) \\ &= T_{[0]-[\infty],q\circ f'}(r) + O(1) \\ &= T_{[0]-[\infty],g\circ f}(r) + O(1) \\ &= O(1), \end{aligned}$$

which implies the proposition. □

Definition 12.7. The **height function** of f relative to a line sheaf $\mathscr{L}$ on X is defined to be $T_{\mathscr{L},f}(r) = T_{D,f}(r) + O(1)$ for any divisor D such that $\mathscr{O}(D) \cong \mathscr{L}$ and such that the support of D does not contain the image of f. It is defined only up to $O(1)$.

One can obtain a precise height function (defined exactly, not up to $O(1)$), by fixing a Weil function for any such D, or by choosing a metric on $\mathscr{L}$. It is also possible to use a $(1,1)$-form associated to such a metric (the **Ahlfors-Shimizu height**), but this will not be used in these notes.

Continuing on with the development of the height, we have the following counterpart to Theorem 10.5.

Theorem 12.8. *(a)* ***Functoriality:*** *If $\phi\colon X \to X'$ is a morphism of complete complex varieties and if $\mathscr{L}$ is a line sheaf on X', then*

$$T_{\phi^*\mathscr{L},f}(r) = T_{\mathscr{L},\phi\circ f}(r) + O(1).$$

(b) ***Additivity:*** *If $\mathscr{L}_1$ and $\mathscr{L}_2$ are line sheaves on X, then*

$$T_{\mathscr{L}_1\otimes\mathscr{L}_2,f}(r) = T_{\mathscr{L}_1,f}(r) + T_{\mathscr{L}_2,f}(r) + O(1).$$

(c) ***Base locus:*** *If the image of f is not contained in the base locus of the complete linear system $|D|$, then $T_{D,f}(r)$ is bounded from below.*
(d) ***Globally generated line sheaves:*** *If $\mathscr{L}$ is a line sheaf on X, and is generated by its global sections, then $T_{\mathscr{L},f}(r)$ is bounded from below.*

The implicit constants can probably also be made to depend only on the geometric data and the choice of height functions (and not on f), but this is not very important since it is the independence of r that is useful.

The following three results correspond to similar results in the end of Sect. 10.

Corollary 12.9. *If $\mathscr{L}$ is an ample line sheaf on X, then $T_{\mathscr{L},f}(r)$ is bounded from below, is bounded from above if and only if f is constant, and is $O(\log r)$ if and only if f is algebraic.*

Proof. When $X = \mathbb{P}^1$, see [29, Chap. 1, Thm. 6.4] for the second assertion. The general case is left as an exercise for the reader. □

Proposition 12.10. *Let $\mathscr{L}$ and $\mathscr{M}$ be line sheaves on X, with $\mathscr{L}$ ample. Then there is a constant C, depending only on $\mathscr{L}$ and $\mathscr{M}$, such that*

$$T_{\mathscr{M},f}(r) \le C\,T_{\mathscr{L},f}(r) + O(1).$$

Proof. This is true for essentially the same reasons as Proposition 10.7. The details are left to the reader. □

Proposition 12.11. *Let $\mathscr{L}$ and $\mathscr{M}$ be line sheaves on X, with $\mathscr{L}$ big. Then there is a constant C and a proper Zariski-closed subset Z of X, depending only on $\mathscr{L}$ and $\mathscr{M}$, such that*

$$T_{\mathscr{M},f}(r) \le C\,T_{\mathscr{L},f}(r) + O(1),$$

provided that the image of f is not contained in Z.

Proof. Similar to the proof of Proposition 10.11; details are again left to the reader. □

Remark 12.12. For an arbitrary complete variety X over $\mathbb{C}$ and a holomorphic curve $f\colon \mathbb{C} \to X$, one can define a **big height** to be a real-valued function $T_{\mathrm{big},f}(r)$ with the property that if Z is the Zariski closure of the image of f, if $\widetilde{Z} \to Z$ is a proper birational morphism with $\widetilde{Z}$ projective, if $\mathscr{L}$ is an ample line sheaf on $\widetilde{Z}$, and if $\tilde{f}\colon \mathbb{C} \to \widetilde{Z}$ is a lifting of f, then there are constants $c > 0$ and C such that $T_{\mathrm{big},f}(r) \ge c\,T_{\mathscr{L},f}(r) + C$ for all $r > 0$. This condition is independent of the choices of $\mathscr{L}$ and $\widetilde{Z}$. This height satisfies the same properties as in Remark 10.12. (There is no list of subvarieties in this case since in Nevanlinna theory f is usually fixed; however, one could define the big height instead by using the same $U_1, \dots, U_n$, $\overline{U}_1, \dots, \overline{U}_n$, and $\mathscr{L}_1, \dots, \mathscr{L}_n$ as in Remark 10.12; then extend $f^{-1}(U_i) \to U_i$ to a map $\mathbb{C} \to \overline{U}_i$ for i such that U_i contains the generic point of Z.)

13 Integral Points

Weil functions can be used to study integral points on varieties. This includes not only affine varieties, but also non-affine varieties. Integral points on non-affine varieties come up in some important applications, such as moduli spaces of abelian varieties.

To begin, let k be a number field and recall that a point $(x_1,\dots,x_n)\in\mathbb{A}^n(k)$ is an **integral point** if all x_i lie in $\mathscr{O}_k$. More generally, if $S\supseteq S_\infty$ is a finite set of places of k, then $(x_1,\dots,x_n)$ as above is an **S-integral point** if all x_i lie in the ring

$$\mathscr{O}_{k,S} := \{x\in k : \|x\|_v \le 1 \text{ for all } v\notin S\} \tag{58}$$

of S-integers. Algebraic points $(x_1,\dots,x_n)\in\mathbb{A}^n(\bar{k})$ are **integral** (resp. **S-integral**) if all of the x_i are integral over $\mathscr{O}_k$ (resp. $\mathscr{O}_{k,S}$). (Of course, $\mathscr{O}_k=\mathscr{O}_{k,S_\infty}$, so only one definition is really needed.) These definitions are inherited by points on a closed subvariety X of $\mathbb{A}^n_k$.

Given an abstract affine variety X over k, however, the situation becomes a little more complicated. Indeed, for any rational point $x\in X(k)$, there is a closed embedding into $\mathbb{A}^n_k$ for some n that takes x to an integral point. The same is true for algebraic points.

Instead, therefore, we refer to integrality of a *set* of points [74, § 1.3]: Let X be an affine variety over k. Then a set $\Sigma\subseteq X(k)$ (resp. $\Sigma\subseteq X(\bar{k})$) is **$S$-integral** if there is a closed immersion $i\colon X\hookrightarrow\mathbb{A}^n_k$ for some n and a nonzero element $a\in k$ such that, for all $x\in\Sigma$, all coordinates of $i(x)$ lie in $(1/a)\mathscr{O}_{k,S}$ (resp. a times all coordinates are integral over $\mathscr{O}_{k,S}$).

As noted above, this definition is meaningful only if Σ is an infinite set.

This definition can be phrased in geometric terms using Weil functions. Indeed, we identify $\mathbb{A}^n_k$ with the complement of the hyperplane $x_0=0$ in $\mathbb{P}^n_k$, by identifying $(x_1,\dots,x_n)\in\mathbb{A}^n_k$ with the point $[1:x_1:\dots:x_n]\in\mathbb{P}^n_k$. Let H denote the hyperplane $x_0=0$ at infinity, and let λ_H be the Weil function (47):

$$\begin{aligned}\lambda_{H,v}([1:x_1:\dots:x_n]) &= -\log\frac{\|1\|_v}{\max\{\|1\|_v,\|x_1\|_v,\dots,\|x_n\|_v\}}\\ &= \log\max\{\|1\|_v,\|x_1\|_v,\dots,\|x_n\|_v\}.\end{aligned} \tag{59}$$

Now let a be a nonzero element of k, let $x\in\bar{k}$, and let L be a number field containing $k(x)$. Then ax is integral over $\mathscr{O}_{k,S}$ if and only if $\|ax\|_w\le 1$ for all places w of M_L lying over places in $M_k\setminus S$, which holds if and only if $\|x\|_w\le\|a\|_w^{-1}$ for all such w. Thus, by (59), $\Sigma\subseteq X(\bar{k})$ is S-integral if and only if there is a closed immersion $i\colon X\hookrightarrow\mathbb{A}^n_k$ for some n and an M_k-constant (c_v) with the following property. For all $x\in\Sigma$, $\lambda_{H,w}(x)\le c_w$ for all places w of $M_{k(x)}$ lying over places not in S. (Here, as above, we identify $\mathbb{A}^n_k$ with $\mathbb{P}^n_k\setminus H$.)

By functoriality of Weil functions (Theorem 9.8b), this leads to the following definition.

Definition 13.1. Let k be a number field, let $S \supseteq S_\infty$ be a finite set of places of k, let X be a complete variety over k, and let D be an effective Cartier divisor on X. Then a set $\Sigma \subseteq X(\bar{k})$ is a (D,S)**-integral set of points** if (i) no point $x \in \Sigma$ lies in the support of D, and (ii) there is a Weil function λ_D for D and an M_k-constant (c_v) such that

$$\lambda_{D,w}(x) \le c_w$$

for all $x \in \Sigma$ and all places w of $M_{k(x)}$ not lying over places in S.

We may eliminate S from the notation if it is clear from the context, and refer to a D-integral set of points.

From the above discussion, it follows that the condition in the earlier definition of integrality holds for some closed immersion into $\mathbb{A}^n_k$, then it holds for all such closed immersions (with varying n).

Similarly, by Theorem 9.8d, one can use a fixed Weil function λ_D in Definition 13.1 (after adjusting (c_v)). One can also vary the divisor, as follows.

Proposition 13.2. *If k, S, and X are as above, and if D_1 and D_2 are effective Cartier divisors on X with the same support, then a set $\Sigma \subseteq X(\bar{k})$ is D_1-integral if and only if it is D_2-integral.*

Proof. This follows from Theorem 9.8a, e (additivity and boundedness of Weil functions). Details are left to the reader. □

Thus, D-integrality depends only on the support of D. In fact, one can go further: It depends only on the open subvariety $X \setminus \operatorname{Supp} D$:

Proposition 13.3. *Let k and S be as above, let X_1 and X_2 be complete k-varieties, and let D_1 and D_2 be effective Cartier divisors on X_1 and X_2, respectively. Assume that*

$$\phi\colon X_1 \setminus \operatorname{Supp} D_1 \xrightarrow{\sim} X_2 \setminus \operatorname{Supp} D_2$$

is an isomorphism. Then a set $\Sigma \subseteq X_1(\bar{k})$ is a D_1-integral set on X_1 if and only if

$$\phi(\Sigma) := \{\phi(x) : x \in \Sigma\}$$

is a D_2-integral set on X_2.

Proof. By working with the closure of the graph, we may assume that ϕ extends to a morphism from X_1 to X_2. In that case, it follows from Theorem 9.8a, e. □

Definition 13.4. Let k and S be as above, and let U be a variety over k. A set $\Sigma \subseteq U(\bar{k})$ is **integral** if there is an open immersion $i\colon U \to X$ of U into a complete variety X over k and an effective Cartier divisor D on X such that $i(U) = X \setminus \operatorname{Supp} D$ and $i(\Sigma)$ is a D-integral set on X.

Proposition 13.5. *Let $\phi\colon X_1 \to X_2$ be a morphism of complete k-varieties, and let D_1 and D_2 be effective Cartier divisors on X_1 and X_2, respectively. Assume that the* *support of D_2 does not contain the image of ϕ, and that the support of D_1 contains the support of ϕ^*D_2. If Σ is a D_1-integral set on X_1, then*

$$\phi(\Sigma) := \{\phi(x) : x \in \Sigma\}$$

is a D_2-integral set on X_2.

Proof. By Proposition 13.2, we may assume that $D_1 - \phi^*D_2$ is effective, and then the result follows by Theorem 9.8a, e. □

If we let $U_1 = X_1 \setminus \operatorname{Supp} D_1$ and $U_2 = X_2 \setminus \operatorname{Supp} D_2$, then the above conditions on the supports of D_1 and D_2 are equivalent to $\phi(U_1) \subseteq U_2$. Therefore Proposition 13.5 says that integral sets of points on varieties are preserved by morphisms of those varieties. This phenomenon is more obvious when using models over $\operatorname{Spec} \mathscr{O}_k$ to work with integral points, but this will not be explored in these notes.

We also note that Definition 13.4 does not require U to be affine. Indeed, many moduli spaces are neither affine nor projective, and it is often useful to work with integral points on those moduli spaces (although this is usually done using models). In an extreme case, U can be a complete variety. This corresponds to taking $D = 0$ in Definition 13.1, and the integrality condition is therefore vacuous in that case.

When working with rational points, Definition 13.1 can be stated using counting functions instead: $\Sigma \subseteq X(k)$ is integral if and only if $N_S(D,x)$ is bounded for $x \in \Sigma$. This is no longer equivalent when working with algebraic points, or when working over function fields, though.

The discussion of the corresponding notion in Nevanlinna theory is quite short: an (infinite) D-integral set of rational points on a complete k-variety X corresponds to a holomorphic curve f in a complete complex variety X whose image is disjoint from the support of a given Cartier divisor D on X. (In other words, $N_f(D,r) = 0$ for all r.) The next section will discuss an example of this comparison.

Of course, holomorphic curves omitting divisors also behave as in Proposition 13.3: Let $\phi : X_1 \to X_2$ be a morphism of complete complex varieties, let D_1 and D_2 be effective Cartier divisors on X_1 and X_2, respectively, with

$$\phi^{-1}(\operatorname{Supp} D_2) \subseteq \operatorname{Supp} D_1,$$

and let $f : \mathbb{C} \to X_1$ be a holomorphic curve which omits D_1. Then $\phi \circ f : \mathbb{C} \to X_2$ omits D_2, for trivial reasons.

Now consider the situation where $\phi : X' \to X$ is a morphism of complete complex varieties, D is an effective Cartier divisor on X, and D' is an effective Cartier divisor on X' with $\operatorname{Supp} D' = \phi^{-1}(\operatorname{Supp} D)$. Assume that ϕ is étale outside of $\operatorname{Supp} D'$. Then any holomorphic curve $f : \mathbb{C} \to X \setminus \operatorname{Supp} D$ lifts to a holomorphic curve $f' : \mathbb{C} \to X' \setminus \operatorname{Supp} D'$ such that $\phi \circ f' = f$, essentially for topological reasons.

What is surprising is that this situation carries over to the number field case. Indeed, let $\phi : X' \to X$ be a morphism of complete k-varieties, and let D and D' be as above, with ϕ étale outside of $\operatorname{Supp} D'$. If Σ is a set of D-integral points in $X(k)$, then $\phi^{-1}(\Sigma)$ is a set of integral points in $X'(\bar{k})$. The Chevalley-Weil theorem extends to integral points by Serre [74], Sect. 4.2 or Vojta [87], Sect. 5.1, and implies that although the points of Σ' may not lie in $X'(k)$, the ramification of the fields $k(x)$ over k is bounded uniformly for all $x \in X'(k)$. Combining this with the Hermite-Minkowski theorem, it then follows that there is a number field $L \supseteq k$ such that $\Sigma' \subseteq X'(L)$.

14 Units and the Borel Lemma

Units in a number field k can be related to integral points on the affine variety $xy = 1$ in $\mathbb{P}^2_k$: u is a unit if and only if there is a point (u, v) on this variety with $u, v \in \mathscr{O}_k$. This variety is isomorphic to $\mathbb{P}^1$ minus two points, which we may take to be 0 and ∞. More generally, a set of rational points on $\mathbb{P}^1 \setminus \{0, \infty\}$ is integral if and only if it is contained in finitely many cosets of the units in the group k^*.

Units therefore correspond to entire functions that never vanish. An entire function f never vanishes if and only if it can be written as e^g for an entire function g. This leads to what is called the "Borel lemma" in Nevanlinna theory.

Theorem 14.1. [7] *If $g_1, \dots, g_n$ are entire functions such that*

$$e^{g_1} + \cdots + e^{g_n} = 1, \tag{60}$$

then some g_j is constant.

Proof. We may assume that $n \geq 2$. The homogeneous coordinates $[e^{g_1} : \ldots : e^{g_n}]$ define a holomorphic curve $f \colon \mathbb{C} \to \mathbb{P}^{n-1}(\mathbb{C})$. The image of this map omits the n coordinate hyperplanes, and also omits the hyperplane $x_1 + \cdots + x_n = 0$ (expressed in homogeneous coordinates $[x_1 : \ldots : x_n]$). Therefore $N_f(H_j, r) = 0$, as H_j varies over these $n+1$ hyperplanes. This contradicts (39) unless the image of f is contained in a hyperplane (note that n is different in (39)). One can then use the linear relation between the coordinates of f to eliminate one of the terms e^{g_j} and then conclude by induction. □

In fact, by induction, it can be shown that some nontrivial subsum of the terms on the left-hand side of (60) must vanish.

To find the counterpart of this result in number theory, change the e^{g_j} to units. This theorem is due to van der Poorten and Schlickewei [84], and independently to Evertse [19].

Theorem 14.2. (Unit Theorem) *Let k be a number field and let $S \supseteq S_\infty$ be a finite set of places of k. Let $\mathscr{U}$ be a collection of n-tuples $(u_1, \dots, u_n)$ of S-units in k that satisfy the equation*

$$a_1 u_1 + \cdots + a_n u_n = 1, \tag{61}$$

where $a_1, \dots, a_n$ are fixed nonzero elements of k. Then all but finitely many elements of $\mathscr{U}$ have the property that there is a nonempty proper subset J of $\{1, 2, \dots, n\}$ such that $\sum_{j \in J} a_j u_j = 0$.

Proof. Assume that the theorem is false, and let $\mathscr{U}'$ be the set of all $(u_1,\dots,u_n)$ for which there is no such J as above. Then $\mathscr{U}'$ is infinite.

If we regard each $(u_1,\dots,u_n) \in \mathscr{U}'$ as a point $[u_1 : \dots : u_n] \in \mathbb{P}_k^{n-1}$, then by looking directly at the formula (40) for Weil functions, we see that $N_S(H_j,x)$ is bounded as x varies over $\mathscr{U}'$, for the same set of $n+1$ hyperplanes as in the previous proof. This gives $m_S(H_j,x) = h_k(x) + O(1)$ for all $x \in U'$ and all j, contradicting Theorem 8.9 unless all points in $\mathbb{P}^{n-1}$ corresponding to points in $\mathscr{U}'$ lie in a finite union of proper linear subspaces.

Consider one of those linear subspaces containing infinitely many points of $\mathscr{U}'$. Combining the equation of some hyperplane containing that subspace with (61) allows one to eliminate one or more of the u_j, since by assumption there is no set J as in the statement of the theorem. We then proceed by induction on n (the base case $n = 1$ is trivial). □

Example 14.3. The condition with the set J is essential because, for example, the unit equation (61) with $n = 3$ and $a_1 = a_2 = a_3 = 1$ has solutions $u + (-u) + 1 = 1$ for infinitely many units u (if k or S is large enough). Geometrically, if $H_1,\dots,H_4$ are the hyperplanes in $\mathbb{P}_k^n$ occurring in the proofs of Theorems 14.1 and 14.2, then the possible sets $J = \{1,2\}$, $J = \{1,3\}$, and $J = \{2,3\}$ correspond to the line joining the points $H_1 \cap H_2$ and $H_3 \cap H_4$, the line joining the points $H_1 \cap H_3$ and $H_2 \cap H_4$ and the line joining the points $H_1 \cap H_4$ and $H_2 \cap H_3$. Each such line meets the divisor $D := \sum H_j$ in only two points, so if we map $\mathbb{P}^1$ to that line in such a way that 0 and ∞ are taken to those two points, then integral points on $\mathbb{P}_k^1 \setminus \{0,\infty\}$ (i.e., units) are taken to integral points on $\mathbb{P}_k^2 \setminus D$.

Finally, we note that theorems on exponentials of entire functions that can be reduced to Theorem 14.1 by elementary geometric arguments can be readily translated to theorems on units, by replacing the use of Theorem 14.1 with Theorem 14.2. For example, Theorem 14.4 below leads directly to Theorem 14.5.

Theorem 14.4. [16, Théorème XVI]; see also [26] and [30]. *Let $f\colon \mathbb{C} \to \mathbb{P}^n$ be a holomorphic curve that omits $n+m$ hyperplanes in general position, $m \ge 1$. Then the image of f is contained in a linear subspace of dimension $\le [n/m]$, where $[\cdot]$ denotes the greatest integer function.*

Theorem 14.5. [50, Cor. 3] *Let $\Sigma \subseteq \mathbb{P}^n(k)$ be a set of D-integral points, where D is the sum of $n+m$ hyperplanes in general position, $m \ge 1$. Then Σ is contained in a finite union of linear subspaces of dimension $\le [n/m]$.*

15 Conjectures in Nevanlinna Theory and Number Theory

Since the canonical line sheaf $\mathscr{K}$ of $\mathbb{P}^1$ is $\mathscr{O}(-2)$, the main inequality of Theorem 5.5 can be stated in the form

$$m_f(D,r) + T_{\mathscr{K},f}(r) \le_{\text{exc}} O(\log^+ T_f(r)) + o(\log r),$$

leading to a general conjecture in Nevanlinna theory. This first requires a definition.

Definition 15.1. A subset Z of a smooth complex variety X is said to have **normal crossings** if each $P \in X(\mathbb{C})$ has an open neighborhood U and holomorphic local coordinates $z_1, \dots, z_n$ in U such that $Z \cap U$ is given by $z_1 = \cdots = z_r = 0$ for some r $(0 \le r \le n)$. A divisor on X is **reduced** if all multiplicities occurring in it are either 0 or 1. Finally, a **normal crossings divisor** on X is a reduced divisor whose support has normal crossings.

(Note that not all authors assume that a normal crossings divisor is reduced.)

Conjecture 15.2. Let X be a smooth complex projective variety, let D be a normal crossings divisor on X, let $\mathscr{K}$ be the canonical line sheaf on X, and let $\mathscr{A}$ be an ample line sheaf on X. Then:

(a) The inequality

$$m_f(D,r) + T_{\mathscr{K},f}(r) \le_{\text{exc}} O(\log^+ T_{\mathscr{A},f}(r)) + o(\log r) \tag{62}$$

holds for all holomorphic curves $f\colon \mathbb{C} \to X$ with Zariski-dense image.

(b) For any $\varepsilon > 0$ there is a proper Zariski-closed subset Z of X, depending only on X, D, $\mathscr{A}$, and ε, such that the inequality

$$m_f(D,r) + T_{\mathscr{K},f}(r) \le_{\text{exc}} \varepsilon T_{\mathscr{A},f}(r) + C \tag{63}$$

holds for all nonconstant holomorphic curves $f\colon \mathbb{C} \to X$ whose image is not contained in Z, and for all $C \in \mathbb{R}$.

The form of this conjecture is the same as the (known) theorem for holomorphic maps to Riemann surfaces. It has also been shown to hold, with a possibly weaker error term, for holomorphic maps $\mathbb{C}^d \to X$ if $d = \dim X$ and the jacobian determinant of the map is not identically zero; see [82] and [10]. The conjecture itself is attributed to Griffiths, although he seems not to have put it in print anywhere.

Conjecture 15.2 has been proved for curves (Theorem 23.2 and Corollary 29.7), but in higher dimensions very little is known. If $X = \mathbb{P}^n$ and D is a sum of hyperplanes, then the normal crossings condition is equivalent to the hyperplanes being in general position, and in that case the first part of the conjecture reduces to Cartan's theorem (Theorem 8.6). The second part is also known in this case [90].

A consequence of Conjecture 15.2 concerns holomorphic curves in varieties of general type, or of log general type.

Proposition 15.3. *Assume that either part of Conjecture 15.2 is true. If X is a smooth variety of general type, then a holomorphic curve $f\colon \mathbb{C} \to X$ cannot have Zariski-dense image. More generally, if X is a smooth variety, D is a normal crossings divisor on X, and $X \setminus D$ is a variety of log general type, then a holomorphic curve $f\colon \mathbb{C} \to X \setminus D$ cannot have Zariski-dense image.*

Proof. Assume that part (a) of the conjecture is true. The proof for (b) is similar and is left to the reader.

As was the case with (24) and (39), (62) can be rephrased as

$$N_f(D,r) \geq_{\text{exc}} T_{\mathscr{K}(D),f}(r) - O(\log^+ T_{\mathscr{A},f}(r)) - o(\log r). \tag{64}$$

In this case, since f misses D, the left-hand side is zero. By the definition of log general type, the line sheaf $\mathscr{K}(D) := \mathscr{K} \otimes \mathscr{O}(D)$ is big. Therefore, this inequality contradicts Proposition 12.11. □

This consequence is also unknown in general. It is known, however, in the special case where X is a closed subvariety of a semiabelian variety and $D = 0$. Indeed, if X is a closed subvariety of a semiabelian variety and is not a translate of a semiabelian subvariety, then a holomorphic curve $f : \mathbb{C} \to X$ cannot have Zariski-dense image. See [41, 32, 78] for the case of abelian varieties, and [61] for the more general case of semiabelian varieties. All of these references build on work of Bloch [5].

Conjecture 15.2b is also known if X is an abelian variety and D is ample [79]. The theorem has been extended to semiabelian varieties again by Noguchi [63], but only applies to holomorphic curves whose image does not meet the divisor at infinity. Again, these proofs build on work of Bloch [5].

Conjecture 15.2 will be discussed further once its counterpart in number theory has been introduced. This, in turn, requires some definitions.

Definition 15.4. Let X be a nonsingular variety. A divisor D on X is said to have **strict normal crossings** if it is reduced, if each irreducible component of its support is nonsingular, and if those irreducible components meet transversally (i.e., their defining equations are linearly independent in the Zariski cotangent space at each point). We say that D has **normal crossings** if it has strict normal crossings locally in the étale topology. This means that for each $P \in X$ there is an étale morphism $\phi : X' \to X$ with image containing P such that $\phi^* D$ has strict normal crossings.

This definition is discussed more in [95, Sect. 7].

Definition 15.5. [83] Let X be a variety. A subset of $X(\bar{k})$ is **generic** if all infinite subsets are Zariski-dense in X.

The number-theoretic counterpart to Conjecture 15.2 is then the following.

Conjecture 15.6. Let k be a number field, let $S \supseteq S_\infty$ be a finite set of places of k, let X be a smooth projective variety over k, let D be a normal crossings divisor on X, let $\mathscr{K}$ be the canonical line sheaf on X, and let $\mathscr{A}$ be an ample line sheaf on X. Then:

(a) Let Σ be a generic subset of $X(k) \setminus \operatorname{Supp} D$. Then the inequality

$$m_S(D,x) + h_{\mathscr{K},k}(x) \leq O(\log^+ h_{\mathscr{A},k}(x)) \tag{65}$$

holds for all $x \in \Sigma$.

(b) For any $\varepsilon > 0$ there is a proper Zariski-closed subset Z of X, depending only on X, D, $\mathscr{A}$, and ε, such that for all $C \in \mathbb{R}$ the inequality

$$m_S(D,x) + h_{\mathscr{K},k}(x) \le \varepsilon\, h_{\mathscr{A},k}(x) + C \tag{66}$$

holds for almost all $x \in (X \setminus Z)(k)$.

By Remark 10.12, one can replace $h_{\mathscr{A},k}$ in this conjecture with a big height (after possibly adjusting Z and ε in part (b)). One can then relax the condition on X to be just a smooth complete variety. The resulting conjecture actually would follow from the original Conjecture 15.6 by Chow's lemma, resolution of singularities, and Proposition 25.2 (without reference to $d_S(x)$ in the latter, since lifting a rational point to the cover does not involve passing to a larger number field in this case). This can also be done for Conjecture 15.2.

Except for error terms, the cases in which Conjecture 15.6 is known correspond closely to those cases for which Conjecture 15.2 is known. Indeed, Conjecture 15.6b is known for curves by Roth's theorem, by a theorem of Lang [44], Thm. 2, and by Faltings's theorem on the Mordell conjecture [21, 22], for genus 0, 1, and > 1, respectively. For curves, part (a) of the conjecture is identical to part (b) except for the error term. Also, Schmidt's Subspace Theorem (Theorem 8.10) proves Conjecture 15.6 except for the error term in part (a), and the assertion on the dependence of the set Z in part (b). As noted earlier, however, the latter assertion is also known (without the dependence on $\mathscr{A}$ and ε).

Remark 15.7. Conjecture 15.6 (and also Conjecture 15.2) are compatible with taking products. Indeed, let $X = X_1 \times_k X_2$ be the product of two smooth projective varieties, with projection morphisms $p_i \colon X \to X_i$ $(i = 1,2)$. Let D_1 and D_2 be normal crossings divisors on X_1 and X_2, respectively, and let $\mathscr{K}$, $\mathscr{K}_1$ and $\mathscr{K}_2$ be the canonical line sheaves on X, X_1, and X_2, respectively. We have $\mathscr{K} \cong p_1^*\mathscr{K}_1 \otimes p_2^*\mathscr{K}_2$. Then the conjecture for D_1 on X_1 and for D_2 on X_2 imply the conjecture for $p_1^*D_1 + p_2^*D_2$ on X.

Remark 15.8. One may ask whether one can make the same change to this conjecture as was done in going from Theorem 8.9 to 8.10 (and likewise in the Nevanlinna case). One can, but it would not make the conjecture any stronger. Indeed, suppose that $D_1, \dots, D_q$ are normal crossings divisors on X. There exists a smooth projective variety X' over k and a proper birational morphism $\pi \colon X' \to X$ such that the support of the divisor $\sum \pi^* D_i$ has normal crossings. Let D' be the reduced divisor on X' for which $\operatorname{Supp} D' = \operatorname{Supp} \sum \pi^* D_i$, and let $\mathscr{K}'$ and $\mathscr{K}$ be the canonical line sheaves of X' and X, respectively. By Proposition 25.2, we have

$$\sum_{v \in S} \max\{\lambda_{D_i,v}(x) : i = 1,\dots,q\} + h_{\mathscr{K},k}(x) \le m_S(D',x') + h_{\mathscr{K}',k}(x') + O(1) \tag{67}$$

for all $x' \in X'(k)$, where $x = \pi(x')$. Therefore, if the left-hand side of (65) or (66) were replaced by the left-hand side of (67), then the resulting conjecture would be a consequence of Conjecture 15.6 applied to D' on X'.

Theorems 8.10 and 8.11 are still needed, though, because Conjectures 15.6 and 15.2 have not been proved for blowings-up of $\mathbb{P}^n$.

Corresponding to Proposition 15.3, we also have

Proposition 15.9. *Assume that either part of Conjecture 15.6 is true. Let k and S be as in Conjecture 15.6, let X be a smooth projective variety over k, and let D be a normal crossings divisor on X. Assume that $X \setminus D$ is of log general type. Then no set of S-integral k-rational points on $X \setminus D$ is Zariski dense.*

Proof. As in the earlier proof, (65) is equivalent to

$$N_S(D,x) \geq h_{\mathscr{K}(D),k}(x) - O(\log^+ h_{\mathscr{A},k}(x)),$$

and (66) can be rephrased similarly. For points x in a Zariski-dense set of k-rational S-integral points, $N_S(D,x)$ would be bounded, contradicting Proposition 10.11 since $\mathscr{K}(D)$ is big. □

This proof shows how Conjecture 15.6 is tied to the Mordell conjecture.

As was the case in Nevanlinna theory, the conclusion of Proposition 15.9 has been shown to hold for closed subvarieties of semiabelian varieties, by Faltings [24] in the abelian case and Vojta [89] in the semiabelian case.

In addition, Conjecture 15.6b has been proved when X is an abelian variety and D is ample [23]. This has been extended to semiabelian varieties [92], but in that case (66) was shown only to hold for sets of integral points on the semiabelian variety.

In parts (b) of Conjectures 15.2 and 15.6, the exceptional set Z must depend on ε; this is because of the following theorem.

Theorem 15.10. [52] *There are examples of smooth projective surfaces X containing infinitely many rational curves Z_i for which the restrictions of (62) and (65) fail to hold.*

These examples do not contradict parts (b) of the conjectures of this section, since the degrees of the curves increase to infinity. Nor do they preclude the sharper error terms in parts (a) of the conjectures. However, they do prevent one from combining the two halves of each conjecture.

Lang [48, Chap. I, Sect. 3] has an extensive conjectural framework concerning how the exceptional set in part (b) of Conjectures 15.2 and 15.6 may behave, especially for varieties of general type (which would not include the examples of Theorem 15.10). See also Sect. 17. Note, however, that the exceptional sets of that section refer only to integral points (or holomorphic curves missing D), so the exceptional sets referenced here are more general.

As a converse of sorts, there are numerous examples of theorems in analysis that apply only to "very generic" situations; i.e., they exclude a countable union of proper analytic subsets. One could pose Conjecture 15.2a in such a setting as well. Such a change would not be meaningful for Conjecture 15.6a, however, since the set of rational (and even algebraic) points on a variety is at most countable.

The formulation of Conjecture 15.6 suggests that, in a higher-dimensional setting, the correct counterpart in number theory for a holomorphic curve with Zariski-dense image is not just an infinite set of rational (or algebraic) points, but an infinite *generic* set. Corresponding to holomorphic curves whose images need not be Zariski-dense, we also make the following definition.

Definition 15.11. Let X be a variety over a number field k. If Z is a closed subvariety of X, then a **Z-generic subset** of $X(\bar{k})$ is a generic subset of $Z(\bar{k})$. Also, a **semi-generic** subset of $X(\bar{k})$ is a Z-generic subset of $X(\bar{k})$ for some closed subvariety Z of X.

A version of Conjecture 15.6 has also been posed for algebraic points. See Conjecture 25.1.

16 Function Fields

Although function fields are not emphasized in these lectures, they provide useful insights, especially when discussing Arakelov theory or use of models. They are briefly introduced in this section. Most results are stated without proof.

Mahler [54] and Osgood [64] showed that Roth's theorem is false for function fields of positive characteristic. Therefore these notes will discuss only function fields of characteristic zero.

For the purposes of these notes, a function field is a finitely-generated field extension of a "ground field" F, of transcendence degree 1. Such a field is called a "function field over F."

If k is a function field over F, it is the function field $K(B)$ for a unique (up to isomorphism) nonsingular projective curve B over F. For each closed point b on B, the local ring $\mathscr{O}_{B,b}$ is a discrete valuation ring whose valuation v determines a non-archimedean place of k with a corresponding norm given by $\|x\|_v = 0$ if $x = 0$ and by the formula

$$\|x\|_v = e^{-[\kappa(b):F]v(x)}$$

if $x \neq 0$. Here v is assumed to be normalized so that its image is $\mathbb{Z}$. We set $N_v = 0$, so that axioms (3) hold.

Let L be a finite extension of k. Then it, too, is a function field over F, and (4) and (5) hold in this context. If B' is the nonsingular projective curve over F corresponding to L, then the inclusion $k \subseteq L$ uniquely determines a finite morphism $B' \to B$ over F.

The field F is not assumed to be algebraically closed. In this context, note that the degree of a divisor on B is defined to be

$$\deg \sum n_b \cdot b = \sum n_b [\kappa(b) : F]. \tag{68}$$

This degree depends on F, since if $F \subseteq F' \subseteq k$ and k is also of transcendence degree 1 over F', then F' is necessarily finite over F,[1] and the degree is divided by $[F' : F]$ if it is taken relative to F' instead of to F.

With this definition of degree, principal divisors have degree 0, which implies that the Product Formula (6) holds, where the set M_k is the set of closed points on the corresponding nonsingular curve B. The Product Formula is the primary condition for k to be a **global field** (for the full set of conditions, see [2, Chap. 12]). There are many other commonalities between function fields and number fields; for example, the affine ring of any nonempty open affine subset of B is a Dedekind ring.

A function field is always implicitly assumed to be given with the subfield F, since (for example) $\mathbb{C}(x,y)$ can be viewed as a function field with either $F = \mathbb{C}(x)$ or $F = \mathbb{C}(y)$, with very different results.

For the remainder of this section, k is a function field of characteristic 0 over a field F, and B is a nonsingular projective curve over F with $k = K(B)$.

A key benefit of working over function fields is the ability to explore diophantine questions using standard tools of algebraic geometry, using the notion of a model.

Definition 16.1. A **model** for a variety X over k is an integral scheme $\mathscr{X}$, given with a flat morphism $\pi\colon \mathscr{X} \to B$ of finite type and an isomorphism $\mathscr{X} \times_B \operatorname{Spec} k \cong X$ of schemes over k. The model is said to be **projective** (resp. **proper**) if the morphism π is projective (resp. proper).

If X is a projective variety over k, then a projective model can be constructed for it by taking the closure in $\mathbb{P}^N_B$. Likewise, a proper model for a complete variety exists, by Nagata's embedding theorem. In either case the model may be constructed so that any given finite collection of Cartier divisors and line sheaves extends to the same sorts of objects on the model [96].

If $\mathscr{X}$ is a proper model over B for a complete variety X over k, then rational points in $X(k)$ correspond naturally and bijectively to sections $i\colon B \to \mathscr{X}$ of $\pi\colon \mathscr{X} \to B$. Indeed, if i is such a section then it takes the generic point of B to a point on the generic fiber of π, which is X. Conversely, given a point in $X(k)$, one can take its closure to get a curve in $\mathscr{X}$; it is then possible to show that the restriction of π to this curve is an isomorphism.

More generally, if L is a finite extension of k, and B' is the nonsingular projective curve over F corresponding to L, then points in $X(L)$ correspond naturally and bijectively to morphisms $B' \to \mathscr{X}$ over B. This follows by applying the above argument to $\mathscr{X} \times_B B'$, which is a proper model for $X \times_k L$ over B'.

With this notation, we can define Weil functions in the function field case as follows.

---

[1] Let $t \in k$ be transcendental over F. Then $F(t)$ and F' are linearly disjoint over F, and therefore $[F' : F] = [F'(t) : F(t)] \le [k : F(t)] < \infty$.

Definition 16.2. Let X be a complete variety over k, let D be a Cartier divisor on X, and let $\pi\colon \mathscr{X} \to B$ be a proper model for X. Assume that D extends to a Cartier divisor on $\mathscr{X}$, also denoted by D. Let L be a finite extension of k, let $x \in X(L)$ be a point not lying on $\operatorname{Supp} D$, let $i\colon B' \to \mathscr{X}$ be the corresponding morphism, as above, and let w be a place of L, corresponding to a closed point b' of B'. Then the image of i is not contained in $\operatorname{Supp} D$ on $\mathscr{X}$, so i^*D is a Cartier divisor on B'. Let n_w be the multiplicity of b' in i^*D. We then define

$$\lambda_{D,w}(x) = n_w[\kappa(b') : F].$$

(One may be tempted to require the notation to indicate the choice of F, but this is not necessary since the choice of F is encapsulated in the place w, which comes with a norm $\|\cdot\|_w$ that depends on F.)

It is possible to show that this definition satisfies the conditions of Definition 9.6 (where now k is a function field). Consequently, Theorems 9.8 and 9.10 hold in this context. Moreover, this definition is compatible with (50) (corresponding to changing B').

In the case of Theorem 9.8, though, a bit more is true: the M_k-constants are not necessary when one works with Cartier divisors on the model. Indeed, if D is an effective Cartier divisor on a model $\mathscr{X}$ of a k-variety X, then (in the notation of Definition 16.2) i^*D is an effective divisor on B', so $n_b \ge 0$ for all b, hence $\lambda_{D,w}(x) \ge 0$ for all w and all $x \notin \operatorname{Supp} D$. Similarly, suppose that D is a Cartier divisor on X, and that λ_D and λ'_D are two Weil functions obtained from extensions D and D' of D to models $\mathscr{X}$ and $\mathscr{X}'$, respectively, using Definition 16.2. We may reduce to the situation where the two models are the same: let $\mathscr{X}''$ be the closure of the graph of the birational map between $\mathscr{X}$ and $\mathscr{X}'$, and pull back D and D' to $\mathscr{X}''$. But now the difference $D - D'$ is a divisor on $\mathscr{X}$ which does not meet the generic fiber, so it is supported only on a finite sum of closed fibers of $\pi\colon \mathscr{X} \to B$. Therefore the Weil function associated to $D - D'$, and hence the difference between λ_D and λ'_D, is bounded by an M_k-constant.

Following Sect. 10, one can then define the height of points $x \notin \operatorname{Supp} D$, starting from a model $\mathscr{X}$ for X over B and a Cartier divisor D on $\mathscr{X}$: with notation as in Definition 16.2, we have

$$\begin{aligned} h_{D,k}(x) &= \frac{1}{[L:k]} \sum_{w \in M_L} \lambda_{D,w}(x) \\ &= \frac{1}{[L:k]} \sum_{w \in M_L} n_w[\kappa(b') : F] \\ &= \frac{\deg i^*D}{[L:k]} \end{aligned} \tag{69}$$

by (68).

Therefore, heights can be expressed using intersection numbers. It is this observation that led to the development of Arakelov theory, which defines models over $\mathscr{O}_k$

of varieties over number fields k, with additional information at archimedean places which again allows heights to be described using suitable intersection numbers.

Returning to function fields, heights defined as in (69) are defined exactly (given a model of the variety and an extension of the Cartier divisor to that model). Except for Theorem 10.8 (Northcott's theorem), all of the results of Sect. 10 extend to the case of heights defined as in (69). In particular, if $\mathscr{L}$ is a line sheaf on a model $\mathscr{X}$ for X, then

$$h_{\mathscr{L},k}(x) = \frac{\deg i^*\mathscr{L}}{[L:k]}. \tag{70}$$

Northcott's theorem is false over function fields (unless F is finite). Instead, however, it is true that the set (55) is parametrized by a scheme of finite type over F.

Models also provide a very geometric way of looking at integral points. For example, consider the situation with rational points. Let S be a finite set of places of k; this corresponds to a proper Zariski-closed subset of B, which we also denote by S. Let X, $\pi\colon \mathscr{X} \to B$, $x \in X(k)$, and $i\colon B \to \mathscr{X}$ be as in Definition 16.2 (with $L = k$), let D be an effective Cartier divisor on $\mathscr{X}$, and let λ_D be the corresponding Weil function as in Definition 16.2. Then, for any place $v \in M_k$, we have $\lambda_{D,v}(x) > 0$ if and only if $i(b)$ lies in the support of D, where $b \in B$ is the closed point associated to v. Thus, a rational point satisfies the condition of Definition 13.1 with λ_D as above and $c_v = 0$ if and only if it corresponds to a section of the map $\pi^{-1}(B \setminus S) \to B \setminus S$. A similar situation holds with algebraic points, which then correspond to multisections of $\pi^{-1}(B \setminus S) \to B \setminus S$.

Conversely, given a set of integral points as in Definition 13.1, by performing some blowings-up one can construct a model and an effective Cartier divisor on that model for which each of the given integral points corresponds to a section (or multisection) as above.

This formalism works also over number fields, without the need for Arakelov theory.

17 The Exceptional Set

The exceptional set mentioned in Conjectures 15.2b and 15.6b leads to interesting questions of its own, even when working only with rational points (or integral points) in the contexts of Propositions 15.3 or 15.9. This question has been explored in more detail by Lang; this is the main topic of this section. For references, see [47], [48, Chap. I, Sect. 3], and [51].

Definition 17.1. Let X be a complete variety over a field k.

(a) The **exceptional set** $\operatorname{Exc}(X)$ is the Zariski closure of the union of the images of all nonconstant rational maps $G \dashrightarrow X$, where G is a group variety over an extension field of k.

(b) If k is a finitely generated extension field of $\mathbb{Q}$, then the **diophantine exceptional set** $\mathrm{Exc}_{\mathrm{dio}}(X)$ is the smallest Zariski-closed subset Z of X such that $(X \setminus Z)(L)$ is finite for all fields L finitely generated over k.
(c) If $k = \mathbb{C}$ then the **holomorphic exceptional set** $\mathrm{Exc}_{\mathrm{hol}}(X)$ is the Zariski closure of the union of the images of all nonconstant holomorphic curves $\mathbb{C} \to X$.

Each of these sets (when defined) may be empty, all of X, or something in between. For each of these types of exceptional set, Lang has conjectured that the exceptional set is a proper subset if the variety X is of general type (but not conversely – see below). He also has conjectured that $\mathrm{Exc}_{\mathrm{dio}}(X) = \mathrm{Exc}(X)$ if k is a finitely-generated extension of $\mathbb{Q}$, and that $\mathrm{Exc}_{\mathrm{hol}}(X) = \mathrm{Exc}(X)$ if $k = \mathbb{C}$. And, finally, if k is finitely generated over $\mathbb{Q}$, then he conjectured that

$$\mathrm{Exc}_{\mathrm{dio}}(X) \times_k \mathbb{C} = \mathrm{Exc}_{\mathrm{hol}}(X \times_k \mathbb{C})$$

for all embeddings $k \hookrightarrow \mathbb{C}$.

The main example in which this conjecture is known is in the context of closed subvarieties of abelian varieties [41]:

Theorem 17.2. (Kawamata Structure Theorem) *Let X be a closed subvariety of an abelian variety A over $\mathbb{C}$. The **Kawamata locus** of X is the union $Z(X)$ of all translated abelian subvarieties of A contained in X. It is a Zariski-closed subset of X, and is a proper subset if and only if X is not fibered by (nontrivial) abelian subvarieties of A.*

This theorem is also true for semiabelian varieties, and by induction on dimension it follows from [61] that the image of a nonconstant holomorphic curve $\mathbb{C} \to X$ must be contained in $Z(X)$. Similarly, if X is a closed subvariety of a semiabelian variety A over a number field k, then any set of integral points on X can contain only finitely many points outside of $Z(X)$. It is also known that a closed subvariety X of a semiabelian variety A is of log general type if and only if it is not fibered by nontrivial semiabelian subvarieties of A. Thus, (restricting to A abelian) Lang's conjectures have been verified for closed subvarieties of abelian varieties.

In a similar vein, the finite collections of proper linear subspaces of positive dimension in Remark 8.12 are the same in Theorems 8.10 and 8.11.

In the context of integral points or holomorphic curves missing divisors, one can also define the same three types of exceptional sets. The changes are obvious, except possibly for $\mathrm{Exc}(X \setminus D)$: In this case it should be the Zariski closure of the union of the images of all non-constant strictly rational maps $G \dashrightarrow X \setminus D$, where G is a group variety. A **strictly rational map** [39, Sect. 2.12] is a rational map $X \dashrightarrow Y$ such that the closure of the graph is proper over X. This variation has not been studied much, though.

More conjectures relating the geometry of a variety and its diophantine properties are described by Campana [9]. He further classifies varieties in terms of fibrations. For example, let $X = C \times \mathbb{P}^1$ where C is a smooth projective curve of genus ≥ 2. This is an example of a variety which is not of general type, but for which all of

Lang's exceptional sets are the entire variety. Yet, for any given number field, $X(k)$ is not Zariski dense, and there are no Zariski-dense holomorphic curves in X. Campana's framework singles out the projection $X \to C$. This projection has general type base, and fibers have Zariski-dense sets of rational points over a large enough field (depending on the fiber).

For varieties of negative Kodaira dimension, the diophantine properties are studied in conjectures of Manin concerning the rate of growth of sets of rational points of height $\le B$, as B varies. This is a very active area of number theory, but is beyond the scope of these notes.

18 Comparison of Problem Types

Before the analogy with Nevanlinna theory came on the scene, things were quite simple: You tried to prove something over number fields, and if you got stuck you tried function fields. If you succeeded over function fields, then you tried to translate the proof over to the number field case. For example, the Mordell conjecture was first proved for function fields by Manin, then Grauert modified his proof. But, those proofs used the absolute tangent bundle, which has no known counterpart over number fields. Ultimately, though, Faltings' proof of the Mordell conjecture did draw upon work over function fields, of Tate, Szpiro, and others.

The analogy with Nevanlinna theory gives a second way of working by analogy, although it is more distant. Also, more recent work has placed more emphasis on higher dimensional varieties, lending more importance to the exceptional set. Thus, a particular diophantine problem leads one to a number of related problems which may be easier and whose solutions may provide some insight into the original problem. These can be (approximately) linearly ordered, as follows. In each case, one looks at a class of pairs (X,D) consisting of a smooth complete variety X over the appropriate field and a normal crossings divisor D on X. Each class has been split into a qualitative part (A) and a quantitative part (B).

1A: Find the exceptional set $\operatorname{Exc}(X \setminus D)$.
B: For each $\varepsilon > 0$, find the exceptional subset Z in Conjectures 15.2 and 15.6. This should be the Zariski closure of the union of all closed subvarieties $Y \subseteq X$ such that, after resolving singularities of Y and of $D\big|_Y$, the main inequality (63) or (66) on Y is weaker than that obtained by restricting the same inequality on X.
2A: Prove that, given any smooth projective curve Y over a field of characteristic zero, and any finite subset $S \subseteq Y$, the set of maps $Y \setminus S \to X \setminus D$ whose image is not contained in the exceptional set, is parametrized by a finite union of varieties.
B: Prove Conjecture 15.6 in the split function field case of characteristic zero.
3A: Prove that all holomorphic curves $\mathbb{C} \to X \setminus D$ must lie in the exceptional set.
B: Prove Conjecture 15.2 for holomorphic curves $\mathbb{C} \to X$.

4A: In the (general) function field case of characteristic zero, prove that the set of integral points on $X \setminus D$ outside of the exceptional set is finite.
B: Prove Conjecture 15.6 over function fields of characteristic zero.
5A: Prove over number fields that the set of integral points on $X \setminus D$ outside of the exceptional set is finite.
B: Prove Conjecture 15.6 (in the number field case).

In the case of the Mordell conjecture, for example, X would lie in the class of smooth projective curves of genus > 1 over the appropriate field, and D would be zero.

As another example, see Corollary 29.9, in going from 2A or 2B to 3A or 3B.

In each of the above items except 1A and 1B, one might also consider algebraic points of bounded degree (or holomorphic functions from a finite ramified covering of $\mathbb{C}$). See Sects. 25 and 27.

19 Embeddings

A major goal of these lectures is to describe recent work on partial proofs of Conjecture 15.6 (as well as Conjecture 15.2). One general approach is to use embeddings into larger varieties to sharpen the inequalities. This can only work if the conjecture is known on the larger variety, and if the exceptional set is also known. At the present time, all work on this has used Schmidt's Subspace Theorem (and Cartan's theorem).

This section will discuss some of the issues involved, before delving into some of the specific methods in following sections.

We begin by considering the example where $X = \mathbb{P}^2_k$ and D is a normal crossings divisor of degree ≥ 4. Then $X \setminus D$ is of log general type, and therefore (if k is a number field) integral sets of points on $X \setminus D$ cannot be Zariski dense, or (if $k = \mathbb{C}$) holomorphic curves $\mathbb{C} \to X \setminus D$ cannot have Zariski-dense image. From now on we will refer only to the number-theoretic case; the version in Nevanlinna theory is similar.

If D is a smooth divisor, then there is no clue on how to proceed. In the other extreme, if D is a sum of at least four lines (in general position), then Schmidt's Theorem gives the answer; see Sect. 14.

If D is a sum of three lines and a conic, some results are known. For example, if L_1, L_2, and L_3 are linear forms defining the lines and Q is a homogeneous quadratic polynomial defining the conic, then all L_i^2/Q must be units (or nearly so) at integral points. Since they are algebraically dependent, we can apply the unit theorem; see [31] and [87, Cor. 2.4.3].

More recently, nontrivial approximation results have been obtained for conics and higher-degree divisors in projective space by using r-uple embeddings.

For example, under the 2-uple embedding $\mathbb{P}^2 \hookrightarrow \mathbb{P}^5$, the image of a conic is contained in a hyperplane. Therefore Schmidt's Subspace Theorem can be applied to $\mathbb{P}^5$ to give an approximation result (provided there are sufficiently many other components in the divisor). Under the 3-uple embedding $\mathbb{P}^2 \hookrightarrow \mathbb{P}^9$, things are better: the image of a conic spans a linear subspace of codimension 3, so there are three linearly independent hyperplanes containing it.

More generally, suppose D is a divisor of degree d in $\mathbb{P}^n$. We consider its image under the r-uple embedding $\mathbb{P}^n \hookrightarrow \mathbb{P}^{\binom{r+n}{n}-1}$. This image spans a linear subspace of codimension $\binom{r+n-d}{n}$, because there are that many monomials of degree $r-d$ in $n+1$ variables (they then get multiplied by the form defining D to get homogeneous polynomials of degree r in the homogeneous coordinate variables in $\mathbb{P}^n$, hence hyperplanes in the image space).

Applying Schmidt's Subspace Theorem to $\mathbb{P}^{\binom{r+n}{n}-1}$ would would then give an inequality of the form

$$\binom{r+n-d}{n} m(D,x) + \cdots \le \left(\binom{r+n}{n} \cdot r + \varepsilon \right) h_k(x) + O(1)$$

for $x \in \mathbb{P}^n(k)$ outside of a finite union of proper subvarieties of degree $\le r$. The idea is to take r large. As $r \to \infty$, the ratio of the coefficients in the above inequality tends to 0, because

$$\frac{\binom{r+n-d}{n}}{\binom{r+n}{n} \cdot r} = \frac{(r-d+n)\cdots(r-d+1)}{(r+n)\cdots(r+1)r} = \frac{r^n + O(r^{n-1})}{r(r^n + O(r^{n-1}))} \to \frac{1}{r} \to 0. \tag{71}$$

This is not useful, but we can try harder. Some hyperplanes in the image space can be made to contain D twice, or three times, etc. After taking this into account, the inequality improves to

$$\left(\binom{r+n-d}{n} + \binom{r+n-2d}{n} + \dots \right) m(D,x) + \dots$$
$$\le \left(\binom{r+n}{n} \cdot r + \varepsilon \right) h_k(x) + O(1).$$

To estimate the factor in front of $m(D,x)$, we have

$$\binom{r-kd+n}{n} = \frac{(r-kd+n)\cdots(r-kd+1)}{n!} = \frac{(r-kd)^n + O_n((r-kd)^{n-1})}{n!}$$

and therefore the coefficient in front of the proximity term is

$$\sum_{k=1}^{[r/d]} \binom{r-kd+n}{n} = \frac{(r-d)^n + \cdots + (r-[r/d]d)^n + O_{n,d}(r^n)}{n!}.$$

As $r \to \infty$, the ratio of this coefficient to the one in front of the height term now converges to

$$\frac{\sum_{k=1}^{[r/d]}\binom{r-kd+n}{n}}{r\binom{r+n}{n}} \approx \frac{\frac{r^{n+1}}{(n+1)d\cdot n!}}{\frac{r^{n+1}}{n!}} \to \frac{1}{(n+1)d}. \tag{72}$$

This indeed gives a nontrivial inequality

$$\frac{1}{d}m_S(D,x)+\cdots \le (n+1+\varepsilon)h_k(x)+O(1). \tag{73}$$

If $d=1$ then this is best possible (but of course is not new, since then D is already a hyperplane in $\mathbb{P}^n$). If $d>1$ then it is less than ideal, but is still new and noteworthy.

If D is the only component in the divisor, then (73) will never lead to a useful inequality, however, since the left-hand side is always bounded by $h_k(x)$. This approach only works if the divisor in question has more than one irreducible component. Having more than one component in the divisor, however, introduces some additional complications.

Suppose, for example, that there are two divisor components D_1 and D_2, and their images under the r-uple embedding span linear subspaces L_1 and L_2 of codimensions ρ_1 and ρ_2, respectively. We have

$$\operatorname{codim}(L_1\cap L_2)\le \rho_1+\rho_2 \tag{74}$$

(assuming $L_1\cap L_2\ne\emptyset$). If this inequality is strict then this causes problems. Indeed, let y denote the image of x under the d-uple embedding. If y is close to L_1 at some place $v\in S$, and also close to L_2 at that same place, then it is necessarily close to $L_1\cap L_2$. If $L_1\cap L_2$ is too large, though, then we will not be able to choose enough hyperplanes containing it to fully utilize both $m(D_1,x)$ and $m(D_2,x)$.

Indeed, choose ρ_1 generic hyperplanes containing L_1 and ρ_2 generic hyperplanes containing L_2. If these $\rho_1+\rho_2$ hyperplanes are in general position then this implies

$$\operatorname{codim}(L_1\cap L_2)\ge \rho_1+\rho_2.$$

So if this inequality does not hold, then the $\rho_1+\rho_2$ hyperplanes cannot (collectively) be in general position, and the max on the left-hand side of (44) will not be as large as one would hope.

So, in order to apply the reasoning leading up to (73) independently for each irreducible component of the divisor, equality must hold in (74) for each pair of components. (Similar considerations also apply to triples of components, etc.)

However, the standard computation of Hilbert functions using short exact sequences gives (for sufficiently large n) that the codimension of the linear span of $D_1\cap D_2$ is

$$\binom{n-d_1+r}{n}+\binom{n-d_2+r}{n}-\binom{n-d_1-d_2+r}{n}.$$

This is too small by $\binom{n-d_1-d_2+r}{n}$. So, any use of this approach would have to take this into account, and would also have to incorporate the changes made in going from (71) to (72).

As noted earlier, the purpose of this section is not to actually prove anything, but merely to highlight the general idea, together with some of the stumbling blocks.

20 Schmidt's Subspace Theorem Implies Siegel's Theorem

One way to avoid the difficulties mentioned in the last section is to restrict to curves, since in that case irreducible divisors are just points, so they do not intersect. Of course, using Schmidt's theorem to imply Roth's theorem would not be interesting, since the latter is already a special case. However, if one restricts to a curve contained in projective space, then one can get a nontrivial result by applying the methods of Sect. 19. This was done by Corvaja and Zannier [13], and gave a new proof of Siegel's theorem.

Theorem 20.1. (Siegel) *Let C be a smooth affine curve over a number field k. Assume that C has at least 3 points at infinity (i.e., at least three points need to be added to obtain a nonsingular projective curve). Then all sets of integral points on C are finite.*

Proof. By expanding k, if necessary, we may assume that the points at infinity are k-rational. Let $\overline{C}$ be the nonsingular projective closure of C, let g be its genus, let $Q_1,\dots,Q_r$ be the points at infinity, and let D be the divisor $Q_1+\dots+Q_r$. Pick N large, and embed $\overline{C}$ into $\mathbb{P}^{M-1}$ by the complete linear system $|ND|$; we have $M=Nr+1-g$. Assume that $\{P_1,P_2,\dots\}$ is an infinite S-integral set of points on C, for some finite set $S\supseteq S_\infty$ of places of k. After passing to an infinite subsequence, we may assume that for each $v\in S$ there is an index $j(v)\in\{1,\dots,r\}$ such that each P_i is at least as close to $Q_{j(v)}$ as to any other Q_j in the v-adic topology.

For all $\ell\in\mathbb{N}$, we have $h^0(\overline{C},\mathscr{O}(ND-\ell Q_j))\ge Nr-\ell+1-g$, so we can choose $Nr-\ell+1-g$ linearly independent hyperplanes in $\mathbb{P}^{M-1}$ vanishing to order $\ge\ell$ at Q_j. For each $v\in S$, do this with $j=j(v)$, obtaining one hyperplane vanishing to order $Nr-g$, a second vanishing to order $Nr-g-1$, etc. Obtaining M hyperplanes in this way for each v, and applying Schmidt's Subspace Theorem, we obtain

$$\sum_{v\in S}\sum_{\ell=0}^{Nr-g}\ell\lambda_{Q_{j(v)},v}(P_i)\le(M+\varepsilon)h_k(P_i)+O(1) \tag{75}$$

outside of a finite union of proper linear subspaces of $\mathbb{P}^{M-1}$. Here the height $h_k(P_i)$ is taken in $\mathbb{P}^{M-1}$. The finitely many linear subspaces correspond to only finitely many points on C, and they can be removed from the set of integral points.

By the assumption on the distance from P_i to $Q_{j(v)}$ (and the fact that the points Q_j are separated by a distance independent of i), we have $\lambda_{Q_j,v}(P_i)=O(1)$ for all $j\ne j(v)$ and all $v\in S$. Therefore

$$\lambda_{Q_{j(v)},v}(P_i)=\lambda_{D,v}(P_i)+O(1)$$

for all $v \in S$, with the constant independent of i. Also, since the embedding in $\mathbb{P}^{M-1}$ is obtained from the complete linear system $|ND|$, we have $h_k(P_i) = h_{ND,k}(P_i) + O(1)$. Therefore (75) becomes

$$\frac{(Nr-g)(Nr-g+1)}{2} m_S(D,P_i) \le N(Nr-g+1+\varepsilon)h_{D,k}(P_i) + O(1).$$

Since the P_i are integral points, though, we have $m_S(D,P_i) = h_{D,k}(P_i) + O(1)$, and the inequality becomes

$$\left(\frac{(Nr-g)(Nr-g+1)}{2} - N(Nr-g+1) - N\varepsilon\right) h_{D,k}(P_i) \le O(1).$$

If N is large and ε is small, then the quantity in parentheses is negative (since $r \ge 2$), leading to a contradiction since D is ample. □

Of course, if $g \ge 1$ then Siegel only required $r > 0$. This can be proved by reducing to the above case. Indeed, embed $\overline{C}$ in its Jacobian and pull back by multiplication by 2. This gives an étale cover of $\overline{C}$ of degree at least 4, so the pull-back of D will have at least three points. Integral points on C will pull back to integral points on the pull-back of C in the étale cover, by the Chevalley-Weil theorem for integral points (see the end of Sect. 13).

21 The Corvaja-Zannier Method in Higher Dimensions

Corvaja and Zannier further developed their method to higher dimensions; see for example [14]. It did not provide the full strength of Conjecture 15.6, even when $X = \mathbb{P}^n$, but it did provide noteworthy new answers. The key to their method can be summarized in the following definition, which is due to Levin [51].

Definition 21.1. Let X be a nonsingular complete variety over a field. A divisor D on X is **very large** if D is effective and, for all $P \in X$, there is a basis B of $L(D)$ such that

$$\sum_{f \in B} \operatorname{ord}_E f > 0 \tag{76}$$

for all irreducible components E of D passing through P. A divisor D is **large** if it is effective and has the same support as some very large divisor.

In the following discussion, it will be useful to have the following functoriality property of large divisors.

Proposition 21.2. *Let X' and X be nonsingular complete varieties over fields L and k, respectively, with $L \supseteq k$, and let $\phi: X' \to X$ be a morphism of schemes such that the diagram*

$$\begin{array}{ccc} X' & \xrightarrow{\phi} & X \\ \downarrow & & \downarrow \\ \operatorname{Spec} L & \longrightarrow & \operatorname{Spec} k \end{array}$$

commutes. Let D be a very large divisor on X, and let $D' = \phi^* D$ be the corresponding divisor on X'. Assume that the natural map

$$\alpha\colon H^0(X, \mathscr{O}(D)) \otimes_k L \to H^0(X', \mathscr{O}(D')) \tag{77}$$

is an isomorphism. Then D' is very large.

Proof. Let P' be a point on X', and let B be a basis for $L(D)$ that satisfies (76) for D at $\phi(P') \in X$. We have a commutative diagram

$$\begin{array}{ccc} L(D) \otimes_k L & \xrightarrow{(\cdot \mathbf{1}_D) \otimes_k L} & H^0(X, \mathscr{O}(D)) \otimes_k L \\ \downarrow \beta & & \downarrow \alpha \\ L(D') & \xrightarrow{\cdot \mathbf{1}_{D'}} & H^0(X', \mathscr{O}(D')) \end{array}$$

in which β is an isomorphism because the other three arrows are isomorphisms. Therefore we let $B' = \{\beta(f \otimes 1) : f \in B\}$; it is a basis of $L(D')$.

Now let E' be an irreducible component of D' passing through P'. For each irreducible component E of D passing through $\phi(P')$, let n_E be the multiplicity of E' in $\phi^* E$. For all nonzero $f \in L(D)$, we have

$$\operatorname{ord}_{E'} \beta(f \otimes 1) \ge \sum_E n_E \operatorname{ord}_E f, \tag{78}$$

where the sum is over all irreducible components E of D passing through $\phi(P')$ (and in particular it includes all irreducible components of D containing $\phi(E')$). Indeed, to verify (78), note that if s is a nonzero element of $H^0(X, \mathscr{O}(D))$, then $\operatorname{ord}_{E'} \alpha(s \otimes 1) \ge \sum_E n_E \operatorname{ord}_E s$, with equality if $(s) = D$. (Strictness may arise if E' is exceptional for ϕ and s vanishes along prime divisors containing $\phi(E')$ that do not occur in D.)

By (78), we then have

$$\sum_{f' \in B'} \operatorname{ord}_{E'} f' \ge \sum_E n_E \sum_{f \in B} \operatorname{ord}_E f > 0$$

(since at least one n_E is strictly positive). Thus B' satisfies (76) for E'. □

Corollary 21.3. *Let X be a nonsingular complete variety over a field k, let D be a divisor on X, let L be a field containing k, let $X_L = X \times_k L$ with projection $\phi\colon X_L \to X$, and let $D_L = \phi^* D$. If D is large (resp. very large), then so is D_L.*

Proof. Indeed, we may assume that D is very large, and note that (77) is an isomorphism because L is flat over k [36, III Prop. 9.3]. □

Corollary 21.4. *Let $\phi\colon X' \to X$ be a proper birational morphism*[2] *of nonsingular complete varieties over a field, and let D be a divisor on X. If D is large (resp. very large), then ϕ^*D is a large (resp. very large) divisor on X'.*

Proof. Again, we may assume that D is very large. In this case (77) is an isomorphism because $\phi_*\mathscr{O}_{X'} = \mathscr{O}_X$; see [36, proof of III Cor. 11.4 and III Remark 8.8.1]. □

Remark 21.5. More generally, for a linear subspace V of $L(D)$, one may define V-very large. For this definition, Proposition 21.2 holds with the weaker assumption that (77) is injective.

We also note that the definition of largeness is (vacuously) true for $D = 0$.

Having discussed the definition of large divisor, the theorem that makes the definition useful is the following.

Theorem 21.6. (Corvaja-Zannier) *Let k be a number field, let $S \supseteq S_\infty$ be a finite set of places of k, let X be a nonsingular complete variety over k, and let D be a nonzero large divisor on X. Then any set of (D,S)-integral points on X is not Zariski-dense.*

Proof. By Proposition 13.2, we may assume that D is very large.

Since $D \ne 0$, there is a component E as in the definition of very large, so $\ell(D) > 1$. Therefore there is a nontrivial rational map $\Phi\colon X \dashrightarrow \mathbb{P}_k^{\ell(D)-1}$.

We may assume that Φ is a morphism. Indeed, let X' be a desingularization of the closure of the graph of Φ. Replace X with X' and D with its pull-back. By Corollary 21.4 the pull-back remains very large. Moreover, the notion of integral set of points remains unchanged, by functoriality of Weil functions (or by the fact that the desingularization may be chosen such that the map $X' \to X$ is an isomorphism away from the support of D).

Now suppose that there is a Zariski-dense set $\{P_i\}$ of integral points. After passing to a Zariski-dense subset, we may assume that for each $v \in S$ there is a point $P_v \in X(k_v)$ such that the P_i converge to P_v in the v-topology. For each such v, let $B_v = \{f_{v,1},\dots,f_{v,\ell(D)}\}$ be a basis for $L(D)$ that satisfies (76) at the point P_v. Let $H_{v,1},\dots,H_{v,\ell(D)}$ be the corresponding hyperplanes in $\mathbb{P}_k^{\ell(D)-1}$. (Corollary 21.3 is not really needed here, but can be used if the reader prefers. When finding P_v, one should think of P_v and the P_i as points on the (complex, real, or p-adic) manifold $X(k_v)$, and when defining B_v one should realize that P_v is a morphism from $\operatorname{Spec} k_v$ to X, whose image is a point in X (which may be the generic point).)

To obtain a contradiction, it will suffice to find an $\varepsilon > 0$ such that

$$\sum_{v\in S}\sum_{j=1}^{\ell(D)} \lambda_{H_{v,j},v}(\Phi(P_i)) \ge (\ell(D)+\varepsilon)h_k(\Phi(P_i)) + O(1) \tag{79}$$

[2] This requires ϕ to be a morphism, not just a rational map.

for all i. Indeed, Schmidt's Subspace Theorem would then imply that the $\Phi(P_i)$ are contained in a finite union of proper linear subspaces of $\mathbb{P}_k^{\ell(D)-1}$, contradicting the fact that the P_i are Zariski-dense in X and that $\Phi(X)$ is not contained in any proper linear subspace of $\mathbb{P}_k^{\ell(D)-1}$.

Let μ be the largest multiplicity of a component of D, and let $\varepsilon = 1/\mu$, so that $\varepsilon \operatorname{ord}_E D \le 1$ for all irreducible components E of D.

Let $v \in S$, and let E be an irreducible component of D. First suppose that E contains P_v. Then $\sum \operatorname{ord}_E \Phi^* H_{v,j} > \ell(D) \operatorname{ord}_E D$ (by (76)), so

$$\sum_{j=1}^{\ell(D)} \operatorname{ord}_E \Phi^* H_{v,j} \ge \ell(D) \operatorname{ord}_E D + 1 \ge (\ell(D)+\varepsilon) \operatorname{ord}_E D$$

and therefore

$$\sum_{j=1}^{\ell(D)} (\operatorname{ord}_E \Phi^* H_{v,j}) \lambda_{E,v}(P_i) \ge (\ell(D)+\varepsilon)(\operatorname{ord}_E D) \lambda_{E,v}(P_i) + O(1)$$

since $\lambda_{E,v}(P_i) \ge O(1)$. This latter inequality also holds if E does not contain P_v, since in that case $\lambda_{E,v}(P_i) = O(1)$. Therefore, we have

$$\begin{aligned}
\sum_{j=1}^{\ell(D)} \lambda_{H_{v,j},v}(\Phi(P_i)) &\ge \sum_{j=1}^{\ell(D)} \sum_E (\operatorname{ord}_E \Phi^* H_{v,j}) \lambda_{E,v}(P_i) + O(1) \\
&\ge (\ell(D)+\varepsilon) \sum_E (\operatorname{ord}_E D) \lambda_{E,v}(P_i) + O(1) \\
&= (\ell(D)+\varepsilon) \lambda_{D,v}(P_i) + O(1),
\end{aligned}$$

where the sums over E are sums over all irreducible components E of D.

Summing over $v \in S$ then gives

$$\sum_{v\in S} \sum_{j=1}^{\ell(D)} \lambda_{H_{v,j},v}(\Phi(P_i)) \ge (\ell(D)+\varepsilon) m_S(D,P_i) + O(1) = (\ell(D)+\varepsilon) h_{D,k}(P_i) + O(1)$$

since the P_i are (D,S)-integral. This is equivalent to (79) by functoriality of heights. □

This proof does not carry over directly to Nevanlinna theory, because it relies on the finiteness of S; moreover, the whole idea of passing to a subsequence is unsuited to Nevanlinna theory. In fact, the proof does not even work for function fields over infinite fields, since in such cases the local fields are not locally compact.

The following lemma, essentially due to Levin [51], works around this problem. The key idea is that it suffices to use only finitely many bases in Definition 21.1.

Lemma 21.7. *Let X be a nonsingular complete variety over $\mathbb{C}$, and let D be an effective divisor on X. Let σ_0 be the set of prime divisors occurring in D, and let Σ be the set of subsets σ of σ_0 for which $\bigcap_{E\in\sigma} E$ is nonempty. For each $\sigma \in \Sigma$ let D_σ be the sum of those components of D* not *lying in σ, with the same multiplicities as they have in D. Pick a Weil function for each such D_σ. Then there is a constant C, depending only on X and D, such that*

$$\min_{\sigma\in\Sigma} \lambda_{D_\sigma}(P) \le C \tag{80}$$

for all $P \in X(\mathbb{C})$.

Proof. The conditions imply that

$$\bigcap_{\sigma\in\Sigma} \operatorname{Supp} D_\sigma = \emptyset,$$

since for all $P \in X$ the set $\sigma := \{E \in \sigma_0 : E \ni P\}$ is an element of Σ, and then $P \notin \operatorname{Supp} D_\sigma$. The lemma then follows from Lemma 9.9, since Σ is a finite set. □

Lemma 21.8. *Let X be a nonsingular complete variety over $\mathbb{C}$, let D be a very large divisor on X whose complete linear system is base-point-free, and let $\Phi\colon X \to \mathbb{P}^{\ell(D)-1}$ be a corresponding morphism to projective space. Then there is a finite collection $H_1,\dots,H_q$ of hyperplanes and $\varepsilon > 0$ such that, given choices $\lambda_{H_1},\dots,\lambda_{H_q}$ and λ_D of Weil functions on $\mathbb{P}^{\ell(D)-1}$ and X, respectively, we have*

$$\max_J \sum_{j\in J} \lambda_{H_j}(\Phi(P)) \ge (\ell(D)+\varepsilon)\lambda_D(P) + O(1), \tag{81}$$

where the implicit constant in $O(1)$ is independent of $P \in X(\mathbb{C})$. Here, as in Theorem 8.11, J varies over all subsets of $\{1,\dots,q\}$ corresponding to subsets of $\{H_1,\dots,H_q\}$ that lie in general position.

Proof. Let Σ be the (finite) set of Lemma 21.7, and for each $\sigma \in \Sigma$ let B_σ be a basis for $L(D)$ that satisfies (76) at some (and hence all) points $P \in \bigcap_{E\in\sigma} E$. Let $H_1,\dots,H_q$ be the distinct hyperplanes in $\mathbb{P}^{\ell(D)-1}$ corresponding to elements of the union $\bigcup_{\sigma\in\Sigma} B_\sigma$, and choose Weil functions λ_{H_j} for them. For each $\sigma \in \Sigma$ let D_σ be as in Lemma 21.7, and choose a Weil function λ_{D_σ} for it. Let C be a constant that satisfies (80). Finally, choose Weil functions λ_E for each prime divisor E occurring in D.

Let μ be the largest multiplicity of a component of D, and let $\varepsilon = 1/\mu$.

Now let $P \in X(\mathbb{C})$. Pick $\sigma \in \Sigma$ for which

$$\lambda_{D_\sigma} \le C, \tag{82}$$

and let $J \subseteq \{1,\dots,q\}$ be the subset for which $\{H_j : j \in J\}$ are the hyperplanes corresponding to the elements of B_σ. As before, (76) applied to B_σ implies that $\sum_{j\in J} \operatorname{ord}_E \Phi^* H_j > \ell(D) \operatorname{ord}_E D$ for all $E \in \sigma$; hence

$$\sum_{j \in J} \operatorname{ord}_E \Phi^* H_j \geq (\ell(D) + \varepsilon) \operatorname{ord}_E D,$$

and therefore

$$\sum_{j \in J} (\operatorname{ord}_E \Phi^* H_j) \lambda_E(P) \geq (\ell(D) + \varepsilon)(\operatorname{ord}_E D) \lambda_E(P) + O(1) \tag{83}$$

since $\lambda_E(P) \geq O(1)$. Also

$$D = D_\sigma + \sum_{E \in \sigma} (\operatorname{ord}_E D) \cdot E. \tag{84}$$

By (83), (82), and (84), we then have

$$\begin{aligned} \sum_{j \in J} \lambda_{H_j}(\Phi(P)) &\geq \sum_{j \in J} \sum_{E \in \sigma} (\operatorname{ord}_E \Phi^* H_j) \lambda_E(P) + O(1) \\ &\geq (\ell(D) + \varepsilon)(\lambda_{D_\sigma}(P) - C) + (\ell(D) + \varepsilon) \sum_{E \in \sigma} (\operatorname{ord}_E D) \lambda_E(P) + O(1) \\ &= (\ell(D) + \varepsilon) \lambda_D(P) + O(1). \end{aligned}$$

In the above, the constants in $O(1)$ depend only on the choices of B_σ and the choices of Weil functions, and on σ (which has only finitely many choices). Since J is one of the sets in (81), the lemma then follows. □

This then leads to the theorem in Nevanlinna theory corresponding to Theorem 21.6:

Theorem 21.9. [51] *Let X be a nonsingular complete variety over $\mathbb{C}$, let D be a nonzero large divisor on X, and let $f\colon \mathbb{C} \to X$ be a holomorphic curve whose image is disjoint from D. Then the image of f is not Zariski dense.*

Proof. As in the proof of Theorem 21.6, we may assume that D is very large and base point free. Let $\Phi\colon X \to \mathbb{P}_{\mathbb{C}}^{\ell(D)-1}$ be a morphism corresponding to a complete linear system of D. Let $f\colon \mathbb{C} \to X$ be a holomorphic curve whose image does not meet D. Let $H_1, \ldots, H_q$ and ε be as in Lemma 21.8. By that lemma, we then have

$$\begin{aligned} \int_0^{2\pi} \max_J \sum_{j \in J} \lambda_{H_j}(\Phi(f(re^{i\theta}))) \frac{d\theta}{2\pi} &\geq (\ell(D) + \varepsilon) m_f(D, r) + O(1) \\ &= (\ell(D) + \varepsilon) T_{\Phi \circ f}(r) + O(1) \end{aligned}$$

for all $r > 0$. This contradicts Theorem 8.11 unless the image of $\Phi \circ f$ is contained in a proper linear subspace of $\mathbb{P}_{\mathbb{C}}^{\ell(D)-1}$ (since $\ell(D) > 1$). This in turn implies that the image of f cannot be Zariski dense. □

This proof can also be adapted back to the number field case, and it also works over function fields.

Some concrete examples of large divisors follow. First, in order to show that Theorem 20.1 is a consequence of Theorem 21.6, we have the following.

Proposition 21.10. *Let C be a smooth projective curve over a field k, and let D be an effective divisor supported on (distinct) rational points $Q_1,\dots,Q_r$. If $r=0$ or $r\ge 3$ then D is large.*

Proof. If $r=0$ then $D=0$, which is already known to be large.

Assume $r\ge 3$. It will suffice to show that if $D=N(Q_1+\dots+Q_r)$ then D is very large for sufficiently large integers N. As in the proof of Theorem 20.1, we have $h^0(C,\mathscr{O}(D-\ell Q_j))\ge Nr-\ell+1-g$ for all $\ell\in\mathbb{N}$ and all $j=1,\dots,r$, where g is the genus of C. For each such j there is a basis $(s_1,\dots,s_{Nr+1-g})$ of $H^0(C,\mathscr{O}(D))$ such that s_ℓ vanishes to order $\ge \ell-1$ at Q_j. Dividing each such s_ℓ by the canonical section 1_D then gives a basis $(f_1,\dots,f_{Nr+1-g})$ of $L(D)$ such that $\operatorname{ord}_{Q_j} f_\ell\ge \ell-1-N$ for all ℓ. Thus

$$\sum_{\ell=1}^{Nr+1-g}\operatorname{ord}_{Q_j} f_\ell \ge (Nr+1-g)\left(\frac{Nr-g}{2}-N\right)$$

if $N>(g-1)/r$, and is strictly positive if also $N>g/(r-2)$. □

Proposition 21.11. *Let X_1 and X_2 be smooth complete varieties over a field k, and let D_1 and D_2 be divisors on X_1 and X_2, respectively. If D_1 and D_2 are large (resp. very large), then $p_1^*D_1+p_2^*D_2$ is a large (resp. very large) divisor on $X_1\times_k X_2$, where $p_i\colon X_1\times_k X_2\to X_i$ is the projection ($i=1,2$).*

Proof. It will suffice to show that if D_1 and D_2 are very large, then so is $p_1^*D_1+p_2^*D_2$. Write $D=p_1^*D_1+p_2^*D_2$.

We first claim that the natural map

$$H^0(X_1,\mathscr{O}(D_1))\otimes_k H^0(X_2,\mathscr{O}(D_2))\longrightarrow H^0(X_1\times_k X_2,\mathscr{O}(D)) \tag{85}$$

is an isomorphism. Indeed, the projection formula [36, II Ex. 5.1d] gives an isomorphism

$$\mathscr{O}(D_1)\otimes_k H^0(X_2,\mathscr{O}(D_2))\xrightarrow{\sim}(p_1)_*\mathscr{O}(p_1^*D_1+p_2^*D_2),$$

of sheaves on X_1, and taking global sections gives (85).

To show that D is very large, let $P\in X_1\times_k X_2$. Let B_1 and B_2 be bases of $L(D_1)$ and $L(D_2)$ satisfying (76) with respect to the points $p_1(P)$ and $p_2(P)$, respectively. By (85), $\{p_1^*f_1\cdot p_2^*f_2 : f_1\in B_1,\ f_2\in B_2\}$ is a basis for $L(D)$; call it B. Let E be an irreducible component of D passing through P. If $E=p_1^*E_1$ for a component E_1 of D_1, then

$$\sum_{h\in B}\operatorname{ord}_E h=\ell(D_2)\sum_{f\in B_1}\operatorname{ord}_{E_1} f>0,$$

so (76) is satisfied for E. Otherwise, we must have $E=p_2^*E_2$ for an irreducible component E_2 of D_2, and (76) is satisfied for a symmetrical reason. Thus D is very large. □

The following gives a slightly more complicated example of large divisors. For this example, recall that a Cartier divisor D on a complete variety X over a field k is **nef** ("numerically effective") if $\deg j^*\mathscr{O}(D) \ge 0$ for all maps $j\colon C \to X$ from a curve C over k to X.

Theorem 21.12. [51, Thm. 9.2] *Let X be a nonsingular projective variety of dimension q, and let $D = \sum D_i$ be a divisor on X for which all D_i are effective and nef. Assume also that all irreducible components of D are nonsingular, and that $D^q > 2qD^{q-1}D_P$ for all $P \in \operatorname{Supp} D$, where $D_P = \sum_{\{i:D_i \ni P\}} D_i$. Then D is large.*

For the proof, see [51]. Note that this generalizes Proposition 21.10.

As another example of this method, we note another theorem of Levin.

Definition 21.13. A variety V over a number field k is **Mordellic** if for all number fields $L \supseteq k$ and all finite sets $S \supseteq S_\infty$ of places of L, there are no infinite sets of S-integral L-rational points on $V_L := V \times_k L$. A variety V over k is **quasi-Mordellic** if there is a proper Zariski-closed subset Z of V such that, for all L and S as above, and for all S-integral sets of L-rational points on V_L, almost all points in the set are contained in Z_L.

Theorem 21.14. [51, Thm. 9.11A] *Let X be a projective variety over a number field k. Let $D = \sum_{i=1}^r D_i$ be a divisor on X such that each D_i is an effective Cartier divisor, and the intersection of any $m+1$ of the supports of the D_i is empty. Then:*

(a) If D_i is big for all i and $r > 2m\dim X$ then $X \setminus D$ is quasi-Mordellic.
(b) If D_i is ample for all i and $r > 2m\dim X$ then $X \setminus D$ is Mordellic.

The proof of this theorem, as well as its counterpart in Nevanlinna theory, appear in [51].

Again, we note that if X is a nonsingular curve, then this reduces to the combination of Theorem 21.6 and Proposition 21.10.

22 Work of Evertse and Ferretti

Evertse and Ferretti also found a way of using Schmidt's Subspace Theorem in combination with d-uple embeddings to get partial results on more general varieties, with respect to more general divisors. Their method is based on using Mumford's degree of contact, which was originally developed to study moduli spaces, but which is also well suited for this application. It uses a bit more machinery than the method of Corvaja and Zannier, and this machinery makes direct comparisons more difficult.

Because of the machinery, we offer here only a sketch of the methods, without proofs.

The idea originated from a paper of Ferretti [25], and was further developed jointly with Evertse; see for example [20]. This work was translated into Nevanlinna theory by Ru [70], solving a conjecture of Shiffman.

Throughout this section, k is a field of characteristic 0 and $X \subseteq \mathbb{P}^N_k$ is a projective variety over k of dimension n and degree Δ.

Definition 22.1. The **Chow form** of X is the unique (up to scalar multiple) polynomial

$$F_X \in k[\mathbf{u}_0, \dots, \mathbf{u}_n] = k[u_{00}, \dots, u_{0N}, u_{10}, \dots, u_{nN}],$$

homogeneous of degree Δ in each block $\mathbf{u}_i$, characterized by the condition

$$F_X(\mathbf{u}_0, \dots, \mathbf{u}_n) = 0 \iff X \cap H_{\mathbf{u}_0} \cap \dots \cap H_{\mathbf{u}_n} \neq \emptyset,$$

where $H_{\mathbf{u}_i}$ is the hyperplane in $\mathbb{P}^N_k$ corresponding to $\mathbf{u}_i \in (\mathbb{P}^N_k)^*$.

For more details on Chow forms, see Hodge and Pedoe [38, Vol. II, Chap. X, Sect. 6–8].

Definition 22.2. Let $\mathbf{c} = (c_0, \dots, c_N) \in \mathbb{R}^{N+1}$ and let F_X be as above. For an indeterminate t, write

$$F_X(t^{c_0} u_{00}, \dots, t^{c_N} u_{0N}, \dots, t^{c_N} u_{nN}) = t^{e_0} G_0(\mathbf{u}_0, \dots, \mathbf{u}_n) + \dots + t^{e_r} G_r(\mathbf{u}_0, \dots, \mathbf{u}_n),$$

where $G_0, \dots, G_r$ are nonzero polynomials in $k[u_{00}, \dots, u_{nN}]$ and $e_0 > \dots > e_r$. Then the **Chow weight** of X with respect to $\mathbf{c}$ is $e_X(\mathbf{c}) = e_0$.

If I is the (prime) homogeneous ideal in $k[x_0, \dots, x_N]$ corresponding to $X \subseteq \mathbb{P}^N_k$, then recall that the **Hilbert function** $H_X(m)$ for $m \in \mathbb{N}$ is defined by

$$H_X(m) = \dim_k k[x_0, \dots, x_N]_m / I_m,$$

where the subscript m denotes the homogeneous part of degree m.

Recall also that the Hilbert polynomial of X (which agrees with the Hilbert function for $m \gg 0$) has leading term $\Delta m^n / n!$.

Definition 22.3. Let I be as above, and let $\mathbf{c} \in \mathbb{R}^{N+1}$. The the **Hilbert weight** of X with respect to $\mathbf{c}$ is

$$S_X(m, \mathbf{c}) = \max \left(\sum_{\ell=1}^{H_X(m)} \mathbf{a}_\ell \cdot \mathbf{c} \right),$$

where the max is taken over all collections $(\mathbf{a}_1, \dots, \mathbf{a}_{H_X(m)})$ with $\mathbf{a}_\ell \in \mathbb{N}^{N+1}$ for all ℓ, whose corresponding monomials $\mathbf{x}^{\mathbf{a}_1}, \dots, \mathbf{x}^{\mathbf{a}_{H_X(m)}}$ give a basis (over k) when mapped to $k[x_0, \dots, x_N]_m / I_m$. (Here $\mathbf{x}^{\mathbf{a}_\ell}$ denotes $x_0^{a_{\ell 0}} \cdots x_N^{a_{\ell N}}$, and the conditions necessarily imply that $a_{\ell 0} + \dots + a_{\ell N} = m$ for all ℓ.)

Mumford showed that

$$S_X(m,\mathbf{c}) = e_X(\mathbf{c}) \cdot \frac{m^{n+1}}{(n+1)!} + O(m^n),$$

and Evertse and Ferretti showed further that if $m > \Delta$ then

$$\frac{S_X(m,\mathbf{c})}{H_X(m)} \ge \frac{m}{(n+1)\Delta} e_X(\mathbf{c}) - (2n+1)\Delta \max_{0\le j\le N} c_j \tag{86}$$

[20, Prop. 3.2]. (To compare these two inequalities, note that $H_X(m) \sim \Delta m^n/n!$.)

In diophantine applications, k is a number field and $S \supseteq S_\infty$ is a finite set of places of k. The following is a slight simplification of the main theorem of Evertse and Ferretti [20].

Theorem 22.4. *Assume that* $n = \dim X > 0$. *For each* $v \in S$ *let* $D_0^{(v)},\dots,D_n^{(v)}$ *be a system of effective divisors on* $\mathbb{P}^N_k$ *satisfying*

$$X \cap \bigcap_{j=0}^{n} \operatorname{Supp} D_j^{(v)} = \emptyset.$$

Then for all ε *with* $0 < \varepsilon \le 1$, *there are hypersurfaces* $G_1,\dots,G_u$ *in* $\mathbb{P}^N$, *not containing* X *and of degree*

$$\deg G_i \le 2(n+1)(2n+1)(n+2)\Delta d^{n+1}\varepsilon^{-1}, \tag{87}$$

where d *is the least common multiple of the degrees of the* $D_j^{(v)}$, *such that all solutions* $x \in X(k)$ *of the inequality*

$$\sum_{v\in S}\sum_{j=0}^{n} \frac{\lambda_{D_j^{(v)},v}(x)}{\deg D_j^{(v)}} \ge (n+1+\varepsilon)h_k(x) + O(1) \tag{88}$$

lie in the union of the G_i. *In particular, these solutions are not Zariski-dense.*

(Evertse and Ferretti also prove a more quantitative version of this theorem, which gives explicit bounds on u, if one ignores solutions of (88) of height below a given explicit bound. They obtain a weaker bound than (87), because of this added strength.)

Note that this result is weaker than Conjecture 15.6, since the latter conjecture does not divide the Weil functions by the degrees of the divisors. It is also stronger, though, in the sense that the sum of the divisors does not have to have normal crossings.

Here we will restate this theorem in a way that translates more readily into Nevanlinna theory. The proof will roughly follow Evertse and Ferretti [20], with substantial simplifications since we are not bounding u. In particular, the "twisted heights" (which are related to the first successive minima in Schmidt's original proof) are not needed here. Because of these simplifications, it may be easier to follow Ru [70], even though he is working in Nevanlinna theory.

Theorem 22.5. *Assume that $n = \dim X > 0$. Let $D_0, \dots, D_q$ be effective divisors on $\mathbb{P}^N_k$, whose supports do not contain X. Let $\mathscr{J}$ be the set of all $(n+1)$-element subsets J of $\{0, \dots, q\}$ for which*

$$X \cap \bigcap_{j \in J} \operatorname{Supp} D_j = \emptyset, \tag{89}$$

and assume that $\mathscr{J}$ is not empty. Then for all ε with $0 < \varepsilon \le 1$, all constants $C \in \mathbb{R}$, and all choices of Weil functions $\lambda_{D_j,v}$, there are hypersurfaces $G_1, \dots, G_u$, as before, such that all solutions $x \in X(k)$ of the inequality

$$\sum_{v \in S} \max_{J \in \mathscr{J}} \sum_{j \in J} \frac{\lambda_{D_j,v}(x)}{\deg D_j} \ge (n+1+\varepsilon) h_k(x) + C \tag{90}$$

lie in the union of the G_i.

Proof (sketch). First, by replacing each D_j with a suitable positive integer multiple, we may assume that all of the D_j have the same degree d.

Next, we reduce to $d = 1$, as follows. Let $\phi \colon \mathbb{P}^N_k \to \mathbb{P}^M_k$ be the d-uple embedding, where $M = \binom{N+d}{N} - 1$, and let $Y = \phi(X)$. Then Y has dimension n, degree Δd^n, and ϕ multiplies the projective height by d. Moreover, there are hyperplanes $E_0, \dots, E_q$ on $\mathbb{P}^M_k$ such that $\phi^* E_j = D_j$ for all j. Thus if $y = \phi(x)$ then (90) is equivalent to

$$\sum_{v \in S} \max_{J} \sum_{j \in J} \lambda_{E_j,v}(y) \ge (n+1+\varepsilon) h_k(y) + C' \tag{91}$$

for a suitable constant C' independent of x. Applying Theorem 22.5 with $d = 1$ to Y and $E_0, \dots, E_q$ then gives hypersurfaces $G'_1, \dots, G'_u$ in $\mathbb{P}^M_k$, of degrees bounded by $(8n+6)(n+2)^2 \Delta d^n \varepsilon^{-1}$, not containing Y, but containing all solutions $y \in Y(k)$ of the inequality (91). Pulling these hypersurfaces back to $\mathbb{P}^N_k$ multiplies their degrees by d, so these pull-backs satisfy (91).

By a further linear embedding of $\mathbb{P}^N_k$, we may assume that $D_0, \dots, D_q$ are the coordinate hyperplanes $x_0 = 0, \dots, x_q = 0$, respectively. We may also assume that all of the Weil functions occurring in (90) are nonnegative.

Now assume, by way of contradiction, that the set of solutions of (90) is not contained in a finite union of hypersurfaces G_i satisfying (87). By a partitioning argument [20, Lemma 5.3] there is a subset Σ of $X(k)$, not contained in a finite union of hypersurfaces G_i as before, $(n+1)$-element subsets J_v of $\{0, \dots, q\}$ satisfying (89) for each $v \in S$, and nonnegative real constants $c_{j,v}$ for all $v \in S$ and all $j \in J_v$, such that

$$\sum_{v \in S} \sum_{j \in J_v} c_{j,v} = 1 \tag{92}$$

and such that the inequality

$$\lambda_{D_j,v}(x) \geq c_{j,v}\left(n+1+\frac{\varepsilon}{2}\right)h_k(x) \tag{93}$$

holds for all $v \in S$, all $j \in J_v$, and all $x \in \Sigma$. Also let $c_{j,v} = 0$ if $v \in S$ and $j \notin J_v$, so that (93) holds for all $v \in S$ and all $j = 0,\ldots,q$.

Now for some $m > \Delta$ (its exact value is given by (103)), let

$$\phi_m\colon \mathbb{P}^N_k \to \mathbb{P}^{R_m}_k$$

be the m-uple embedding; here $R_m = \binom{N+m}{m} - 1$. Let X_m be the linear subspace of $\mathbb{P}^{R_m}_k$ spanned by $\phi_m(X)$. We have

$$\dim X_m = H_X(m) - 1.$$

For each $v \in S$ let

$$\mathbf{c}_v = (c_{0,v},\ldots,c_{q,v},0,\ldots,0) \in \mathbb{R}^{N+1}_{\geq 0},$$

and let $\mathbf{a}_{1,v},\ldots,\mathbf{a}_{H_X(m),v}$ be the elements of $\mathbb{N}^{N+1}$ for which the monomials $\mathbf{x}^{\mathbf{a}_{\ell,v}}$, $\ell = 1,\ldots,H_X(m)$, give a basis for $k[x_0,\ldots,x_N]_m/I_m$ satisfying

$$S_X(m,\mathbf{c}_v) = \sum_{\ell=1}^{H_X(m)} \mathbf{a}_{\ell,v}\cdot\mathbf{c}_v. \tag{94}$$

For each v the monomials $\mathbf{x}^{\mathbf{a}_{\ell,v}}$, $\ell = 1,\ldots,H_X(m)$, define linear forms $L_{\ell,v}$ in the homogeneous coordinates on $\mathbb{P}^{R_m}_k$ which are linearly independent on $\phi_m(X)$, and therefore on X_m. For each v and ℓ choose Weil functions $\lambda_{L_{\ell,v}}$ on $\mathbb{P}^{R_m}_k$. We have

$$\lambda_{L_{\ell,v},v}(\phi_m(x)) \geq \sum_{j=0}^{q} a_{\ell,v,j}\lambda_{D_j,v}(x) + O(1)$$

for all v and ℓ. After adjusting the Weil function, we may assume that the $O(1)$ term is not necessary.

By (93) and (94), we then have

$$\sum_{\ell=1}^{H_X(m)} \lambda_{L_{\ell,v},v}(\phi_m(x)) \geq S_X(m,\mathbf{c}_v)\left(n+1+\frac{\varepsilon}{2}\right)h_k(x) \tag{95}$$

for all $x \in \Sigma$ and all $v \in S$. Assume for now that there is an integer $m \geq \Delta$ such that

$$m \leq \frac{2(n+1)(2n+1)(n+2)\Delta}{\varepsilon} \tag{96}$$

and such that

$$\left(n+1+\frac{\varepsilon}{2}\right)\sum_{v\in S} S_X(m,\mathbf{c}_v) > mH_X(m). \tag{97}$$

Then, for sufficiently small $\varepsilon' > 0$, (95) will imply

$$\sum_{v\in S}\sum_{\ell=1}^{H_X(m)} \lambda_{L_{\ell,v},v}(\phi_m(x)) \ge (H_m(x)+\varepsilon')mh_k(x)+O(1) \tag{98}$$

for all $x\in\Sigma$. Note that $h_k(\phi_m(x)) = mh_k(x)+O(1)$. Applying Schmidt's Subspace Theorem to X_m (via some chosen isomorphism $X_m \cong \mathbb{P}_k^{H_X(m)-1}$) it follows that there is a finite union of hyperplanes in X_m, and hence in $\mathbb{P}_k^{R_m}$, containing $\phi_m(\Sigma)$. These pull back to give homogeneous polynomials G_i of degree m on $\mathbb{P}_k^N$; they satisfy (87) by (96).

We now show that (97) holds for some m satisfying (96).

By (86), we have

$$\sum_{v\in S} S_X(m,\mathbf{c}_v) \ge H_X(m)\sum_{v\in S}\left(\frac{m}{(n+1)\Delta}e_X(\mathbf{c}_v) - (2n+1)\Delta\max_{j\in J_v} c_{j,v}\right). \tag{99}$$

By [20, Lemma 5.1], we have

$$e_X(\mathbf{c}_v) \ge \Delta\sum_{j\in J_v} c_{j,v}, \tag{100}$$

and therefore

$$\sum_{v\in S} e_X(\mathbf{c}_v) \ge \Delta\sum_{v\in S}\sum_{j\in J_v} c_{j,v} = \Delta \tag{101}$$

by (92). Thus (99) becomes

$$\begin{aligned}\sum_{v\in S} S_X(m,\mathbf{c}_v) &\ge H_X(m)\left(\frac{m}{n+1} - (2n+1)\Delta\sum_{v\in S}\max_{j\in J_v} c_{j,v}\right)\\ &\ge H_X(m)\left(\frac{m}{n+1}-(2n+1)\Delta\right).\end{aligned} \tag{102}$$

Now let

$$m = \left\lfloor \frac{2(n+1)(2n+1)(n+2)\Delta}{\varepsilon}\right\rfloor. \tag{103}$$

This clearly satisfies (96); in addition, we have

$$m < \left(n+1+\frac{\varepsilon}{2}\right)\left(\frac{m}{n+1}-(2n+1)\Delta\right). \tag{104}$$

Thus (102) becomes

$$\left(n+1+\frac{\varepsilon}{2}\right)\sum_{v\in S}S_X(m,\mathbf{c}_v) > mH_X(m),$$

which is (97). □

In Nevanlinna theory, the counterpart to Theorem 22.5 was proved by Ru [70]. Here we give a slightly stronger version of his theorem: X is not required to be nonsingular, and we incorporate the set $\mathscr{J}$. This stronger version still follows from his proof without essential changes, though.

Theorem 22.6. *Let $k=\mathbb{C}$ and assume that $n=\dim X>0$. Let $D_0,\dots,D_q$ be effective divisors on $\mathbb{P}^N_{\mathbb{C}}$, whose supports do not contain X. Let $\mathscr{J}$ be the set of all $(n+1)$-element subsets J of $\{0,\dots,q\}$ for which*

$$X\cap\bigcap_{j\in J}\operatorname{Supp}D_j=\emptyset,$$

and assume that $\mathscr{J}$ is not empty. Fix $\varepsilon\in\mathbb{R}$ with $0<\varepsilon\le 1$, fix $C\in\mathbb{R}$, and choose Weil functions λ_{D_j} for all j. Let $f\colon\mathbb{C}\to X(\mathbb{C})$ be a holomorphic curve whose image is not contained in any hypersurface in $\mathbb{P}^N$ not containing X of degree $\le 2(n+1)(2n+1)(n+2)\Delta d^{n+1}\varepsilon^{-1}$. Then

$$\int_0^{2\pi}\max_{J\in\mathscr{J}}\sum_{j\in J}\frac{\lambda_{D_j}(f(re^{i\theta}))}{\deg D_j}\frac{d\theta}{2\pi}\le_{\text{exc}}(n+1+\varepsilon)T_f(r)+C. \tag{105}$$

Proof (sketch). This proof uses the same general outline as the proof of Theorem 22.5, but there is an essential difference in that one cannot take a subsequence in order to define constants $c_{j,v}$, since the interval $[0,2\pi]$ is not a finite set. Instead, however, it is possible to drop the condition (92); then (101) is no longer valid. However, (100) still holds, and is homogeneous in the components of $\mathbf{c}$. Therefore, we may omit the step of dividing by the height, and just let the components of $\mathbf{c}$ be the Weil functions themselves (assumed nonnegative).

In detail, as before we assume that $D_0,\dots,D_q$ are restrictions of the coordinate hyperplanes $x_0=0$, …, $x_q=0$, and that the Weil functions λ_{D_j} are nonnegative.

For each $r>0$ and $\theta\in[0,2\pi]$ let $J_{r,\theta}$ be an element of $\mathscr{J}$ for which

$$\sum_{j\in J_{r,\theta}}\lambda_{D_j}(f(re^{i\theta}))$$

is maximal, for each $j\in J_{r,\theta}$ let

$$c_{j,r,\theta}=\lambda_{D_j}(f(re^{i\theta})),$$

and for each $j\in\{0,\dots,N\}\setminus J_{r,\theta}$ let $c_{j,r,\theta}=0$. Let

$$\mathbf{c}_{r,\theta}=(c_{0,r,\theta},\dots,c_{N,r,\theta})\in\mathbb{R}^{N+1}_{\ge 0}.$$

Then, as before, [20, Lemma 5.1] gives

$$e_X(\mathbf{c}_{r,\theta}) \ge \Delta \sum_{j\in J_{r,\theta}} c_{j,r,\theta}. \tag{106}$$

Let m be as in (103). By (86), (106), and nonnegativity of $c_{j,r,\theta}$, we have

$$\begin{aligned} \frac{1}{H_X(m)} \int_0^{2\pi} S_X(m,\mathbf{c}_{r,\theta})\,\frac{d\theta}{2\pi} &\ge \int_0^{2\pi} \left(\frac{m}{(n+1)\Delta} e_X(\mathbf{c}_{r,\theta}) - (2n+1)\Delta \max_{j\in J_{r,\theta}} c_{j,r,\theta}\right)\frac{d\theta}{2\pi} \\ &\ge \int_0^{2\pi} \left(\frac{m}{n+1} \sum_{j\in J_{r,\theta}} c_{j,r,\theta} - (2n+1)\Delta \max_{j\in J_{r,\theta}} c_{j,r,\theta}\right)\frac{d\theta}{2\pi} \\ &\ge \left(\frac{m}{n+1} - (2n+1)\Delta\right) \int_0^{2\pi} \left(\sum_{j\in J_{r,\theta}} c_{j,r,\theta}\right)\frac{d\theta}{2\pi}. \end{aligned} \tag{107}$$

By (104) there is an $\varepsilon' > 0$ such that

$$\frac{m}{n+1} - (2n+1)\Delta \ge \frac{m}{n+1+\varepsilon/2} \cdot \frac{H_X(m)+\varepsilon'}{H_X(m)}\;;$$

hence (107) becomes

$$\left(n+1+\frac{\varepsilon}{2}\right) \int_0^{2\pi} S_X(m,\mathbf{c}_{r,\theta})\,\frac{d\theta}{2\pi} \ge m(H_X(m)+\varepsilon') \int_0^{2\pi} \left(\sum_{j\in J_{r,\theta}} c_{j,r,\theta}\right)\frac{d\theta}{2\pi}. \tag{108}$$

This corresponds to (97) in the earlier proof.

Now let $\phi_m\colon \mathbb{P}^N_{\mathbb{C}} \to \mathbb{P}^{R_m}_{\mathbb{C}}$ be the m-uple embedding. Then the image of $\phi_m \circ f$ is not contained in any hyperplane not also containing $\phi_m(X)$. As before, let X_m be the linear subspace of $\mathbb{P}^{R_m}_{\mathbb{C}}$ spanned by $\phi_m(X)$.

As before, for each r and θ there are $\mathbf{a}_{1,r,\theta},\dots,\mathbf{a}_{H_X(m),r,\theta} \in \mathbb{N}^{N+1}$, corresponding to a basis of $\mathbb{C}[x_0,\dots,x_N]_m/I_m$, such that

$$S_X(m,\mathbf{c}_{r,\theta}) = \sum_{\ell=1}^{H_X(m)} \mathbf{a}_{\ell,r,\theta}\cdot \mathbf{c}_{r,\theta}. \tag{109}$$

These correspond to linear forms $L_{\ell,r,\theta}$ on $\mathbb{P}^{R_m}_{\mathbb{C}}$, $\ell = 1,\dots,H_X(m)$, which are linearly independent on X_m for each r and θ, and which satisfy

$$\lambda_{L_{\ell,r,\theta}}(\phi_m(f(re^{i\theta}))) \ge \sum_{j=0}^{q} a_{\ell,r,\theta,j}\lambda_{D_j}(f(re^{i\theta})) \tag{110}$$

for suitable choices of Weil functions $\lambda_{L_{\ell,r,\theta}}$.

By the definitions of $J_{r,\theta}$ and of $c_{j,r,\theta}$, by (108), by (109), by (110), and by applying Cartan's Theorem 8.11 to X_m, we then have

$$\begin{aligned}
&\int_0^{2\pi} \max_{J\in\mathscr{J}} \sum_{j\in J} \lambda_{D_j}(f(re^{i\theta}))\,\frac{d\theta}{2\pi} \\
&\quad = \int_0^{2\pi} \sum_{j\in J_{r,\theta}} \lambda_{D_j}(f(re^{i\theta}))\,\frac{d\theta}{2\pi} \\
&\quad = \int_0^{2\pi} \left(\sum_{j\in J_{r,\theta}} c_{j,r,\theta}\right)\frac{d\theta}{2\pi} \\
&\quad \le \frac{n+1+\varepsilon/2}{m(H_X(m)+\varepsilon')} \int_0^{2\pi} S_X(m,\mathbf{c}_{r,\theta})\,\frac{d\theta}{2\pi} \\
&\quad = \frac{n+1+\varepsilon/2}{m(H_X(m)+\varepsilon')} \int_0^{2\pi} \left(\sum_{\ell=1}^{H_X(m)} \mathbf{a}_{\ell,r,\theta}\cdot\mathbf{c}_{r,\theta}\right)\frac{d\theta}{2\pi} \\
&\quad \le \frac{n+1+\varepsilon/2}{m(H_X(m)+\varepsilon')} \int_0^{2\pi} \left(\sum_{\ell=1}^{H_X(m)} \lambda_{L_{\ell,r,\theta}}(\phi_m(f(re^{i\theta})))\right)\frac{d\theta}{2\pi} \\
&\quad \le_{\mathrm{exc}} \frac{n+1+\varepsilon/2}{m} T_{\phi_m\circ f}(r) + C' \\
&\quad \le \left(n+1+\frac{\varepsilon}{2}\right) T_f(r) + C,
\end{aligned}$$

where C' is chosen so that the last inequality holds. □

(This is better than (105) by $\varepsilon/2$, since we have removed the partitioning argument. Thus, the bound (87) can be improved by a factor of 2. This can be done in the number field case too, also by eliminating the partitioning argument there. We decided to keep the partitioning argument in that case, though, since such arguments are common in number theory and it is useful to know how to translate them into Nevanlinna theory.)

23 Truncated Counting Functions and the abc Conjecture

Many results in Nevanlinna theory, when expressed in terms of counting functions instead of proximity functions, hold also in strengthened form using what are called *truncated counting functions.* As usual, one can define truncated counting functions in the number field case as well, and this leads to deep conjectures of high current interest. Perhaps the best-known such conjecture is the abc conjecture of Masser and Oesterlé.

Definition 23.1. Let X be a complete complex variety, let D be an effective Cartier divisor on X, let $f\colon \mathbb{C} \to X$ be a holomorphic curve whose image is not contained in the support of D, and let $n \in \mathbb{Z}_{>0}$. Then the **n-truncated counting function** with respect to D is

$$N_f^{(n)}(D,r) = \sum_{0<|z|<r} \min\{\operatorname{ord}_z f^*D, n\} \log \frac{r}{|z|} + \min\{\operatorname{ord}_0 f^*D, n\} \log r.$$

As with the earlier counting function, the n-truncated counting function is functorial and nonnegative. It is not additive in D, though, due to the truncation. We only have an inequality: If D_1 and D_2 are effective, then

$$N_f^{(n)}(D_1+D_2,r) \le N_f^{(n)}(D_1,r) + N_f^{(n)}(D_2,r).$$

In Nevanlinna theory, the Second Main Theorem for curves has been extended to truncated counting functions:

Theorem 23.2. *Let X be a smooth complex projective curve, let D be a reduced effective divisor on X, let $\mathscr{K}$ be the canonical line sheaf on X, and let $\mathscr{A}$ be an ample line sheaf on X. Then the inequality*

$$N_f^{(1)}(D,r) \ge_{\text{exc}} T_{\mathscr{K}(D),f}(r) - O(\log^+ T_{\mathscr{A},f}(r)) - o(\log r)$$

holds for all nonconstant holomorphic curves $f\colon \mathbb{C} \to X$.

Of course, X needs to be a curve of genus ≤ 1 for this to be meaningful, since otherwise there is no function f. However, if the domain is a finite ramified covering of $\mathbb{C}$, then X can have large genus; see Conjecture 27.5 and the discussion following it.

Also, in Theorem 8.6, the counting functions in (39) can be replaced by n-truncated counting functions:

$$\sum_{j=1}^{q} N_f^{(n)}(H_j,r) \ge_{\text{exc}} (q-n-1)T_f(r) - O(\log^+ T_f(r)) - o(\log r) \tag{111}$$

Theorem 8.11 is not suitable for using truncated counting functions, though, due to its emphasis on the proximity function.

Conjecture 15.2, though, should also be true with counting functions. The question arises, however: truncation to what? Note that (111) is false if the terms $N_f^{(n)}(H_j,r)$ are replaced by $N_f^{(1)}(H_j,r)$, unless one allows the exceptional set to contain hypersurfaces of degree greater than 1 (see Example 23.6). I do believe that Conjecture 15.2 should be true with 1-truncated counting functions, though, even though it would involve substantial complications.

The translation of the above into number theory is straightforward (except that the counterparts to Theorem 23.2 and 111 are only conjectural).

Definition 23.3. Let k be a number field, let $S \supseteq S_\infty$ be a finite set of places of k, let X be a complete variety over k, let D be an effective Cartier divisor on X, and let $n \in \mathbb{Z}_{>0}$. For every place $v \notin S$ (necessarily non-archimedean), let $\mathfrak{p}_v$ denote the corresponding prime ideal in $\mathscr{O}_k$. Then the **n-truncated counting function** with respect to D is

$$N_S^{(n)}(D,x) = \sum_{v \notin S} \min\{\lambda_{D,v}(x), n\log(\mathscr{O}_k : \mathfrak{p}_v)\}$$

for all $x \in X(k)$ not lying in the support of D. If $x \in X(\bar{k})$ lies outside the support of D, then we let $L = \kappa(x)$, let T be the set of places of L lying over S, and define

$$N_S^{(n)}(D,x) = \frac{1}{[L:k]} \sum_{w \notin T} \min\{\lambda_{D,v}(x), n\log(\mathscr{O}_L : \mathfrak{p}_w)\}. \tag{112}$$

Truncation does not respect (48) at ramified places, so (112) is not independent of the choice of L. It is independent of the choice of Weil function, up to $O(1)$. As in the case of Nevanlinna theory, this truncated counting function is functorial and nonnegative, and obeys an inequality

$$N_S^{(n)}(D_1+D_2,x) \le N_S^{(n)}(D_1,x) + N_S^{(n)}(D_2,x)$$

if D_1 and D_2 are effective.

We conjecture that a counterpart to Conjecture 15.6 holds with truncated counting functions:

Conjecture 23.4. Let k be a number field, let $S \supseteq S_\infty$ be a finite set of places of k, let X be a smooth projective variety over k, let D be a normal crossings divisor on X, let $\mathscr{K}$ be the canonical line sheaf on X, and let $\mathscr{A}$ be an ample line sheaf on X. Then:

(a) Let Σ be a generic subset of $X(k) \setminus \operatorname{Supp} D$. Then the inequality

$$N_S^{(1)}(D,x) \ge h_{\mathscr{K}(D),k}(x) - O\left(\sqrt{h_{\mathscr{A},k}(x)}\right) \tag{113}$$

holds for all $x \in \Sigma$.

(b) For any $\varepsilon > 0$ there is a proper Zariski-closed subset Z of X, depending only on X, D, $\mathscr{A}$, and ε, such that for all $C \in \mathbb{R}$ the inequality

$$N_S^{(1)}(D,x) \ge h_{\mathscr{K}(D),k}(x) - \varepsilon h_{\mathscr{A},k}(x) - C \tag{114}$$

holds for almost all $x \in (X \setminus Z)(k)$.

Note that the error term in (113) is weaker than in (65); see Stewart and Tijdeman [81] and van Frankenhuijsen [86].

Unlike the situation in Nevanlinna theory, this conjecture is not known in *any* case over number fields (other than those for which $X(k)$ is not Zariski dense).

The simplest (nontrivial) case of this conjecture is when $X = \mathbb{P}^1_k$ and D is a divisor consisting of three points, say $D = [0] + [1] + [\infty]$. In that case, it is equivalent to the "abc conjecture" of Masser and Oesterlé [62, (9.5)]. This conjecture can be stated (over $\mathbb{Q}$ for simplicity, and with a weaker error term) as follows.

Conjecture 23.5. Fix $\varepsilon > 0$. Then there is a constant C_ε such that there are only finitely many triples $(a,b,c) \in \mathbb{Z}^3$ satisfying $a+b+c=0$, $\gcd(a,b,c)=1$, and

$$\log\max\{|a|,|b|,|c|\} \le (1+\varepsilon)\sum_{p\mid abc}\log p + C_\varepsilon. \tag{115}$$

To see the equivalence with the above-mentioned special case of Conjecture 23.4, let (a,b,c) be a triple of relatively prime rational integers satisfying $a+b+c=0$, and let $x \in \mathbb{P}^2_{\mathbb{Q}}$ be the corresponding point with homogeneous coordinates $[a:b:c]$. Then the left-hand side of (115) is just the height $h_{\mathbb{Q}}(x)$.

Now let D be the divisor consisting of the coordinate hyperplanes H_0, H_1, and H_2 in $\mathbb{P}^2_{\mathbb{Q}}$ (defined respectively by $x_0 = 0$, $x_1 = 0$, and $x_2 = 0$). Since $\gcd(a,b,c) = 1$, we have

$$\lambda_{H_0,p}(x) = -\log\frac{\|a\|_p}{\max\{\|a\|_p,\|b\|_p,\|c\|_p\}} = \operatorname{ord}_p(a)\log p,$$

for all (finite) rational primes p, where $\operatorname{ord}_p(a)$ denotes the largest integer m for which $p^m \mid a$. Thus

$$N^{(1)}_{\{\infty\}}(H_0,x) = \sum_{p\mid a}\log p.$$

Similarly

$$N^{(1)}_{\{\infty\}}(H_1,x) = \sum_{p\mid b}\log p \qquad \text{and} \qquad N^{(1)}_{\{\infty\}}(H_2,x) = \sum_{p\mid c}\log p.$$

Therefore, by relative primeness,

$$N^{(1)}_{\{\infty\}}(D,x) = \sum_{p\mid abc}\log p,$$

so (115) is equivalent to

$$h_{\mathbb{Q}}(x) \le (1+\varepsilon)N^{(1)}_{\{\infty\}}(D,x) + C_\varepsilon.$$

Since $a+b+c=0$, the points $[a:b:c]$ all lie on the line $x_0+x_1+x_2=0$ in $\mathbb{P}^2_{\mathbb{Q}}$. Choosing an isomorphism of this line with $\mathbb{P}^1_{\mathbb{Q}}$ such that the restriction of D corresponds to the divisor $[0]+[1]+[\infty]$ on $\mathbb{P}^1_{\mathbb{Q}}$, it follows by functoriality of $h_{\mathbb{Q}}$ and $N^{(1)}_{\{\infty\}}(D,x)$ that Conjecture 23.5 is equivalent to the special case of Conjecture 23.4 with $k = \mathbb{Q}$, $S = \{\infty\}$, $X = \mathbb{P}^1_{\mathbb{Q}}$, $D = [0]+[1]+[\infty]$, and with a weaker error term.

Some effort has been expended on finding a higher-dimensional counterpart to the abc conjecture (in the spirit of Cartan's and Schmidt's theorems). One major decision, for example, is how to extend the condition on relative primeness. For the equation $a+b+c=0$, the condition $\gcd(a,b,c)=1$ is equivalent to pairwise relative primeness, since $p\mid a$ and $p\mid b$ easily implies $p\mid c$. With more terms, such as $a+b+c+d=0$, though, the two variants are no longer equivalent. For the sake of the present discussion, we use weaker condition of overall relative primeness. This is all that is needed for the largest absolute value to be equivalent to the (multiplicative) height.

So let $a_0,\dots,a_{n+1}$ be integers with $\gcd(a_0,\dots,a_{n+1})=1$ and $a_0+\dots+a_{n+1}=0$. Such an $(n+2)$-tuple gives a point $x:=[a_0:\ldots:a_n]\in\mathbb{P}^n_{\mathbb{Q}}$ with

$$h(x)=\log\max\{|a_0|,\dots,|a_{n+1}|\}+O(1).$$

Let D be the divisor on $\mathbb{P}^n$ consisting of the sum of the coordinate hyperplanes and the hyperplane $x_0+\dots+x_n=0$. Then we have

$$N^{(1)}_{\{\infty\}}(D,x)=\sum_{p\mid a_0\cdots a_{n+1}} p+O(1).$$

Since the canonical line sheaf $\mathscr{K}$ on $\mathbb{P}^n$ is $\mathscr{O}(-n-1)$ and D has degree $n+2$, we have $\mathscr{K}(D)\cong\mathscr{O}(1)$ and therefore (114) would (if true) give

$$\sum_{p\mid a_0\cdots a_{n+1}}\log p\ge(1-\varepsilon)\log\max\{|a_0|,\dots,|a_{n+1}|\}-C \tag{116}$$

for all $\varepsilon>0$ and all C, for almost all rational points $x=[a_0:\ldots:a_n]$ outside of some proper Zariski-closed subset of $\mathbb{P}^n_{\mathbb{Q}}$ depending on ε.

The following example, due to Brownawell and Masser [8, p. 430], then shows that an obvious extension of the abc conjecture to hyperplanes in $\mathbb{P}^n$ is false with 1-truncated counting functions, unless one allows exceptional hypersurfaces of degree >1.

Example 23.6. Let $n\in\mathbb{Z}_{>0}$, and consider the map $\phi\colon\mathbb{P}^1_{\mathbb{Q}}\to\mathbb{P}^n_{\mathbb{Q}}$ given by

$$\phi([x_0:x_1])=\left[x_0^n:\binom{n}{1}x_0^{n-1}x_1:\binom{n}{2}x_0^{n-2}x_1^2:\ldots:x_1^n\right].$$

Let D be the divisor $(x_0x_1(x_0+x_1))=[0]+[-1]+[\infty]$ on $\mathbb{P}^1_{\mathbb{Q}}$ and let D' be the (similar) divisor $(y_0\cdots y_n(y_0+\dots+y_n))$ on $\mathbb{P}^n_{\mathbb{Q}}$. Note that $\operatorname{Supp}\phi^*D'=\operatorname{Supp}D$, so that

$$N^{(1)}_{\{\infty\}}(D',\phi([x_0:x_1]))=N^{(1)}_{\{\infty\}}(D,[x_0:x_1])$$

if $x_0\ne0$ and $x_1\ne0$, and that $h_{\mathbb{Q}}(\phi([x_0:x_1]))=n\,h_{\mathbb{Q}}([x_0:x_1])+O(1)$. It is known that there are infinitely many pairs (a,b) of relatively prime integers for which

$$N^{(1)}_{\{\infty\}}(D,[a:b]) \le h_{\mathbb{Q}}([a:b]) ;$$

therefore we have

$$N^{(1)}_{\{\infty\}}(D',\phi([a:b])) \le \frac{1}{n} h_{\mathbb{Q}}(\phi([a:b])) + O(1),$$

contrary to (116). The points $\phi([a:b])$ are not contained in any hyperplane in $\mathbb{P}^n_{\mathbb{Q}}$. They are, of course, contained in the image of ϕ, hence are not Zariski dense.

The abc conjecture is still unsolved, and is regarded to be quite deep. This is so even though its counterpart in Nevanlinna theory is already known (and has been known for decades). The remainder of these notes discuss extensions of Conjecture 15.6 that all have the property of implying the abc conjecture.

24 On Discriminants

This section discusses some facts about discriminants of number fields. These will be used to formulate a diophantine conjecture for algebraic points in Sect. 25.

Definition 24.1. Let $L \supseteq k$ be number fields, and let D_L denote the absolute discriminant of L. Then the **logarithmic discriminant** of L (relative to k) is

$$d_k(L) = \frac{1}{[L:k]} \log|D_L| - \log|D_k|.$$

Also, if X is a variety over k and $x \in X(\bar{k})$, then let

$$d_k(x) = d_k(\kappa(x)).$$

The discriminant of a number field k is related to the different $\mathscr{D}_{k/\mathbb{Q}}$ of k over $\mathbb{Q}$ by the formulas

$$|D_k| = (\mathbb{Z} : N^k_{\mathbb{Q}} \mathscr{D}_{k/\mathbb{Q}}) = (\mathscr{O}_k : \mathscr{D}_{k/\mathbb{Q}}).$$

By multiplicativity of the different in towers, we therefore have

$$\begin{aligned} d_k(L) &= \frac{1}{[L:k]} \log(\mathscr{O}_L : \mathscr{D}_{L/k}) \\ &= \frac{1}{[L:k]} \sum_{\substack{\mathfrak{q} \in \operatorname{Spec} \mathscr{O}_L \\ \mathfrak{q} \ne (0)}} \operatorname{ord}_{\mathfrak{q}} \mathscr{D}_{L/k} \cdot \log(\mathscr{O}_L : \mathfrak{q}) \end{aligned} \tag{117}$$

for number fields $L \supseteq k$.

This expression can be used to define a “localized” log discriminant term:

Definition 24.2. Let $L \supseteq k$ be number fields, and let $S \supseteq S_\infty$ be a finite set of places of k. Let $\mathscr{O}_{L,S}$ denote the localization of $\mathscr{O}_L$ away from (finite) places of L lying over places in S; note that $\mathscr{O}_{L,S} = \mathscr{O}_L \otimes_{\mathscr{O}_k} \mathscr{O}_{k,S}$ (cf. (58)). Then we define

$$d_S(L) = \frac{1}{[L:k]} \sum_{\substack{\mathfrak{q} \in \operatorname{Spec} \mathscr{O}_{L,S} \\ \mathfrak{q} \neq (0)}} \operatorname{ord}_{\mathfrak{q}} \mathscr{D}_{L/k} \cdot \log(\mathscr{O}_{L,S} : \mathfrak{q}). \tag{118}$$

Also, if X is a variety over k and $x \in X(\bar{k})$, then let

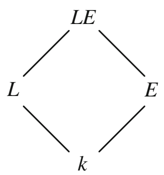
$$d_S(x) = d_S(\kappa(x)).$$

By (117), if $S = S_\infty \subseteq M_k$ then $d_S(L) = d_k(L)$. Other than in this section, these notes will be concerned only with $d_k(\cdot)$. In fact, for any rational prime p the portion of $(\log|D_k|)/[k:\mathbb{Q}]$ coming from primes over p is bounded by $(1+\log_p[k:\mathbb{Q}])\log p$ [73, Chap. III, Remark 1 following Prop. 13]. These notes will be primarily concerned with number fields of bounded degree over $\mathbb{Q}$, so the difference between d_S and d_k is bounded and can be ignored. However, for this section (only) the general situation will be considered, since the results may be useful elsewhere and may not be available elsewhere.

For the remainder of this section, k is a number field and $S \supseteq S_\infty$ is a finite set of places of k.

The following lemma shows that the discriminant is not increased by taking the compositum with a given field.

Lemma 24.3. *Let*



be a diagram of number fields. Then

$$d_S(LE) - d_S(E) \le d_S(L).$$

Proof. We first show that

$$\mathscr{D}_{LE/E} \mid \mathscr{D}_{L/k} \cdot \mathscr{O}_{LE}, \tag{119}$$

which really amounts to showing that $\mathscr{D}_{L/k} \subseteq \mathscr{D}_{LE/E}$. Recall that $\mathscr{D}_{L/k}$ is the ideal in $\mathscr{O}_L$ generated by all elements $f'(\alpha)$, as α varies over the set $\{\alpha \in \mathscr{O}_L : L = k(\alpha)\}$

and f is the (monic) irreducible polynomial for α over k. Such α also lie in $\mathscr{O}_{LE}$, and generate LE over E. Let g be the irreducible polynomial for α over E; we note that $g \mid f$ and therefore $f = gh$ for a monic polynomial $h \in \mathscr{O}_E[t]$. We also have $f'(\alpha) = g'(\alpha)h(\alpha)$. Since $h(\alpha) \in \mathscr{O}_{LE}$, it then follows that $f'(\alpha) \in \mathscr{D}_{LE/E}$, which then implies (119).

It then suffices to show that

$$\operatorname{ord}_{\mathfrak{q}} \mathscr{D}_{L/k} \cdot \log(\mathscr{O}_L : \mathfrak{q}) = \frac{1}{[LE : L]} \sum_{\substack{Q \in \operatorname{Spec} \mathscr{O}_{LE} \\ Q \mid \mathfrak{q}}} \operatorname{ord}_Q(\mathscr{D}_{L/k}\mathscr{O}_{LE}) \cdot \log(\mathscr{O}_{LE} : Q)$$

for all nonzero $\mathfrak{q} \in \operatorname{Spec} \mathscr{O}_L$. But this follows immediately from the classical fact that $[LE : L] = \sum e_{Q/\mathfrak{q}} f_{Q/\mathfrak{q}}$. □

The Chevalley-Weil theorem may be generalized to a situation where ramification is allowed. This involves a proximity function for the ramification divisor, which is defined as follows.

Definition 24.4. Let $\phi \colon X \to Y$ be a generically finite, dominant morphism of nonsingular complete varieties over a field k. Assume that the function field extension $K(X)/K(Y)$ is separable. Then the natural map $\phi^*\Omega_{Y/k} \to \Omega_{X/k}$ induces a natural map $\phi^*\mathscr{K}_Y \to \mathscr{K}_X$ of canonical sheaves, which in turn defines a natural map $\mathscr{O}_X \to \mathscr{K}_X \otimes \phi^*\mathscr{K}_Y^\vee$. This latter map defines a section of $\mathscr{K}_X \otimes \phi^*\mathscr{K}_Y^\vee$, whose divisor is the **ramification divisor** of X over Y. It is an effective divisor R, and we have $\mathscr{K}_X \cong \mathscr{K}_Y(R)$.

(The remainder of this section will be quite technical. Most readers will likely be interested only in the statement of Theorem 24.11, and should now skip to the statement of that theorem and to Theorem 24.13.)

The following definition will also be needed to generalize Chevalley-Weil.

Definition 24.5. Let M be a finitely generated module over a noetherian ring R, and let

$$R^m \xrightarrow{f} R^n \longrightarrow M \longrightarrow 0 \tag{120}$$

be a presentation of M. Then the 0th **Fitting ideal** of M is the ideal $F_0(M)$ in R generated by the determinants of all $n \times n$ submatrices of the $n \times m$ matrix representing f. It is independent of the presentation [17, 20.4]. This globalizes: if $\mathscr{F}$ is a coherent sheaf on a noetherian scheme X, then locally one can form presentations (120) and glue them to give a sheaf of ideals $F_0(\mathscr{F})$.

Lemma 24.6. *Let $\mathscr{F}$ be a coherent sheaf on a noetherian scheme X.*

(a) Forming the Fitting ideal commutes with pull-back: Let $\phi \colon X' \to X$ be a morphism of noetherian schemes. Then

$$F_0(\phi^*\mathscr{F}) = \phi^*F_0(\mathscr{F}) \cdot \mathscr{O}_{X'}. \tag{121}$$

(b) If $\mathscr{F} \twoheadrightarrow \mathscr{F}'$ is a surjection, then

$$F_0(\mathscr{F}') \supseteq F_0(\mathscr{F}).$$

Proof. Let $\phi\colon X' \to X$ be as in (a). Since tensoring is right exact, a presentation of $\mathscr{F}$ pulls back to give a presentation of $\phi^*\mathscr{F}$ on X', and (121) follows directly.

If $\mathscr{F} \twoheadrightarrow \mathscr{F}'$ is a surjection, then one can use the same middle terms in the local presentations of $\mathscr{F}$ and of $\mathscr{F}'$, and the first term in the presentation of $\mathscr{F}$ can be a direct summand of the first term in the presentation of $\mathscr{F}'$. In this case, the resulting generators of $F_0(\mathscr{F})$ are a subset of the generators of $F_0(\mathscr{F}')$. This gives (b). □

The ramification divisor can be described using a Fitting ideal.

Lemma 24.7. *Let $\phi\colon X \to Y$ be as in Definition 24.4. Then $F_0(\Omega_{X/Y})$ is a sheaf of ideals, locally principal and generated by functions f which locally generate the ramification divisor as a Cartier divisor.*

Proof. Indeed, the first exact sequence of differentials

$$\phi^*\Omega_{Y/k} \longrightarrow \Omega_{X/k} \longrightarrow \Omega_{X/Y} \longrightarrow 0$$

gives a locally free presentation of $\Omega_{X/Y}$, and the two initial terms have the same rank. □

The proof of Theorem 24.11 will also need the notion of a *model* of a variety over a number field (corresponding to Definition 16.1 in the function field case).

Definition 24.8. Let X be a variety over a number field k. A **model** for X over $\mathscr{O}_k$ is an integral scheme $\mathscr{X}$, flat over $\mathscr{O}_k$, together with an isomorphism $X \cong \mathscr{X} \times_{\mathscr{O}_k} k$.

If X is a complete variety, then a model $\mathscr{X}$ can be constructed using Nagata's embedding theorem. Moreover, $\mathscr{X}$ can be chosen to be proper over $\operatorname{Spec}\mathscr{O}_k$. On such a model, rational points correspond naturally and bijectively to sections $\mathscr{O}_k \to \mathscr{X}$ by the valuative criterion of properness, and algebraic points $\operatorname{Spec} L \to X$ correspond naturally and bijectively to morphisms $\operatorname{Spec}\mathscr{O}_L \to \mathscr{X}$ over $\mathscr{O}_k$.

As is the case over function fields, a key benefit of working with models is the fact that Weil functions, and therefore the proximity, counting, and height functions, can be defined exactly once one has chosen an extension of the given Cartier divisor D to the model. (Such an extension may not always exist, but the model can be chosen so that it does exist.) At archimedean places, these definitions rely on the additional data specified in Arakelov theory. These definitions, however, will not be described in these notes.

For the purposes of this section, however, we do need to define the proximity function relative to a sheaf of ideals.

Definition 24.9. Let X be a complete variety over a number field k, let $\mathscr{X}$ be a proper model for X over $\mathscr{O}_k$, let $\mathscr{I}$ be a sheaf of ideals on $\mathscr{X}$, and let $S \supseteq S_\infty$ be a finite set of places of k. Let $x \in X(\bar{k})$, and assume that x does not lie in the closed subscheme of $\mathscr{X}$ defined by $\mathscr{I}$. Then the **counting function** $N_S(\mathscr{I},x)$ is defined as follows. Let L be some number field containing $\kappa(x)$, let $i\colon \operatorname{Spec}\mathscr{O}_L \to \mathscr{X}$ be the morphism over $\operatorname{Spec}\mathscr{O}_k$ corresponding to x, and let I be the ideal in $\mathscr{O}_L$ corresponding to the ideal sheaf $i^*\mathscr{I}\cdot\mathscr{O}_{\operatorname{Spec}\mathscr{O}_L}$ on $\operatorname{Spec}\mathscr{O}_L$. Then we define

$$\begin{aligned} N_S(\mathscr{I},x) &= \frac{1}{[L:k]}\log(\mathscr{O}_{L,S} : I\mathscr{O}_{L,S}) \\ &= \frac{1}{[L:k]}\sum_{\substack{\mathfrak{q}\in\operatorname{Spec}\mathscr{O}_{L,S}\\ \mathfrak{q}\neq(0)}} \operatorname{ord}_{\mathfrak{q}} I\cdot\log(\mathscr{O}_{L,S}:\mathfrak{q}). \end{aligned}$$

It can be shown (although we will not do so here) that if $\mathscr{X}$ is a proper model over $\operatorname{Spec}\mathscr{O}_k$ for a complete variety X over k, if $\mathscr{I}$ is an ideal sheaf on $\mathscr{X}$, and if the restriction of $\mathscr{I}$ to X corresponds to a Cartier divisor D, then

$$N_S(\mathscr{I},x) = N_S(D,x) + O(1) \tag{122}$$

for all $x \in X(\bar{k})$ not lying in $\operatorname{Supp}D$.

Counting (and proximity and height) functions relative to sheaves of ideals were first introduced by Yamanoi [102], in the context of Nevanlinna theory.

The different can be described via differentials as well. Indeed, $\mathscr{D}_{L/k}$ is the annihilator of the sheaf of relative differentials:

$$\mathscr{D}_{L/k} = \operatorname{Ann}\Omega_{\mathscr{O}_L/\mathscr{O}_k}. \tag{123}$$

In this case $\Omega_{\mathscr{O}_L/\mathscr{O}_k}$ is a torsion sheaf locally generated by one element; hence (118) can be rewritten as

$$\begin{aligned} d_S(L) &= \frac{1}{[L:k]}\sum_{\substack{\mathfrak{q}\in\operatorname{Spec}\mathscr{O}_{L,S}\\ \mathfrak{q}\neq(0)}} \operatorname{length}_{\mathfrak{q}}\Omega_{\mathscr{O}_L/\mathscr{O}_k}\cdot\log(\mathscr{O}_{L,S}:\mathfrak{q}) \\ &= \frac{1}{[L:k]}\log\#H^0(\mathscr{O}_{L,S},\Omega_{\mathscr{O}_{L,S}/\mathscr{O}_{k,S}}). \end{aligned} \tag{124}$$

The following lemma does most of the work in generalizing the Chevalley-Weil theorem. It has been stated as a separate lemma because it will also be used in the Nevanlinna case.

Lemma 24.10. *Let $A \to B$ be a local homomorphism of discrete valuation rings, with B finite over A, let $\phi\colon X \to Y$ be a generically finite morphism of schemes, and let*

$$\begin{array}{ccc} \operatorname{Spec} B & \xrightarrow{j} & X \\ \downarrow & & \downarrow \phi \\ \operatorname{Spec} A & \longrightarrow & Y \end{array} \tag{125}$$

be a commutative diagram. Assume also that the image of j is not contained in the support of $\Omega_{X/Y}$, that $j^*\mathscr{O}_X$ generates the fraction field of B over the fraction field of A, and that the fraction field of B is separable over the fraction field of A. Then

$$\mathscr{D}_{B/A} \supseteq j^* F_0(\Omega_{X/Y}) \cdot B. \tag{126}$$

Proof. The map j factors through the product $\operatorname{Spec} A \times_Y X$:

$$\begin{array}{ccccc} \operatorname{Spec} B & \xrightarrow{j''} & \operatorname{Spec} A \times_Y X & \xrightarrow{q} & X \\ & & \downarrow & & \downarrow \\ & & \operatorname{Spec} A & \longrightarrow & Y\,; \end{array}$$

here $j = q \circ j''$. We may replace $\operatorname{Spec} A \times_Y X$ in this diagram with an open affine neighborhood $\operatorname{Spec} B''$ of $j''(b)$, where b denotes the closed point of $\operatorname{Spec} B$. By Lemma 24.6a, we have

$$j^* F_0(\Omega_{X/Y}) \cdot B = (j'')^* F_0(\Omega_{B''/A}) \cdot B. \tag{127}$$

The map $\operatorname{Spec} B \to \operatorname{Spec} B''$ corresponds to a ring homomorphism $B'' \to B$; let B' denote its image. Then j'' factors through $j' \colon \operatorname{Spec} B \to \operatorname{Spec} B'$ and a closed immersion $\operatorname{Spec} B' \to \operatorname{Spec} B''$. By the second exact sequence for differentials, the map $\Omega_{B''/A}\big|_{B'} \to \Omega_{B'/A}$ is surjective, and by Lemma 24.6b

$$(j'')^* F_0(\Omega_{B''/A}) \cdot B \subseteq (j')^* F_0(\Omega_{B'/A}) \cdot B. \tag{128}$$

By [43, Def. 10.1 and p. 166], $F_0(\Omega_{B'/A})$ is the Kähler different $\mathfrak{d}_K(B'/A)$. Since B' is finite over A and the fraction field of B' is separable over the fraction field of A, [43, Prop. 10.22] gives

$$F_0(\Omega_{B'/A}) \subseteq \mathscr{D}_{B'/A} \tag{129}$$

(note that $\mathscr{D}_{B'/A}$ is $\mathfrak{d}_D(B'/A)$ in Kunz's notation). Finally, since $B' \subseteq B$, it follows directly from the definition (see for example [43, G.9a]) that

$$\mathscr{D}_{B'/A} \subseteq \mathscr{D}_{B/A}. \tag{130}$$

Combining (127)–(130) then gives (126). □

The generalized Chevalley-Weil theorem can now be stated as follows. This was originally proved in [87, Thm. 5.1.6], but the proof there is only valid if X and Y

have good reduction everywhere, or if the points $x \in X(\bar{k})$ have bounded degree over k. Therefore we will give a general proof here.

Theorem 24.11. *Let $\phi: X \to Y$ be a generically finite, dominant morphism of nonsingular complete varieties over k, with ramification divisor R. Then for all $x \in X(\bar{k})$ not lying on* $\operatorname{Supp} R$, *we have*

$$d_S(x) - d_S(\phi(x)) \le N_S(R,x) + O(1). \tag{131}$$

Proof. Let $\mathscr{X}$ and $\mathscr{Y}$ be models for X and Y over $\mathscr{O}_k$, respectively. By replacing $\mathscr{X}$ with the closure of the graph of the rational map $\phi: \mathscr{X} \to \mathscr{Y}$ if necessary, we may assume that ϕ extends to a morphism $\mathscr{X} \to \mathscr{Y}$ over $\operatorname{Spec} \mathscr{O}_k$. By Lemma 24.7 and (122), it will then suffice to show that

$$d_S(x) - d_S(\phi(x)) \le N_S(F_0(\Omega_{\mathscr{X}/\mathscr{Y}}),x). \tag{132}$$

Fix $x \in X(\bar{k})$ as above, and let $E = \kappa(x)$ and $L = \kappa(\phi(x))$. Then $d_S(x) = d_S(E)$ and $d_S(\phi(x)) = d_S(L)$. Let w be a place of E with $w \nmid S$, and let v be a place of L lying under w. Let $\mathscr{O}_w$ and $\mathscr{O}_v$ denote the localizations of $\mathscr{O}_E$ and $\mathscr{O}_L$ at the primes corresponding to w and v, respectively. Let $j: \operatorname{Spec} \mathscr{O}_w \to \mathscr{X}$ be the restriction of the map $\operatorname{Spec} \mathscr{O}_E \to \mathscr{X}$ over $\mathscr{O}_k$ corresponding to x. By (118), multiplicativity of the different in towers, Definition 24.9, and compatibility of various things with localization, it suffices to show that

$$\mathscr{D}_{\mathscr{O}_w/\mathscr{O}_v} \supseteq j^* F_0(\Omega_{\mathscr{X}/\mathscr{Y}}) \cdot \mathscr{O}_w. \tag{133}$$

The point $\phi(x) \in Y(L)$ determines a map $\operatorname{Spec} \mathscr{O}_v \to \mathscr{Y}$, so there is a diagram

$$\begin{array}{ccc} \operatorname{Spec} \mathscr{O}_w & \xrightarrow{\ j\ } & \mathscr{X} \\ \downarrow & & \downarrow \\ \operatorname{Spec} \mathscr{O}_v & \longrightarrow & \mathscr{Y}. \end{array}$$

This satisfies the conditions of Lemma 24.10, which implies (133). □

Remark 24.12. More generally, the above proof shows that if one replaces (131) with (132), then Theorem 24.11 still holds without the assumptions that X and Y are nonsingular.

The counterpart to this theorem in Nevanlinna theory is the following. (This is much easier in the special case $\dim X = \dim Y = 1$. The general case is more complicated because then the ramification divisor may not be easy to describe.)

Theorem 24.13. *Let B be a connected (nonempty) Riemann surface, let $\phi: X \to Y$ be a generically finite, dominant morphism of smooth complete complex varieties,* *with ramification divisor R, and let $f\colon B\to X$ be a holomorphic function whose image is not contained in $\phi(\operatorname{Supp} R)$. Then there is a connected Riemann surface B', a proper surjective holomorphic map $\pi\colon B'\to B$ of degree bounded by $[K(X):K(Y)]$, and a holomorphic function $g\colon B'\to X$ such that the diagram*

$$\begin{array}{ccc} B' & \xrightarrow{\ g\ } & X \\ \Big\downarrow{\scriptstyle\pi} & & \Big\downarrow{\scriptstyle\phi} \\ B & \xrightarrow{\ f\ } & Y \end{array} \tag{134}$$

commutes. Moreover, if e is the ramification index of B' over B at at any given point of B', then $e-1$ is bounded by the multiplicity of the analytic divisor g^*R at that point. (In other words, the ramification divisor of π is bounded by g^*R, relative to the cone of effective analytic divisors.)

Proof. Let $d=[K(X):K(Y)]$. Let $B_0=\{(b,x)\in B\times X : f(b)=\phi(x)\}$; it is an analytic variety, of degree d over B (i.e., fibers of the projection $B_0\to B$ have at most d points, and some fibers have exactly d points). Let B' be the normalization of B_0 [35, R13]; again B' is of degree d over B. After replacing B' with one of its connected components, we may assume that B' is connected (of degree $\le d$ over B). We then have holomorphic functions $\pi\colon B'\to B$ and $g\colon B'\to X$ as in (134). Also, B' is a Riemann surface [35, Q13].

Now fix $b'\in B'$, and let $b=\pi(b')\in B$. Fix holomorphic local coordinates z' at b' and z at b, vanishing at the respective points. Via the local coordinate z, we identify an open neighborhood of b in B with an open neighborhood of 0 in $\mathbb{C}$, and identify the ring $\mathscr{O}$ of germs of holomorphic functions on B at b with the ring of germs of holomorphic functions on $\mathbb{C}$ at 0. Also let $\mathscr{O}'$ be the ring of germs of holomorphic functions on B' at b', and let e be the ramification index of π at b'. Then the germ of the analytic variety B' at b' is a finite branched covering of the germ of B at b, of covering order e, and we identify $\mathscr{O}$ with a subring of $\mathscr{O}'$ via π. By [35, C5 and C8], there is a canonically defined monic polynomial $P\in\mathscr{O}[t]$ of degree e such that $P(z')=0$, and

$$\mathscr{O}'\cong\mathscr{O}[t]/P(t).$$

Since B' is regular at b', the germ of the variety B' at b' is irreducible, so $\mathscr{O}'$ is an entire ring [35, B6]. Therefore $P(t)$ is irreducible, and by the Weierstrass Preparation Theorem [35, A4] it is a Weierstrass polynomial. This means that all non-leading coefficients vanish at b. A straightforward computation then gives

$$\Omega_{\mathscr{O}'/\mathscr{O}}=\mathscr{O}'/(z')^{e-1}.$$

By [35, A8 and G20], $\mathscr{O}$ and $\mathscr{O}'$ are discrete valuation rings, hence Dedekind, and then (123) gives

$$\mathscr{D}_{\mathscr{O}'/\mathscr{O}}=(z')^{e-1}.$$

Moreover, we have a commutative diagram

$$\begin{array}{ccc} \operatorname{Spec}\mathscr{O}' & \xrightarrow{j} & X \\ \downarrow & & \downarrow \phi \\ \operatorname{Spec}\mathscr{O} & \longrightarrow & Y \end{array}$$

which satisfies the conditions of Lemma 24.10. Therefore,

$$\mathscr{D}_{\mathscr{O}'/\mathscr{O}} \supseteq j^* F_0(\Omega_{X/Y}) \cdot \mathscr{O}'.$$

This, together with Lemma 24.7, implies the theorem. □

Remark 24.14. As was the case in Remark 24.12, Theorem 24.13 remains true when X and Y are allowed to be singular, provided that the conclusion is replaced by the assertion that the ramification divisor of π is bounded by the analytic divisor associated to $g^* F_0(\Omega_{X/Y})$.

25 A Diophantine Conjecture for Algebraic Points

This section describes an extension of Conjecture 15.6 to allow algebraic points instead of rational points. This comes at a cost of adding a discriminant term to the inequality.

This conjecture is subject to some doubt: see Remark 27.6.

Conjecture 25.1. Let k be a number field, let $S \supseteq S_\infty$ be a finite set of places of k, let X be a smooth projective variety over k, let D be a normal crossings divisor on X, let $\mathscr{K}$ be the canonical line sheaf on X, let $\mathscr{A}$ be an ample line sheaf on X, and let r be a positive integer. Then:

(a) Let Σ be a generic subset of $X(\bar{k}) \setminus \operatorname{Supp} D$ such that $[\kappa(x) : k] \le r$ for all $x \in \Sigma$. Then the inequality

$$m_S(D,x) + h_{\mathscr{K},k}(x) \le d_k(x) + O(\log^+ h_{\mathscr{A},k}(x)) \tag{135}$$

holds for all $x \in \Sigma$.

(b) For any $\varepsilon > 0$ there is a proper Zariski-closed subset Z of X, depending only on X, D, $\mathscr{A}$, and ε, such that for all $C \in \mathbb{R}$ the inequality

$$m_S(D,x) + h_{\mathscr{K},k}(x) \le d_k(x) + \varepsilon h_{\mathscr{A},k}(x) + C \tag{136}$$

holds for almost all $x \in (X \setminus Z)(\bar{k})$ with $[\kappa(x) : k] \le r$.

When $r = 1$, this just reduces to Conjecture 15.6, since then $\kappa(x) = k$ for all x. Other than with $r = 1$, no case of this conjecture is known (for number fields). Over function fields, some parts are known.

One may also ask if the conjecture is true without the bound on r. This would require changing the quantization of C in part (b): for example, there are infinitely many roots of unity, which all have height zero. It would also require changing the $d_k(x)$ terms to $d_S(x)$. Other than that, though, it seems to be a reasonable conjecture.

As with any mathematical statement, it is often useful to be aware of how its strength varies with the parameters. For this conjecture, replacing k with a larger number field (and S with the corresponding set), adding places to S, increasing D, increasing r, or (in the case of part (b)) decreasing ε results in a stronger statement.

As was the case with Conjecture 15.6, this Conjecture 25.1 can also be posed for smooth complete varieties X, provided that $h_{\mathscr{A},k}$ is replaced by a big height.

Remark 15.7 does not extend trivially to Conjecture 25.1, though, since the discriminant terms may not add up.

By [87, Prop. 5.4.1],[3] Conjecture 25.1b with $D = 0$ implies the full Conjecture 25.1b. This uses a covering construction.

Next, we show how Conjecture 25.1 relates to generically finite ramified covers.

Proposition 25.2. *Let k and S be as in Conjecture 25.1, let $\pi\colon X' \to X$ be a surjective generically finite morphism of complete nonsingular varieties over k, and let D be a normal crossings divisor on X. Let $D' = (\pi^*D)_{\text{red}}$ (this means the reduced divisor with the same support as π^*D), and assume that it too has normal crossings. Let $\mathscr{K}$ and $\mathscr{K}'$ denote the canonical line sheaves on X and X', respectively. Then, for all $x \in X'(\bar{k})$ not lying on* Supp D' *or on the support of the ramification divisor,*

$$m_S(D,\pi(x)) + h_{\mathscr{K},k}(\pi(x)) - d_S(\pi(x)) \le m_S(D',x) + h_{\mathscr{K}',k}(x) - d_S(x) + O(1). \tag{137}$$

In particular, since the pull-back of any big line sheaf on X to X' remains big, either part of Conjecture 25.1 for D' on X' implies that same part for D on X.

Proof. Let R be the ramification divisor for X' over X; since $\mathscr{K}' \cong \pi^*\mathscr{K} \otimes \mathscr{O}(R)$, we then have

$$h_{\mathscr{K},k}(\pi(x)) - h_{\mathscr{K}',k}(x) = -m_S(R,x) - N_S(R,x) + O(1).$$

Also, Theorem 24.11 gives

$$d_S(x) - d_S(\pi(x)) \le N_S(R,x) + O(1).$$

Finally, by [87, Lemma 5.2.2], $\pi^*D - (\pi^*D)_{\text{red}} \le R$ (relative to the cone of effective divisors). Therefore

$$m_S(D,\pi(x)) - m_S(D',x) \le m_S(R,x) + O(1).$$

[3] The proposition is actually valid in more generality than its statement indicates. However, the proof has an error. The functions $f_1,\dots,f_n$ must be chosen such that each point of Supp D has an open neighborhood U such that $D = (f_i)$ on U for some i.

Adding this equation and the two inequalities then gives (137). □

It was this proposition that motivated the original version of Conjecture 25.1 [87, p. 63].

Finally, we note that Conjecture 25.1 can also be posed with truncated counting functions.

Conjecture 25.3. Let k, S, X, D, $\mathscr{K}$, $\mathscr{A}$, and r be as in Conjecture 25.1. Then:

(a) Let Σ be a generic subset of $X(\bar{k}) \setminus \operatorname{Supp} D$ such that $[\kappa(x) : k] \le r$ for all $x \in \Sigma$. Then the inequality

$$N_S^{(1)}(D,x) + d_k(x) \ge h_{\mathscr{K}(D),k}(x) - O(\log^+ h_{\mathscr{A},k}(x)) \tag{138}$$

holds for all $x \in \Sigma$.

(b) For any $\varepsilon > 0$ there is a proper Zariski-closed subset Z of X, depending only on X, D, $\mathscr{A}$, and ε, such that for all $C \in \mathbb{R}$ the inequality

$$N_S^{(1)}(D,x) + d_k(x) \ge h_{\mathscr{K}(D),k}(x) - \varepsilon h_{\mathscr{A},k}(x) - C \tag{139}$$

holds for almost all $x \in (X \setminus Z)(\bar{k})$ with $[\kappa(x) : k] \le r$.

Remark 25.4. Using a covering construction, it has been shown that Conjecture 25.3b would follow from Conjecture 25.1b [91]. As noted earlier in this section, the latter would then follow from Conjecture 25.1b with $D = 0$, again using a covering construction. In both of these cases, the coverings involved are generically finite, so the implication holds for varieties of any given dimension. Thus, as is noted in the next section, Conjecture 25.3b has been fully proved for curves over function fields of characteristic 0.

Proposition 25.2 does not extend to the situation of truncated counting functions, though.

26 The $1 + \varepsilon$ Conjecture and the abc Conjecture

The special case of Conjecture 25.1 in which $\dim X = 1$ and $D = 0$ is perhaps the most approachable unsolved special case, and has drawn some attention. It is called the "$1 + \varepsilon$ conjecture."

Conjecture 26.1. Let k be a number field, let X be a smooth projective curve over k, let $\mathscr{K}$ denote the canonical line sheaf on X, let r be a positive integer, let $\varepsilon > 0$, and let $C \in \mathbb{R}$. Then

$$h_{\mathscr{K},k}(x) \le (1+\varepsilon) d_k(x) + C$$

for almost all $x \in X(\bar{k})$ with $[\kappa(x) : k] \le r$.

This conjecture was recently proved over function fields of characteristic 0 by McQuillan [57] and (independently) by Yamanoi [103]. See also McQuillan [58] and Gasbarri [27]. Thus, Conjectures 25.1 and 25.3 hold for curves over function fields of characteristic 0, by Remark 25.4; see also [103, Thm. 5]

Conjecture 26.1 is known to imply the abc conjecture [87, pp. 71–72].

Proposition 26.2. *If Conjecture 26.1 holds, then so does the abc conjecture.*

Proof. Let $\varepsilon > 0$, and let a,b,c be relatively prime integers with $a+b+c=0$. For large integers n, there is an associated point

$$P_n = \left[\sqrt[n]{a} : \sqrt[n]{b} : \sqrt[n]{c}\right] \in X_n(\mathbb{Q}),$$

where X_n is the nonsingular curve $x_0^n + x_1^n + x_2^n = 0$ in $\mathbb{P}^2_{\mathbb{Q}}$. This point has height

$$h_{\mathscr{K},\mathbb{Q}}(P_n) = \frac{n-3}{n} \log\max\{|a|,|b|,|c|\} + O(1),$$

since the canonical line sheaf $\mathscr{K}$ on X_n is the restriction of $\mathscr{O}(n-3)$. Here the implicit constant depends only on n. We also have

$$d_{\mathbb{Q}}(P_n) \le \frac{n-1}{n} \sum_{p|abc} \log p + O(1),$$

where the implicit constant depends only on n. Therefore, applying Conjecture 26.1 to points P_n on X_n gives, for all n and all $\varepsilon' > 0$ a constant $C_{n,\varepsilon'} \in \mathbb{R}$ such that

$$\frac{n-3}{n} \log\max\{|a|,|b|,|c|\} \le \left(\frac{n-1}{n} + \varepsilon'\right) \sum_{p|abc} \log p + C_{n,\varepsilon'}.$$

The proof concludes by taking n sufficiently large and ε' sufficiently small so that

$$\left(\frac{n-1}{n} + \varepsilon'\right) \Big/ \frac{n-3}{n} < 1 + \varepsilon,$$

and noting that the constants in the above discussion are independent of the triple (a,b,c). □

27 Nevanlinna Theory of Finite Ramified Coverings

In Nevanlinna theory, changing the domain of the holomorphic function from $\mathbb{C}$ to a *finite ramified covering* is the counterpart to working with algebraic points of bounded degree.

References on finite ramified coverings include Lang and Cherry [49], Chap. III and Yamanoi [103].

Throughout this section, B is a connected Riemann surface, $\pi\colon B \to \mathbb{C}$ is a proper surjective holomorphic map, X is a smooth complete complex variety, and $f\colon B \to X$ is a holomorphic function.

Note that π has a well-defined, finite degree, denoted $\deg \pi$.

We again refer to f as a holomorphic curve.

Definition 27.1. Define

$$\begin{aligned} B[r] &= \{b \in B : |\pi(b)| \le r\}, \\ B(r) &= \{b \in B : |\pi(b)| < r\}, \quad \text{and} \\ B\langle r\rangle &= \{b \in B : |\pi(b)| = r\}. \end{aligned}$$

On $B\langle r\rangle$, let σ be the measure

$$\sigma = \frac{1}{\deg \pi} \pi^* \left(\frac{d\theta}{2\pi}\right).$$

Definition 27.2. Let D be an effective divisor on X whose support does not contain the image of f, and let λ_D be a Weil function for D. Then the **proximity function** of f with respect to D is

$$m_f(D,r) = \int_{B\langle r\rangle} \lambda_D \circ f \cdot \sigma.$$

Definition 27.3.

(a) The **counting function** for an analytic divisor $\Delta = \sum_{b\in B} n_b \cdot b$ on B is

$$N_\Delta(r) = \frac{1}{\deg \pi} \left(\sum_{b\in B(r)\setminus \pi^{-1}(0)} n_b \log \frac{r}{|\pi(b)|} + \sum_{b\in \pi^{-1}(0)} n_b \log r \right).$$

(b) If D is a divisor on X whose support does not contain the image of f, then the **counting function** for D is the function

$$N_f(D,r) = N_{f^*D}(r).$$

(c) The **ramification counting function** for π is the counting function for the ramification divisor of π. It is denoted $N_{\mathrm{Ram}(\pi)}(r)$.

If $B = \mathbb{C}$ and π is the identity mapping, then the proximity function of Definition 27.2 and the counting function of Definition 27.3b extend those of Definitions 12.1

and 12.2, respectively. They also satisfy additivity, functoriality, and boundedness properties, as in Proposition 12.3.

If B' is another connected Riemann surface and $\pi' \colon B' \to B$ is another proper surjective holomorphic map, then

$$m_{f\circ\pi'}(D,r) = m_f(D,r) \qquad \text{and} \qquad N_{f\circ\pi'}(D,r) = N_f(D,r). \tag{140}$$

This holds in particular if $B = \mathbb{C}$ and π is the identity map. It is the counterpart to the fact that the proximity and counting functions in number theory are independent of the choice of number field used in Definition 11.1.

In this situation (but without the assumption $B = \mathbb{C}$), we also have

$$N_{\mathrm{Ram}(\pi\circ\pi')}(r) = N_{\mathrm{Ram}(\pi)}(r) + N_{\mathrm{Ram}(\pi')}(r). \tag{141}$$

(Note that the first term on the right-hand side is a counting function on B, while the others are on B'.) This corresponds to multiplicativity of the different in towers. Note also that, although in general $\mathrm{Ram}(\pi)$ has infinite support, its support in any given set $B(r)$ is finite, in parallel with the fact that any given extension of number fields has only finitely many ramified primes. However, given a sequence of algebraic points of bounded degree, the corresponding sequence of number fields will in general have no bound on the number of ramified primes, corresponding to the fact that $\mathrm{Ram}(\pi)$ may have infinite support.

The height is defined similarly to Definitions 12.4 and 12.7:

Definition 27.4. If D is an effective divisor on X whose support does not contain the image of f, then the **height** of f relative to D is defined up to $O(1)$ by

$$T_{D,f}(r) = m_f(D,r) + N_f(D,r).$$

If $\mathscr{L}$ is a line sheaf on X, then the height $T_{\mathscr{L},f}(r)$ is defined to be $T_{D,f}(r)$ for any divisor D on X for which $\mathscr{O}(D) \cong \mathscr{L}$ and whose support does not contain the image of f.

A First Main Theorem holds for the height as defined here, so the height relative to a line sheaf is well defined [49, III Thm. 2.1]. Theorem 12.8 also holds for heights in this context, as do Propositions 12.10 and 12.11. Corollary 12.9 is not meaningful in this context, since B need not be algebraic.

If $\pi' \colon B' \to B$ is as in (140) and D and $\mathscr{L}$ are as in Definition 27.4, then

$$T_{D,f\circ\pi'}(r) = T_{D,f}(r) + O(1) \qquad \text{and} \qquad T_{\mathscr{L},f\circ\pi'}(r) = T_{\mathscr{L},f}(r) + O(1).$$

Griffiths' conjecture (Conjecture 15.2) can be posed in this context, without any changes other than the domain of the holomorphic curve f, and adding terms $N_{\mathrm{Ram}(\pi)}(r)$ to the right-hand sides of (62) and (63). It will not be repeated here.

The variant with truncated counting functions reads as follows.

Conjecture 27.5. Let X be a smooth complex projective variety, let D be a normal crossings divisor on X, let $\mathscr{K}$ be the canonical line sheaf on X, and let $\mathscr{A}$ be an ample line sheaf on X. Then:

(a) The inequality

$$N_f^{(1)}(D,r)+N_{\operatorname{Ram}(\pi)}(r) \geq_{\text{exc}} T_{\mathscr{K}(D),f}(r)-O(\log^+ T_{\mathscr{A},f}(r))-o(\log r) \tag{142}$$

holds for all holomorphic curves $f\colon B \to X$ with Zariski-dense image.

(b) For any $\varepsilon > 0$ there is a proper Zariski-closed subset Z of X, depending only on X, D, $\mathscr{A}$, and ε, such that the inequality

$$N_f^{(1)}(D,r)+N_{\operatorname{Ram}(\pi)}(r) \geq_{\text{exc}} T_{\mathscr{K}(D),f}(r)-\varepsilon T_{\mathscr{A},f}(r)-C \tag{143}$$

holds for all nonconstant holomorphic curves $f\colon B \to X$ whose image is not contained in Z, and for all $C \in \mathbb{R}$.

This is proved in many of the same situations where Conjecture 15.2 is proved (except possibly for the level of truncation of the counting functions). See for example [49], and also Corollary 29.7 for the case when $\dim X = 1$.

Remark 27.6. This conjecture, and therefore also Conjecture 25.1, is doubted by some. For example, McQuillan [56, Example V.1.5] notes that if X is a quotient of the unit ball in $\mathbb{C}^2$, if $f\colon B \to X$ is a one-dimensional geodesic, and if $\pi\colon B \to \mathbb{C}$ is a proper surjective holomorphic map, then

$$T_{\mathscr{K}_X,f}(r) = N_{\operatorname{Ram}(\pi)}(r)+o(T_{\mathscr{K}_X,f}(r)).$$

However, *loc. cit.* does not address how to show that a suitable map π exists, and in subsequent communications McQuillan has referred only to proper ramified coverings of the unit disk. Therefore, this is not strictly speaking a counterexample, but McQuillan finds it persuasive.

28 The $1+\varepsilon$ Conjecture in the Split Function Field Case

This section describes how the $1+\varepsilon$ conjecture can be easily proved in what is called the "split function field case," following early work of de Franchis [46, p. 223].

Throughout this section, F is a field, B is a smooth projective curve over F, and $k = K(B)$ is the function field of B.

If L is a finite separable extension of k, corresponding to a smooth projective curve B' over F and a finite morphism $B' \to B$ over F, then the logarithmic discriminant term in the function field case is defined as

$$d_k(L) = \frac{\deg \mathscr{K}'}{[L:k]} - \deg \mathscr{K}, \tag{144}$$

where $\mathscr{K}'$ and $\mathscr{K}$ are the canonical line sheaves of B' and B, respectively, and degrees are taken relative to F. As before, we then define $d_k(x) = d_k(\kappa(x))$ for $x \in X(\bar{k})$. The discriminant can also be written

$$d_k(L) = \frac{1}{[L:k]} \dim_F H^0(B', \Omega_{B'/B}) \tag{145}$$

(cf. (124)).

This definition is valid for general function fields.

The remainder of this section will restrict to the **split function field case**. This refers to the situation in which X is of the form $X \cong X_0 \times_F k$ for a smooth projective curve X_0 over F, the model $\mathscr{X}$ is a product $X_0 \times_F B$ (so that the model *splits* into a product), and $\pi\colon \mathscr{X} \to B$ is the projection morphism to the second factor.

Following early work of de Franchis [46, p. 223], it is fairly easy to prove the $1+\varepsilon$ conjecture in the split function field case of characteristic 0.

Theorem 28.1. *Let F be a field of characteristic 0, let X_0 be a smooth projective curve over F, let $X = X_0 \times_F k$, and let $\mathscr{X} = X_0 \times_F B$. Let $\mathscr{K}$ be the canonical line sheaf on X_0, and let $\mathscr{A}$ be an ample line sheaf on X_0. View both of these line sheaves as line sheaves on X or on $\mathscr{X}$ by pulling back via the projection morphisms. Then*

$$h_{\mathscr{K},k}(x) \le d_k(x) + \deg \Omega_{B/F} \tag{146}$$

for all $x \in X(\bar{k})$.

Proof. The proof is particularly easy if F is algebraically closed.

In that case, let $q\colon \mathscr{X} \to X_0$ denote the projection morphism. If $q \circ i$ is a constant morphism, then by (70) the left-hand side of (146) is zero. Since the right-hand side is nonnegative by (145), the inequality is true in this case.

If $q \circ i$ is nonconstant, then it is finite and surjective, and we have

$$h_{\mathscr{K},k}(x) = \frac{(2g(X_0)-2)\deg(q\circ i)}{[K(B'):k]} \le \frac{2g(B')-2}{[K(B'):k]} = d_k(x) + 2g(B) - 2.$$

by (70), the Riemann-Hurwitz formula (twice), and by (144). (Here, as usual, $g(B)$, $g(B')$, and $g(X_0)$ denote the genera of these curves.)

The general case proceeds by reducing to the above special case. First, we may assume that F is algebraically closed in k (i.e., that k/F is a *regular* field extension). Indeed, replacing F with a finite extension divides both sides of (146) by the degree of that extension, due to the fact that all quantities are expressed in terms of degrees of line sheaves, which depend in that way on F.

Let x be an algebraic point on X, let B' be the smooth projective curve over F corresponding to $\kappa(x)$, and let $i\colon B' \to \mathscr{X}$ be the morphism over B corresponding to x. Again, $K(B')$ need not be a regular extension of F; let F' be the algebraic closure of F in $K(B')$. We may replace X_0 with $X_0 \times_F \overline{F}$, B with $B \times_F \overline{F}$, $\mathscr{X}$ with $\mathscr{X} \times_F \overline{F}$, and B' with $B' \times_{F'} \overline{F}$. This again does not affect the validity of (146), since both sides are divided by $[F' : F]$. Indeed, replacing B' with $B' \times_F \overline{F}$ would have left both sides of the inequality unchanged, but B' would now be a disjoint union of $[F' : F]$ smooth projective curves. Choosing one of those curves amounts to taking $B' \times_{F'} \overline{F}$ instead of $B' \times_F \overline{F}$ for some choice of embedding $F' \hookrightarrow \overline{F}$. Also, these changes do not affect the fact that $\mathscr{X} = X_0 \times_F B$.

This reduces to the case in which F is algebraically closed. □

A way to look at this proof is to think of the derivative of the map $i\colon B' \to \mathscr{X}$. It takes values in the absolute tangent bundle $T_{\mathscr{X}/F}$. Since $\mathscr{X}$ is a product, though, the tangent bundle is also a product $p^*T_{X_0/F} \times q^*T_{B/F}$, where p and q are the projection morphisms. This allows us to project onto the second factor $T_{X_0/F}$, which gives a way to bound $T_{\mathscr{K},k}(x)$.

In the general (non-split) function field case, there is first of all no bundle $T_{X_0/F}$ to project to. Instead, we have only the relative tangent bundle $T_{\mathscr{X}/B}$. This is a subbundle of the absolute tangent bundle, not a quotient, and there is no canonical projection from $T_{\mathscr{X}/F}$ to $T_{\mathscr{X}/B}$. McQuillan's proof works mainly because, for points of large height, the tangent vectors giving the derivative of $i\colon B' \to \mathscr{X}$ are "more vertical" than for points of smaller height. Therefore, two arbitrarily chosen ways of projecting the absolute tangent bundle to the relative tangent bundle will differ by a smaller amount, measured relative to the size of the tangent vector. This is sufficient to make the argument carry over.

29 Derivatives in Nevanlinna Theory

Generally speaking, proofs of theorems in Nevanlinna theory rely upon either of two methods for their basic proofs. Historically, the first was Nevanlinna's "Lemma on the Logarithmic Derivative" (Theorem 29.1). Slightly more recently, methods using differential geometry, especially focusing on curvature, have also been used. Although the latter has obvious geometric appeal, the method of the lemma on the logarithmic derivative has also been phrased in geometric terms, and (at present) is the preferred method for comparisons with number theory.

Throughout this section, B is a connected Riemann surface and $\pi\colon B \to \mathbb{C}$ is a proper surjective holomorphic map.

Nevanlinna's original Lemma on the Logarithmic Derivative (LLD) is the following.

Theorem 29.1. (Lemma on the Logarithmic Derivative) *Let f be a meromorphic function on $\mathbb{C}$. Then*

$$\int_0^{2\pi} \log^+ \left| \frac{f'(re^{i\theta})}{f(re^{i\theta})} \right| \frac{d\theta}{2\pi} \le_{\text{exc}} O(\log^+ T_f(r)) + o(\log r). \tag{147}$$

More generally, if f is a meromorphic function on B, then

$$\int_{B\langle r\rangle} \log^+ \left| \frac{df/\pi^* dz}{f} \right| \sigma \le_{\text{exc}} O(\log T_f(r) + \log r). \tag{148}$$

Proof. For the first part, see [60, IX 3.3], or [76, Thm. 3.11] for the error term given here. The second part follows from [3, Thm. 2.2]. □

A geometrical adaptation of this lemma has recently been discovered by Kobayashi, McQuillan, Wong, and others. This first requires a definition.

Definition 29.2. Let X be a smooth complex projective variety, and let D be a normal crossings divisor on X. Then the sheaf $\Omega^1_X(\log D)$ is the subsheaf of the sheaf of meromorphic sections of Ω^1_X generated by the holomorphic sections and the local sections of the form df/f, where f is a local holomorphic function that vanishes only on D [15, II 3.1]. This is locally free of rank $\dim X$. The log tangent sheaf $T_X(-\log D)$ is its dual. There are corresponding vector bundles, of the same names.

Theorem 29.3. (Geometric Lemma on the Logarithmic Derivative) *Let X be a smooth complex projective variety, let D be a normal crossings divisor on X, and let $f\colon B \to X$ be a holomorphic curve whose image is not contained in* $\operatorname{Supp} D$. *Let $\mathscr{A}$ be an ample line sheaf on X. Finally, let $|\cdot|$ be a hermitian metric on the log tangent bundle $T_X(-\log D)$, and let $d_D f\colon B \to T_X(-\log D)$ denote the canonical lifting of f (as a meromorphic function). Then*

$$\int_{B\langle r\rangle} \log^+ \left| d_D f(re^{i\theta}) \right| \sigma \le_{\text{exc}} O(\log T_{\mathscr{A},f}(r) + \log r). \tag{149}$$

Proof (Wong). The general idea of the proof is that one can work locally on finitely many open sets to reduce the question to finitely many applications of the classical LLD. This proof presents a geometric rendition of this idea.

We first note that the assertion is independent of the choice of metric, since by compactness any two metrics are equivalent up to nonzero constant factors.

Next, note that the special case $X = \mathbb{P}^1$, $D = [0] + [\infty]$, is equivalent to the classical Lemma on the Logarithmic Derivative. Indeed, in this case $T_X(-\log D) \cong X \times \mathbb{C}$ is just the trivial vector bundle of rank 1. Choose the metric on $T_X(-\log D)$ to be the one corresponding to the obvious metric on $X \times \mathbb{C}$; then (149) reduces to (147).

The assertion of the theorem is preserved by taking products. Indeed, if it holds for holomorphic curves $f_1\colon B \to X_1$ and $f_2\colon B \to X_2$ relative to normal crossings divisors D_1 and D_2 on smooth complex projective varieties X_1 and X_2, respectively, then it is true for the product $(f_1, f_2)\colon B \to X_1 \times X_2$ relative to the normal crossings divisor $D := p_1^* D_1 + p_2^* D_2$, where $p_j\colon X_1 \times X_2 \to X_j$ are the projection morphisms ($j = 1,2$). This follows by choosing the obvious metric on $T_{X_1 \times X_2}(-\log D)$ and applying (31); details are left to the reader.

Finally, let D' be a normal crossings divisor on a smooth complex projective variety X', let Z be a closed subvariety of X that contains the image of f, and let $\phi\colon Z \dashrightarrow X'$ be a rational map. Assume that there is a nonempty Zariski-open subset U of Z and a constant $C > 0$ such that $|\phi_*(v)| \ge C|v|$ for all $v \in T_X(-\log D)$ lying over U, that the holomorphic curve f meets U, and that Theorem 29.3 holds for the holomorphic curve $\phi \circ f$ in X' relative to D'. We then claim that the theorem also holds for f. Indeed, the left-hand side of (149) does not decrease by more than $\log C$ if f is replaced by $\phi \circ f$, and the right-hand sides in the two cases are comparable by Proposition 12.11 and properties of big line sheaves.

Therefore, we may assume that the divisor D has *strict* normal crossings. Indeed, there is a smooth complex projective variety X' and a birational morphism $\pi\colon X' \to X$, isomorphic over $X \setminus \operatorname{Supp} D$, such that $D' := (\pi^* D)_{\text{red}}$ is a strict normal crossings divisor. This is true because one can resolve the singularities of each component of D. Since π_* induces a holomorphic map $T_{X'}(-\log D') \to T_X(-\log D)$, the inverse rational map π^{-1} satisfies the conditions of the claim.

Thus, to prove the theorem, it suffices to let Z be the Zariski closure of the image of f, and find nonzero elements $f_1, \dots, f_n \in K(Z)$ for which the corresponding rational map $\phi\colon Z \dashrightarrow (\mathbb{P}^1)^n$ satisfies the conditions of the claim (for suitable $U \subseteq Z$ and $C > 0$) relative to the divisor $D' = \sum_{j=1}^n p_j^*([0] + [\infty])$, where $p_j\colon (\mathbb{P}^1)^n \to \mathbb{P}^1$ is the projection morphism to the jth factor, $1 \le j \le n$.

To satisfy the conditions of the claim, it suffices to find a finite set $\mathscr{G}$ of nonzero functions in $K(X)$ such that at each closed point $z \in Z$ there is a subset $\mathscr{G}_z \subseteq \mathscr{G}$ such that for each $g \in \mathscr{G}_z$ the differential dg/g determines a regular section of $\Omega_X^1(\log D)$ in a neighborhood of z in X, and such that as g varies over $\mathscr{G}_z$ the differentials dg/g generate $\Omega_X^1(\log D)$ at z.

To construct $\mathscr{G}$, let z be a closed point of Z. For some open neighborhood V of z in X there are regular functions $g_1, \dots, g_r$ on V whose vanishing determines the components of D passing through z in V. Letting $\mathfrak{m}$ denote the maximal ideal of z in X, the strict normal crossings condition implies that the g_j are linearly independent in the complex vector space $\mathfrak{m}/\mathfrak{m}^2$ (the Zariski cotangent space). After shrinking V if necessary, we may choose regular functions $g_{r+1}, \dots, g_d$ on V, such that the functions $g_1, \dots, g_r, g_{r+1} - 1, \dots, g_d - 1$ all vanish at z, and such that their images in $\mathfrak{m}/\mathfrak{m}^2$ form a basis. Then $dg_1/g_1, \dots, dg_d/g_d$ determine regular sections of $\Omega_X^1(\log D)$ in a neighborhood of z in X, and generate the sheaf on that neighborhood. By a compactness argument, one then obtains a finite collection $\mathscr{G}$ satisfying the condition everywhere on Z. □

Remark 29.4. There is no $N_{\operatorname{Ram}(\pi)}(r)$ term in either of these theorems; it appears subsequently. The same is true of the exceptional set (it appears later still).

Remark 29.5. When $B = \mathbb{C}$ and π is the identity map, the error term in (149) can be sharpened to $O(\log^+ T_{\mathscr{A},f}(r)) + o(\log r)$. This will also be true in subsequent results, but will not be explicitly mentioned.

The Geometric LLD leads to an inequality, due originally to McQuillan [55, Thm. 0.2.5]. This inequality presently shows more promise for possible diophantine analogies, since it omits some of the information on the derivative, and since it may be related to parts of the proof of Schmidt's Subspace Theorem.

Before stating the theorem, we note that for the purposes of these notes, if $\mathscr{E}$ is a quasi-coherent sheaf on a scheme X, then

$$\mathbb{P}(\mathscr{E}) = \mathbf{Proj} \bigoplus_{d \ge 0} S^d \mathscr{E}$$

(as in EGA). In particular, if $\mathscr{E}$ is a vector sheaf, then points on $\mathbb{P}(\mathscr{E})$ correspond bijectively to *hyperplanes* (not lines) in the fiber over the corresponding point on X. This scheme comes with a **tautological line sheaf** $\mathscr{O}(1)$, which gives rise to the name of McQuillan's inequality.

If X and D are as in Theorem 29.3, then $\Omega^1_X(\log D)$ is defined as a locally free sheaf on X as an analytic space. This is a coherent sheaf, hence by GAGA [72], it comes from a coherent sheaf on X as a scheme. This latter sheaf is denoted $\Omega_{X/\mathbb{C}}(\log D)$. In fact it is locally free – see the introduction to Sect. 30.

Theorem 29.6. (McQuillan's "Tautological Inequality") *Let X, D, $f\colon B \to X$, and $\mathscr{A}$ be as in Theorem 29.3. Assume also that f is not constant. Let*

$$f' \colon B \to \mathbb{P}(\Omega_{X/\mathbb{C}}(\log D))$$

be the canonical lifting of f, associated to the nonzero map from $f^*\Omega^1_X(\log D)$ to the cotangent sheaf of B. Then

$$T_{\mathscr{O}(1),f'}(r) \le_{\text{exc}} N^{(1)}_f(D,r) + N_{\text{Ram}(\pi)}(r) + O(\log T_{\mathscr{A},f}(r) + \log r). \tag{150}$$

Proof. Let

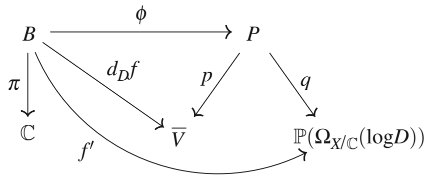
$$V = \mathbb{V}(\Omega_{X/\mathbb{C}}(\log D)) = \mathbf{Spec} \bigoplus_{d \ge 0} S^d \Omega_{X/\mathbb{C}}(\log D).$$

This is the total space of $T_X(-\log D)$. Also let

$$\overline{V} = \mathbb{P}(\Omega_{X/\mathbb{C}}(\log D) \oplus \mathscr{O}_X).$$

We have a natural embedding $V \hookrightarrow \overline{V}$ that realizes $\overline{V}$ as the projective closure on fibers of V.

Let $[\infty]$ denote the (reduced) divisor $\overline{V} \setminus V$. The integrand of (149) can be viewed as a proximity function for $[\infty]$, and the strategy of the proof is to use this to get a bound on $T_{\mathscr{O}(1),f'}(r)$, via the rational map $\overline{V} \dashrightarrow \mathbb{P}(\Omega_{X/\mathbb{C}}(\log D))$. To compare the geometries of these two objects, we use the closure of the graph of this rational map in $\overline{V} \times_X \mathbb{P}(\Omega_{X/\mathbb{C}}(\log D))$, which is the blowing-up of $\overline{V}$ along the image $[0]$ of the zero section. Let $p\colon P \to \overline{V}$ be this blowing-up, let E be its exceptional divisor. Let $q\colon P \to \mathbb{P}(\Omega_{X/\mathbb{C}}(\log D))$ be the projection to the second factor. We have a diagram



There is a unique lifting $\phi : B \to P$ that satisfies $d_D f = p \circ \phi$ and $f' = q \circ \phi$. We also have

$$p^* \mathscr{O}(1) \cong q^* \mathscr{O}(1) \otimes \mathscr{O}(E)$$

(where the first $\mathscr{O}(1)$ is on $\overline{V}$ and the second one is on $\mathbb{P}(\Omega_{X/\mathbb{C}}(\log D))$). This is because any given nonzero rational section s of $\Omega_{X/\mathbb{C}}(\log D)$ on X gives a rational section $(s,1)$ of $\mathscr{O}(1)$ on $\overline{V}$, and also a rational section of $\mathscr{O}(1)$ on $\mathbb{P}(\Omega_{X/\mathbb{C}}(\log D))$. Their pull-backs to P coincide except that the first one also vanishes to first order along E.

We also have $\mathscr{O}([\infty]) \cong \mathscr{O}(1)$ on $\overline{V}$, because the divisor $[\infty]$ is cut out by the section $(0,1)$ of $\Omega_{X/\mathbb{C}}(\log D) \oplus \mathscr{O}_X$.

Thus, we have

$$\begin{aligned}
T_{\mathscr{O}(1),f'}(r) &= T_{q^*\mathscr{O}(1),\phi}(r) + O(1) \\
&= T_{\mathscr{O}(1),d_D f}(r) - T_{\mathscr{O}(E),\phi}(r) + O(1) \\
&= N_{d_D f}([\infty],r) - T_{\mathscr{O}(E),\phi}(r) + m_{d_D f}([\infty],r) + O(1) \\
&\le_{\text{exc}} N_f^{(1)}(D,r) + N_{\text{Ram}(\pi)}(r) + O(\log T_{\mathscr{A},f}(r) + \log r).
\end{aligned}$$

To explain the last step above, $m_{d_D f}([\infty],r)$ is bounded by Theorem 29.3. Since E is effective and does not contain the image of ϕ (since f is not constant), $T_{\mathscr{O}(E),\phi}(r)$ is bounded from below. (It can also be used to subtract a term $N_{\text{Ram}(f)}(r)$ from the right-hand side of (150).)

Now consider $N_{d_D f}([\infty],r)$. Fix a point $b \in B$, let w be a local coordinate on B at b, let z be the coordinate on $\mathbb{C}$, and let $z_1,\dots,z_n$ be local coordinates on X at $f(b)$ such that D is locally given by $z_1 \cdots z_r = 0$ nearby. Then, near $f(b)$, $\overline{V}$ has homogeneous coordinate functions $dz_1/z_1,\dots,dz_r/z_r,dz_{r+1},\dots,dz_n,1$. Relative to these coordinates, the value of $d_D f$ in a punctured neighborhood of b is given by

$$\begin{aligned}
&\left[\frac{d(z_1\circ f)/dz}{z_1\circ f} : \dots : \frac{d(z_r\circ f)/dz}{z_r\circ f} : \frac{d(z_{r+1}\circ f)}{dz} : \dots : \frac{d(z_n\circ f)}{dz} : 1\right] \\
&\qquad = \left[\frac{d(z_1\circ f)/dw}{z_1\circ f} : \dots : \frac{d(z_r\circ f)/dw}{z_r\circ f} : \frac{d(z_{r+1}\circ f)}{dw} : \dots : \frac{d(z_n\circ f)}{dw} : \frac{dz}{dw}\right].
\end{aligned}$$

Then $d_D f$ will meet $[\infty]$ to the extent that there are poles among the first n coordinates or a zero in the last coordinate. Poles among the first n coordinates can only occur in the first r coordinates (using the second representation above), and in that case they will at most be simple poles and will only occur if $f(b) \in \operatorname{Supp} D$. Thus the contribution to $N_{d_D f}([\infty], r)$ from poles in these coordinates is bounded by $N_f^{(1)}(D,r)$. The contribution coming from zeroes in the last coordinate is bounded by $N_{\mathrm{Ram}(\pi)}(r)$. □

As a sample application of this theorem, it implies the Second Main Theorem with truncated counting functions for maps to Riemann surfaces, including the case in which the domain is a finite ramified cover of $\mathbb{C}$. This is the (proved) case $\dim X = 1$ of Conjecture 27.5.

Corollary 29.7. *Let X be a smooth complex projective curve, let D be an effective reduced divisor on X, and let $f\colon B \to X$ be a non-constant holomorphic curve. Then*

$$N_f^{(1)}(D,r) + N_{\mathrm{Ram}(\pi)}(r) \ge_{\mathrm{exc}} T_{\mathscr{K}(D),f}(r) - O(\log T_{\mathscr{A},f}(r) + \log r). \tag{151}$$

Proof. Since X is a curve, the vector sheaf $\Omega_{X/\mathbb{C}}(\log D)$ is isomorphic to the line sheaf $\mathscr{K}(D)$. Therefore the canonical projection $p\colon \mathbb{P}(\Omega_{X/\mathbb{C}}(\log D)) \to X$ is an isomorphism, $\mathscr{O}(1) \cong p^*\mathscr{K}(D)$, and $f' = p^{-1} \circ f$. Thus

$$T_{\mathscr{O}(1),f'}(r) = T_{\mathscr{K}(D),f}(r) + O(1),$$

so (151) is equivalent to (150). □

Remark 29.8. In fact, when $\dim X = 1$, McQuillan's inequality is directly equivalent to Conjecture 27.5, as can be seen from the above proof. This is not true in higher dimension, though (McQuillan's inequality is proved, but Conjecture 27.5 is not).

Cartan's theorem (Theorem 8.6) can also be proved using McQuillan's inequality, but this requires more work than can be included here. See [97]. The modified version (Theorem 8.11) requires a modified form of McQuillan's inequality (involving the same type of change).

It is hoped that other key results in Nevanlinna theory can also be proved using Theorem 29.6.

We end the section with another corollary, which often has applications in Nevanlinna theory. It generalizes the Schwarz lemma, which has played an important role in Nevanlinna theory for a long time; see [80, Thm. 3], where it is proved for jet differentials. The introduction of *op. cit.* also describes some of the history of this result. See also [53, Sect. 4], [93, Cor. 5.2], and [94].

Corollary 29.9. *Let X be a smooth complex projective variety, let D be a normal crossings divisor on X, let $f\colon B \to X$ be a holomorphic map, let $\mathscr{A}$ be an ample line sheaf on X, let $\mathscr{L}$ be a line sheaf on X, let d be a positive integer, and let ω be a* *global section of* $S^d\Omega_{X/\mathbb{C}}(\log D)$*. If* $f^*\omega \ne 0$ *(i.e., it does not vanish everywhere on* B*), then*

$$\frac{1}{d}T_{\mathscr{L},f}(r) \le_{\text{exc}} N_f^{(1)}(D,r) + N_{\text{Ram}(\pi)}(r) + O(\log T_{\mathscr{A},f}(r) + \log r).$$

Proof. Let $f' : B \to \mathbb{P}(\Omega_{X/\mathbb{C}}(\log D))$ be as in Theorem 29.6, and let

$$p\colon \mathbb{P}(\Omega_{X/\mathbb{C}}(\log D)) \to X$$

be the canonical projection. Then ω corresponds to a global section

$$\omega' \in \Gamma\big(\mathbb{P}(\Omega_{X/\mathbb{C}}(\log D)), \mathscr{O}(d) \otimes p^*\mathscr{L}^\vee\big),$$

and $(f')^*\omega' = f^*\omega$. Thus the image of f' is not contained in the base locus of $\mathscr{O}(d) \otimes \mathscr{L}^\vee$, so

$$T_{\mathscr{O}(d),f'}(r) \ge T_{\mathscr{L},f}(r) + O(1)$$

by Theorem 12.8c. The result then follows immediately from (150). □

One can think of this result as a generalization of the fact that if $f\colon C \to X$ is a nonconstant map from a nonsingular projective curve of genus g to a smooth complete variety X, then $\deg f^*\mathscr{K}_X \le 2g-2$, where $\mathscr{K}_X$ is the canonical line sheaf of X. Thus, it is useful in carrying over results from the split function field case to Nevanlinna theory (see Sect. 18). It is used in this manner in the proof of [93, Thm. 5.3].

30 Derivatives in Number Theory

Whether one uses the Lemma on the Logarithmic Derivative or curvature, Nevanlinna theory depends in an essential way on the ability to take the derivative of a holomorphic function. In the number field case, on the other hand, there is currently no known counterpart to the derivative. Even in the function field case, the derivative lives in the absolute tangent bundle, but any counterpart to the derivative as in Nevanlinna theory should live in the *relative* tangent bundle. McQuillan gets around this in his proof of the $1+\varepsilon$ conjecture, by noting that for points of large height the derivative has an approximate projection to the relative tangent bundle that is precise enough to be useful (see the end of Sect. 28). Although this method shows a great deal of promise, it will not be explored further here.

Instead, this section will describe a conjecture in number theory based on McQuillan's tautological inequality. Because of its origin, the name "tautological conjecture" is too good to pass up.

If X is a smooth complete variety over a field k, and if D is a normal crossings divisor on X, then an algebraic definition of $\Omega_{X/k}(\log D)$ is given in [40, 1.7]. Kato

[40, 1.8] also shows this to be locally free in the étale topology of rank $\dim X$. This then descends to a quasi-coherent sheaf on X in the Zariski topology by [33, VIII Thm. 1.1]. It is a vector sheaf (with non-obvious generators) by [34, IV 2.5.2].

Conjecture 30.1. (Tautological Conjecture) Let k be a number field, let $S \supseteq S_\infty$ be a finite set of places of k, let X be a nonsingular complete variety over k with $\dim X > 0$, let D be a normal crossings divisor on X, let r be a positive integer, let $\mathscr{A}$ be an ample line sheaf on X, and let $\varepsilon > 0$. Then, for all $x \in X(\bar{k}) \setminus \operatorname{Supp} D$ with $[\kappa(x) : k] \le r$, there is a closed point $x' \in \mathbb{P}(\Omega_{X/k}(\log D))$ lying over x such that

$$h_{\mathscr{O}(1),k}(x') \le N_S^{(1)}(D,x) + d_k(x) + \varepsilon\, h_{\mathscr{A},k}(x) + O(1). \tag{152}$$

Moreover, given a finite collection of rational maps $g_i \colon X \dashrightarrow W_i$ to varieties W_i, there are finite sets Σ_i of closed points on W_i for each i with the following property. For each x as above, x' may be chosen so that, for each i, if x lies in the domain of g_i and if $g_i(x) \notin \Sigma_i$, then x' lies in the domain of the induced rational map $\mathbb{P}(\Omega_{X/k}(\log D)) \dashrightarrow \mathbb{P}(\Omega_{W_i/k})$. Moreover, the constant implicit in the $O(1)$ term depends only on k, S, X, D, r, $\mathscr{A}$, ε, the rational maps g_i, and the choices of height and counting functions.

This extra condition (involving the rational maps g_i) should perhaps be explained a bit. This condition seems to be necessary in order to ensure that the points x' behave more like derivatives. For example, consider the special case in which $r = 1$, $D = 0$, and X is a product $X_1 \times X_2$. Then the points x must be rational points, and (152) for X is the sum of the same inequality for X_1 and X_2. But then, without the last condition in the conjecture, the conjecture would hold if it held for *either* factor, since one could take x' tangent to the copy of X_1 or X_2 sitting inside of X. This seems a bit unnatural. In addition, the last condition is useful for applications.

McQuillan's work is not the only support for this conjecture. For some time, it has been known that parts of Schmidt's proof of his Subspace Theorem correspond to a proof of Cartan's theorem due to H. and J. Weyl [100], further developed by Ahlfors [1]. Both of these proofs can be divided up into an "old" part (corresponding to an extension to higher dimensions of the proofs of Roth and Nevanlinna of the earlier case on $\mathbb{P}^1$), and a "new" part. In Ahlfors' case, the "new" part consists of working with the associated curves (Frenet formalism); in Schmidt's case, it consists of working with Minkowski's theory of successive minima. In either case, the proof involves geometric constructions on $\bigwedge^p \mathbb{C}^{n+1}$ or $\bigwedge^p k^{n+1}$, respectively. This strongly suggested that the theory of successive minima may be related to the use of derivatives in number theory [87, Chap. 6]. This has been recently refined [97] to more explicitly involve a variant of Conjecture 30.1, and also to use the geometry of flag varieties.

The tie-in between successive minima and the tautological conjecture proceeds as follows. Let X be a nonsingular complete variety over a number field k, let D be a normal crossings divisor on X, let $Y = \operatorname{Spec} \mathscr{O}_k$, and let $\mathscr{X} \to Y$ be a proper model for X. Then we have a relative tangent sheaf $T_{\mathscr{X}/Y}(-\log D)$ on $\mathscr{X}$. This is not necessarily a vector sheaf, since $\mathscr{X}$ need not be smooth over Y and D need not extend as a normal crossings divisor. However, we shall ignore that distinction for the sake of discussion. Via the mechanisms of Arakelov theory, one can assign a Hermitian metric to the sheaf at all archimedean places. If $i\colon \operatorname{Spec} Y \to \mathscr{X}$ is the section of $\mathscr{X} \to Y$ corresponding to a given rational point, then $i^*T_{\mathscr{X}/Y}(-\log D)$ is a vector sheaf on $Y = \operatorname{Spec} \mathscr{O}_k$, which can then be viewed as a lattice in $\mathbb{R}^n$ (if $k = \mathbb{Q}$), or as a lattice in $(k \otimes_{\mathbb{Q}} \mathbb{R})^n$ (generally). Bounds on the metrics at archimedean places correspond to giving a convex symmetric body in $(k \otimes_{\mathbb{Q}} \mathbb{R})^n$, and therefore Minkowski's theory of successive minima can be translated into Arakelov theory as a search for linearly independent nonzero sections of $i^*T_{\mathscr{X}/Y}(-\log D)$, obeying certain upper bounds on its metric at each infinite place. See, for example, [28]. Or, in the function field case, it is known that Minkowski's theory corresponds to a search for nonzero global sections with bounded poles at certain places (an application of Riemann-Roch).

Giving a nonzero section as the first successive minimum is basically equivalent to giving a line subsheaf of largest degree (unless the first two successive minima are close). This corresponds to giving a quotient subbundle $\mathscr{L}$ of $i^*\Omega_{\mathscr{X}/Y}(\log D)$ of smallest degree. This, in turn, corresponds to giving a point $x' \in \mathbb{P}(\Omega_{X/Y}(\log D))$ lying over the rational point in question [36, II 7.12]. The sheaf $\mathscr{L}$ is none other than the pull-back of the tautological bundle $\mathscr{O}(1)$ to Y via the section $i'\colon Y \to \mathbb{P}(\Omega_{\mathscr{X}/Y}(\log d))$ corrsponding to the point x'. Therefore, the degree of $\mathscr{L}$ is the height $h_{\mathscr{O}(1),k}(x')$ (see (70)). This again leads to Conjecture 30.1.

We emphasize that specific bounds on the $O(1)$ term in (152) are not given, so Conjecture 30.1 is meaningful only for an infinite set of points – or, better yet, for a generic or semi-generic set of points (Definitions 15.5 and 15.11).

Remark 30.2. The assertions of Remark 29.8 also apply in the arithmetic situation. When $\dim X = 1$, the extra conditions involving the rational functions g_i in the Tautological Conjecture are automatically satisfied. Therefore, in this case the Tautological Conjecture is equivalent to Conjecture 25.3b (for the same reasons as in Remark 29.8). Thus, by Remark 25.4, the Tautological Conjecture is proved for curves over function fields of characteristic 0. As in Remark 29.8, though, Conjectures 30.1 and 25.3b are not as closely related when $\dim X > 1$.

We also note that points of low height as in Conjecture 30.1 behave like derivatives in the following sense.

Proposition 30.3. (Arithmetic Chain Rule) *Let $f\colon X_1 \to X_2$ be a morphism of complete varieties over a number field k. Then, for all $x \in X_1(\bar{k})$ where f is étale, and for all closed points $x' \in \mathbb{P}(\Omega_{X_1/k})$ lying over x, the rational map $f_*\colon \mathbb{P}(\Omega_{X_1/k}) \dashrightarrow \mathbb{P}(\Omega_{X_2/k})$ takes x' to a point x'_2 (lying over $f(x)$) for which*

$$h_{\mathscr{O}(1),k}(x'_2) \le h_{\mathscr{O}(1),k}(x') + O(1).$$

Moreover, assume that X_1 and X_2 are projective, with ample line sheaves $\mathscr{A}_1$ and $\mathscr{A}_2$, respectively, and that $\varepsilon_1 > 0$ is a positive number for which (152) holds for all x in some set $\Sigma \subseteq X_1(\bar{k})$ (with respect to $\mathscr{A}_1$ and ε_1). Then there is an $\varepsilon_2 > 0$ such that (152) holds for $f(x) \in X_2(\bar{k})$ for all $x \in \Sigma$, with respect to $\mathscr{A}_2$ and ε_2.

Proof. Let $\mathscr{X}_1$ and $\mathscr{X}_2$ be models for X_1 and X_2 over $Y := \operatorname{Spec}\mathscr{O}_k$, respectively, chosen such that f extends as a morphism $f\colon \mathscr{X}_1 \to \mathscr{X}_2$, let $Y' = \operatorname{Spec}\mathscr{O}_{\kappa(x')}$, and let $i\colon Y' \to \mathscr{X}_1$ be the multisection corresponding to x (which factors through $\operatorname{Spec}\mathscr{O}_{\kappa(x)}$). Then x' corresponds to a surjection $i^*\Omega_{\mathscr{X}_1/Y} \twoheadrightarrow \mathscr{L}$ for a line sheaf $\mathscr{L}$ on Y', and $h_{\mathscr{O}(1),k}(x')$ is the Arakelov degree of $\mathscr{L}$ divided by $[\kappa(x') : k]$. We also have a morphism $f^*\Omega_{\mathscr{X}_2/Y} \to \Omega_{\mathscr{X}_1/Y}$, isomorphic at x. This gives a nonzero map $(f\circ i)^*\Omega_{\mathscr{X}_2/Y} \to \mathscr{L}$, so $h_{\mathscr{O}(1),k}(x_2') \le h_{\mathscr{O}(1),k}(x')$ (with heights defined using these models).

The second assertion is immediate from the first assertion, by Proposition 10.13. □

A similar result holds for closed immersions (but without the assumption on étaleness).

The name "Arithmetic Chain Rule" comes from the fact that this result shows that the "derivatives" x' and x_2' are related in the expected way.

31 Another Conjecture Implies abc

Conjecture 23.4, involving truncated counting functions, is of course a vast generalization of the abc conjecture, and Conjecture 25.1, which involved algebraic points, also rather easily implies abc. Actually, though, the (seemingly) weaker Conjecture 15.6 has also been shown to imply the abc conjecture [93]. This implication, however, (necessarily) needs to use varieties of dimension > 1, whereas knowing either of the former two conjectures even for curves would suffice.

This section sketches the proof of the implication mentioned above.

Theorem 31.1. *For any $\varepsilon > 0$ there is a nonsingular projective variety X_ε over $\mathbb{Q}$, a normal crossings divisor D_ε on X_ε, and a real number $\varepsilon' > 0$, such that if Conjecture 15.6b holds for X_ε, D_ε, and ε', then the abc conjecture (Conjecture 23.5) holds for ε.*

Proof (sketch). Fix an integer $n > 3/\varepsilon + 3$. Let X_n be the closed subvariety in $\left(\mathbb{P}^2\right)^n$ in coordinates

$$([x_1 : y_1 : z_1], \dots, [x_n : y_n : z_n])$$

given by the equation

$$\prod_{i=1}^n x_i^i + \prod_{i=1}^n y_i^i + \prod_{i=1}^n z_i^i = 0.$$

There is a rational map $X_n \dashrightarrow \mathbb{P}^2$ given by

$$([x_1 : y_1 : z_1], \dots, [x_n : y_n : z_n]) \mapsto \left[\prod x_i^i : \prod y_i^i : \prod z_i^i\right]. \tag{153}$$

Let Γ_n be the closure of the graph of this rational map in $X_n \times \mathbb{P}^2$, and let $\phi : \Gamma_n \to \mathbb{P}^2$ be the projection to the second factor. The image of ϕ is a line, which we identify with $\mathbb{P}^1$.

Given relatively prime integers a,b,c with $a+b+c=0$, define a point in $X_n(\mathbb{Q})$ as follows. Let

$$x_n = \prod_p p^{[(\operatorname{ord}_p a)/n]} \qquad \text{and} \qquad x_i = \prod_{\operatorname{ord}_p a \equiv i \pmod n} p \quad (i<n).$$

(The brackets in the definition of x_n denote the greatest integer function.) With these definitions, we have $a = \prod x_i^i$, with x_n as large as possible subject to all x_i being integers. Similarly define $y_1,\dots,y_n$ using b and $z_1,\dots,z_n$ using c. This point lifts to a unique point in $\Gamma_n(\mathbb{Q})$, which we denote $P_{a,b,c}$.

Let D be the effective Cartier divisor on Γ_n obtained by pulling back the divisor $x_1\cdots x_n y_1\cdots y_n z_1\cdots z_n = 0$ from X_n, and let E be the divisor on the image $\mathbb{P}^1$ of ϕ, obtained by restricting the coordinate hyperplanes on $\mathbb{P}^2$. The latter divisor is the sum of the points $[1:-1:0]$, $[0:1:-1]$, and $[-1:0:1]$. Let $S=\{\infty\}\subseteq M_{\mathbb{Q}}$. It is possible to show that if p is a rational prime and v is the corresponding place of $\mathbb{Q}$, then

$$\lambda_{E,v}([a:b:c]) = \operatorname{ord}_p(abc)\log p$$

and

$$\lambda_{D,v}(P_{a,b,c}) = \operatorname{ord}_p(x_1\cdots x_n y_1\cdots y_n z_1\cdots z_n)\log p,$$

using Weil functions suitably defined using (40). It then follows that

$$N_S(D,P_{a,b,c}) \le \sum_{p\mid abc}\log p + \frac{1}{n}N_S(E,\phi(P_{a,b,c})).$$

One would like to apply Conjecture 15.6 to the divisor D on Γ_n, but this is not possible since Γ_n is singular. However, there is a nonsingular projective variety Γ_n', a normal crossings divisor D' on Γ_n', and a proper birational morphism $\psi\colon \Gamma_n' \to \Gamma_n$, such that $\operatorname{Supp} D' = \psi^{-1}(\operatorname{Supp} D)$, ψ is an isomorphism over a suitably large set, and $\mathscr{K}_{\Gamma_n'}(D') \ge \psi^*\phi^*\mathscr{O}(1)$ relative to the cone of effective divisors. For details see [93, Lemma 3.9].

Let $P'_{a,b,c}$ be the point on Γ_n' lying over $P_{a,b,c}$. Then one can show that

$$\begin{aligned}
h_{\mathbb{Q}}([a:b:c]) &= h_{\psi^*\phi^*\mathscr{O}(1),\mathbb{Q}}(P'_{a,b,c}) + O(1)\\
&\le h_{\mathscr{K}_{\Gamma_n'}(D'),\mathbb{Q}}(P'_{a,b,c}) + O(1)\\
&\le N_S(D',P'_{a,b,c}) + \varepsilon' h_{\mathscr{A},\mathbb{Q}}(P'_{a,b,c}) + O(1)\\
&\le N_S(D,P_{a,b,c}) + \varepsilon' h_{\mathscr{A},\mathbb{Q}}(P'_{a,b,c}) + O(1)\\
&\le \sum_{p\mid abc}\log p + \frac{3}{n}h_{\mathbb{Q}}([a:b:c]) + \varepsilon'' h_{\mathbb{Q}}([a:b:c]) + O(1)\\
&\le \sum_{p\mid abc}\log p + \frac{\varepsilon}{1+\varepsilon}h_{\mathbb{Q}}([a:b:c]) + O(1),
\end{aligned}$$

and therefore $h_{\mathbb{Q}}([a:b:c]) \le (1+\varepsilon)\sum_{p|abc}\log p + O(1)$. Here we have used the fact that $N_S(D', P'_{a,b,c}) \le N_S(D, P_{a,b,c})$, which follows from the fact that $\psi^*D - D'$ is effective.

This chain of inequalities holds outside of some proper Zariski-closed subset of Γ'_n, but it is possible to show that this set can be chosen so that it only involves finitely many triples (a,b,c). □

The variety X_n admits a faithful action of $\mathbb{G}_{\mathrm{m}}^{2n-2}$, by letting the first $n-1$ coordinates act by $x_i \mapsto tx_i$ and $x_1 \mapsto t^{-i}x_1$ for $i = 2,\dots,n$, and letting the other $n-1$ coordinates act similarly on the y_i. This action respects fibers of the rational map (153), so the action extends to Γ_n, and the construction of Γ'_n can be done so that the group action extends there, too. Since $\dim X_n = 2n-1$, the group acts transitively on open dense subsets of suitably generic fibers of ϕ. It is this group action that allows one to control the proper Zariski-closed subset of Γ'_n arising out of Conjecture 15.6b. The group action also provides some additional structure, and in fact Conjecture 15.2 can be proved in this context [93, Thm. 5.3].

32 An abc Implication in the Other Direction

The preceding sections give a number of ways in which some conjectures imply the abc conjecture. It is also true, however, that the abc conjecture implies parts of the preceding conjectures. While this is mostly a curiousity, since the implied special cases are known to be true whereas the abc conjecture is still a conjecture, this provides some insight into the geometry of the situation.

The implications of this section were first observed by Elkies [18], who showed that "Mordell is as easy as abc," i.e., the abc conjecture implies the Mordell conjecture. This was extended by Bombieri [6], who showed that abc implies Roth's theorem, and then by van Frankenhuijsen [85], who showed that the abc conjecture implies Conjecture 15.6b for curves. In each of these cases, the abc conjecture for a number field k would be needed to imply any given instance of Conjecture 15.6b. Here "the abc conjecture for k" means Conjecture 23.4b over k with $X = \mathbb{P}^1_k$ and $D = [0]+[1]+[\infty]$. It is further true that a "strong abc conjecture," namely Conjecture 25.3b with $X = \mathbb{P}^1_k$ and $D = [0]+[1]+[\infty]$, would imply Conjecture 25.1b for curves; in other words van Frankenhuijsen's implication holds also for algebraic points of bounded degree.

This circle of ideas stems from two observations. The first of these is due to Belyĭ [4]. He showed that a smooth complex projective curve X comes from a curve over $\overline{\mathbb{Q}}$ (i.e., there is a curve X_0 over $\overline{\mathbb{Q}}$ such that $X \cong X_0 \times_{\overline{\mathbb{Q}}} \mathbb{C}$) if and only if there is a finite morphism from X to $\mathbb{P}^1_{\mathbb{C}}$ ramified only over $\{0,1,\infty\}$. For our purposes, this can be adapted as follows.

Theorem 32.1. (Belyĭ) *Let X be a smooth projective curve over a number field k, and let S be a finite set of closed points on X. Then there is a finite morphism $f\colon X \to \mathbb{P}^1_k$ which is ramified only over $\{0,1,\infty\}$, and such that $S \subseteq f^{-1}(\{0,1,\infty\})$.*

Proof. See Belyĭ [4] or Serre [74]. □

The other ingredient is a complement (and actually, a converse) to Proposition 25.2, using truncated counting functions.

Proposition 32.2. *Let k be a number field or function field, let $S \supseteq S_\infty$ be a finite set of places of k, let $\pi\colon X' \to X$ be a surjective generically finite morphism of complete nonsingular varieties over k, and let D be a normal crossings divisor on X. Let $D' = (\pi^* D)_{\mathrm{red}}$, and assume that it too has normal crossings. Assume also that the ramification divisor R of π satisfies*

$$R = \pi^* D - D' \tag{154}$$

(and therefore that π is unramified outside of $\operatorname{Supp} D'$*).*[4] *Let $\mathscr{K}$ and $\mathscr{K}'$ denote the canonical line sheaves on X and X', respectively. Then, for all $x \in X'(\bar{k})$ not lying on* $\operatorname{Supp} D'$,

$$\begin{aligned} &N_S^{(1)}(D',x) + d_S(x) - h_{\mathscr{K}'(D'),k}(x) \\ &\qquad \geq N_S^{(1)}(D,\pi(x)) + d_S(\pi(x)) - h_{\mathscr{K}(D),k}(\pi(x)) + O(1), \end{aligned} \tag{155}$$

with equality if $[\kappa(x) : k]$ is bounded. In particular, either part of Conjecture 25.3 for D' on X' is equivalent to that same part for D on X.

Proof. First, by (154), we have $\mathscr{K}'(D') \cong \pi^*(\mathscr{K}(D))$. Thus

$$h_{\mathscr{K}'(D'),k}(x) = h_{\mathscr{K}(D),k}(\pi(x)) + O(1),$$

so it suffices to show that

$$N_S^{(1)}(D',x) + d_S(x) \geq N_S^{(1)}(D,\pi(x)) + d_S(\pi(x)) + O(1), \tag{156}$$

with equality if $[\kappa(x) : k]$ is bounded.

Now let $Y_S = \operatorname{Spec} \mathscr{O}_{k,S}$ (or, if $S = \emptyset$, which can only happen in the function field case, let Y_S be the smooth projective curve over the field of constants of k for which $K(Y_S) = k$). Let $\mathscr{X}$ be a proper model for X over Y_S for which D extends to an effective Cartier divisor, which will still be denoted D. (This can be obtained by taking a proper model, extending D to it as a Weil divisor, and blowing up the sheaf of ideals of the corresponding reduced closed subscheme.) Let $\mathscr{X}'$ be a proper model for X' over Y_S. We may assume that π extends to a morphism $\pi\colon \mathscr{X}' \to \mathscr{X}$, and that D' extends to an effective Cartier divisor on $\mathscr{X}'$. We assume further that $\operatorname{Supp} D' = \pi^{-1}(\operatorname{Supp} D)$ (on $\mathscr{X}'$). Indeed, one can obtain $\mathscr{X}'$ and D' by blowing up the reduced sheaf of ideals corresponding to $\pi^{-1}(\operatorname{Supp} D)$.

---

[4] This condition is equivalent to (X',D') being *log étale* over (X,D) by [40, (3.12)], using the fact that (X',D') is log smooth over $\operatorname{Spec} k$ by (3.7)(1) of *op. cit.*

For $x \in X'(\bar{k})$, let $L = \kappa(\pi(x))$, let $L' = \kappa(x)$, and let $Y_S^\flat$ and Y_S' be the normalizations of Y_S in L and L', respectively. Then we have a commutative diagram

$$\begin{array}{ccc} Y_S' & \xrightarrow{i'} & \mathscr{X}' \\ \downarrow p & & \downarrow \pi \\ Y_S^\flat & \xrightarrow{i} & \mathscr{X} \end{array} \tag{157}$$

in which the maps i' and i correspond to the algebraic points x and $\pi(x)$, respectively. We use the divisors D and D' to define $N_S^{(1)}(D,\pi(x))$ and $N_S^{(1)}(D',x)$, respectively. Then a place w of L contributes to $N_S^{(1)}(D,\pi(x))$ if and only if the corresponding closed point of $Y_S^\flat$ lies in $i^{-1}(\operatorname{Supp} D)$, and similarly for places w' of L'. By commutativity of (157) and the condition $\operatorname{Supp} D' = \pi^{-1}(\operatorname{Supp} D)$, it follows that

$$N_S^{(1)}(D,\pi(x)) - N_S^{(1)}(D',x) = \frac{1}{[L':k]} \sum_{w'} (e_{w'/w} - 1) \log \#\mathbb{F}_{w'},$$

where the sum is over places w' of L' corresponding to points in $(i')^*(\operatorname{Supp} D')$, w is the place of L lying under w', $e_{w'/w}$ is the ramification index of w' over w, and $\mathbb{F}_{w'}$ is the residue field of w'. (In the function field case, replace $\log \#\mathbb{F}_{w'}$ with $[\mathbb{F}_{w'} : F]$.)

We also have

$$d_S(x) - d_S(\pi(x)) = \frac{1}{[L':k]} \sum_{Q'} \operatorname{ord}_{Q'} \mathscr{D}_{L'/L} \cdot \log \#\mathbb{F}_{w'},$$

where the sum is over nonzero prime ideals Q' of $\mathscr{O}_{L',S}$, and w' is the corresponding place of L'. This sum can be restricted to primes corresponding to points in $(i')^*(\operatorname{Supp} D')$, since the other primes are unramified over L by Lemma 24.10 applied to $\pi \colon \mathscr{X}' \to \mathscr{X}$ and the relevant local rings. Therefore the inequality (156) follows from the elementary fact that

$$\operatorname{ord}_{Q'} \mathscr{D}_{L'/L} \ge e_{Q'/Q} - 1, \tag{158}$$

where $Q = Q' \cap \mathscr{O}_{L,S}$. (This inequality may be strict if Q' is wildly ramified over Q.)

Now if $[L' : k]$ is bounded, then the differences in (158) add up to at most a bounded amount, so (156) holds up to $O(1)$ in that case.

The last assertion of the proposition follows trivially from (155). □

The implications mentioned in the beginning of this section then follow immediately from Theorem 32.1 and Proposition 32.2, upon noting that (154) always holds for finite morphisms of nonsingular curves.

Acknowledgements Partially supported by NSF grant DMS-0500512

References

1. Ahlfors, L.V.: The theory of meromorphic curves. Acta Soc. Sci. Fennicae. Nova Ser. A. **3** (4), 1–31 (1941)
2. Artin, E: Algebraic numbers and algebraic functions. Gordon and Breach Science, New York (1967)
3. Ashline, G.L.: The defect relation of meromorphic maps on parabolic manifolds. Mem. Amer. Math. Soc. **139**(665), 78 (1999). ISSN 0065-9266
4. Belyĭ, G. V.: Galois extensions of a maximal cyclotomic field. Izv. Akad. Nauk SSSR Ser. Mat. **43**(2), 267–276, 479 (1979). ISSN 0373-2436
5. Bloch, A.: Sur les systèmes de fonctions uniformes satisfaisant à l'équation d'une variété algébrique dont l'irrégularité dépasse la dimension. J. de Math. **5**, 19–66 (1926)
6. Bombieri, E.: Roth's theorem and the abc-conjecture. Preprint ETH Zürich (1994)
7. Borel, E.: Sur les zéros des fonctions entières. Acta Math. **20**, 357–396 (1897)
8. Brownawell, W.D., Masser, D.W.: Vanishing sums in function fields. Math. Proc. Cambridge Philos. Soc. **100**(3), 427–434 (1986). ISSN 0305-0041
9. Campana, F.: Orbifolds, special varieties and classification theory. Ann. Inst. Fourier (Grenoble) **54**(3), 499–630 (2004). ISSN 0373-0956
10. Carlson, J., Griffiths, P.: A defect relation for equidimensional holomorphic mappings between algebraic varieties. Ann. Math. **95**(2), 557–584 (1972) ISSN 0003-486X
11. Cartan, H.: Sur les zéros des combinaisons linéaires de p fonctions holomorphes données. Mathematica (Cluj) **7**, 5–29 (1933)
12. Cherry, W., Ye, Z.: Nevanlinna's theory of value distribution. Springer Monographs in Mathematics. Springer, Berlin (2001). ISBN 3-540-66416-5
13. Corvaja, P., Zannier, U.: A subspace theorem approach to integral points on curves. C. R. Math. Acad. Sci. Paris **334**(4), 267–271 (2002). ISSN 1631-073X
14. Corvaja, P., Zannier, U.: On integral points on surfaces. Ann. Math. **160**(2), 705–726 (2004). ISSN 0003-486X
15. Deligne, P.: Équations différentielles à points singuliers réguliers. Lecture Notes in Mathematics, vol. 163. Springer, Berlin (1970)
16. Dufresnoy, J.: Théorie nouvelle des familles complexes normales. Applications à l'étude des fonctions algébroïdes. Ann. Sci. École Norm. Sup. **61**(3), 1–44 (1944) ISSN 0012-9593
17. Eisenbud, D.: Commutative algebra (with a view toward algebraic geometry), vol. 150 In: Graduate texts in mathematics. Springer, New York (1995). ISBN 0-387-94268-8; 0-387-94269-6
18. Elkies, N.D.: *ABC* implies Mordell. Int. Math. Res. Notices, **1991**(7), 99–109 (1991). ISSN 1073-7928
19. Evertse, J.-H.: On sums of S-units and linear recurrences. Compositio Math. **53**(2), 225–244 (1984). ISSN 0010-437X
20. Evertse, J.-H., Ferretti, R.G.: A generalization of the subspace theorem with polynomials of higher degree. In: Diophantine approximation. Dev. Math., vol. 16, pp. 175–198. Springer, NewYork (2008). ArXiv:math.NT/0408381
21. Faltings, G.: Endlichkeitssätze für abelsche Varietäten über Zahlkörpern. Invent. Math. **73**(3), 349–366 (1983). ISSN 0020-9910
22. Faltings, G.: Finiteness theorems for abelian varieties over number fields. In: Arithmetic geometry (Storrs, Conn., 1984), pp. 9–27. Springer, New York (1986). Translated from the German original [Invent. Math. **73**(3), 349–366 (1983); Invent. Math. **75**(2), 381 (1984); MR 85g:11026ab] by Edward Shipz
23. Faltings, G.: Diophantine approximation on abelian varieties. Ann. Math. (2) **133**(3), 549–576 (1991). ISSN 0003-486X
24. Faltings, G.: The general case of S. Lang's conjecture. In: Barsotti Symposium in algebraic geometry (Abano Terme, 1991), Perspect. Math., vol. 15, pp. 175–182. Academic, CA (1994)

25. Ferretti, R.G.: Mumford's degree of contact and Diophantine approximations. Compositio Math. **121**(3), 247–262 (2000). ISSN 0010-437X
26. Fujimoto, H.: Extensions of the big Picard's theorem. Tôhoku Math. J. **24**(2), 415–422 (1972). ISSN 0040-8735
27. Gasbarri, C.: The strong *abc* conjecture over function fields (after McQuillian and Yamanoi). Astérisque, (326): Exp. No. 989, viii, 219–256 (2010). Séminaire Bourbaki. Vol. 2007/2008 (2009)
28. Gillet, H., Soulé, C.: On the number of lattice points in convex symmetric bodies and their duals. Israel J. Math. **74**(2–3), 347–357 (1991). ISSN 0021-2172
29. Goldberg, A.A., Ostrovskii, I.V.: Value distribution of meromorphic functions. In: Translations of mathematical monographs, vol. 236. American Mathematical Society, RI (2008). ISBN 978-0-8218-4265-2; Translated from the 1970 Russian original by Mikhail Ostrovskii, With an appendix by Alexandre Eremenko and James K. Langley
30. Green, M.: Holomorphic maps into complex projective space omitting hyperplanes. Trans. Amer. Math. Soc. **169**, 89–103 (1972). ISSN 0002-9947
31. Green, M.: Some Picard theorems for holomorphic maps to algebraic varieties. Amer. J. Math. **97**, 43–75 (1975). ISSN 0002-9327
32. Green, M., Griffiths, P.: Two applications of algebraic geometry to entire holomorphic mappings. In: The Chern symposium 1979 (Proc. Internat. Sympos., Berkeley, CA, 1979), pp. 41–74. Springer, New York (1980)
33. Grothendieck, A. et al.: Revêtements étales et groupe fondamental. Springer, Berlin (1971). Séminaire de Géométrie Algébrique du Bois Marie 1960–1961 (SGA 1), Dirigé par Alexandre Grothendieck. Augmenté de deux exposés de M. Raynaud, Lecture Notes in Mathematics, vol. 224; arXiv:math.AG/0206203
34. Grothendieck, A., Dieudonné, J.-A.-E.: Éléments de géométrie algébrique. Publ. Math. IHES, 4, 8, 11, 17, 20, 24, 28, and 32, (1960–1967). ISSN 0073-8301.
35. Gunning, R.C.: Introduction to holomorphic functions of several variables. vol. II (Local theory). The Wadsworth and Brooks/Cole Mathematics Series. Wadsworth and Brooks/Cole Advanced Books and Software, CA (1990). ISBN 0-534-13309-6
36. Hartshorne, R.: Algebraic geometry. Springer, New York (1977). ISBN 0-387-90244-9; Graduate Texts in Mathematics, No. 52
37. Hayman, W.K.: Meromorphic functions. Oxford Mathematical Monographs. Clarendon Press, Oxford (1964)
38. Hodge, W.V.D., Pedoe, D.: Methods of algebraic geometry, vol. II. Book III: General theory of algebraic varieties in projective space. Book IV: Quadrics and Grassmann varieties. Cambridge University Press, Cambridge (1952)
39. Iitaka, S.: Algebraic geometry, An introduction to birational geometry of algebraic varieties, Graduate texts in mathematics, vol. 76. Springer, New York (1982). ISBN 0-387-90546-4; North-Holland Mathematical Library, 24
40. Kato, K.: Logarithmic structures of Fontaine-Illusie. In: Algebraic analysis, geometry, and number theory (Baltimore, MD, 1988), pp. 191–224. Johns Hopkins University Press, MD (1989)
41. Kawamata, Y.: On Bloch's conjecture. Invent. Math. **57**(1), 97–100 (1980). ISSN 0020-9910
42. Koblitz, N.: *p*-adic numbers, *p*-adic analysis, and zeta-functions, Graduate texts in mathematics, vol. 58, 2nd edn. Springer, New York (1984). ISBN 0-387-96017-1
43. Kunz, E.: Kähler differentials. Advanced Lectures in Mathematics. Friedr. Vieweg, Braunschweig (1986). ISBN 3-528-08973-3
44. Lang, S.: Integral points on curves. Inst. Hautes Études Sci. Publ. Math. **6**, 27–43 (1960). ISSN 0073-8301
45. Lang, S.: Algebraic number theory. Addison-Wesley, MA (1970)
46. Lang, S.: Fundamentals of diophantine geometry. Springer, New York (1983). ISBN 0-387-90837-4
47. Lang, S.: Hyperbolic and diophantine analysis. Bull. Amer. Math. Soc. (N.S.) **14**(2), 159–205 (1986). ISSN 0273-0979

48. Lang, S.: Number theory III: Diophantine geometry, Encyclopaedia of mathematical sciences, vol. 60. Springer, Berlin (1991). ISBN 3-540-53004-5
49. Lang, S., Cherry, W.: Topics in Nevanlinna theory, Lecture Notes in mathematics, vol. 1433. Springer, Berlin (1990). ISBN 3-540-52785-0; With an appendix by Zhuan Ye
50. Levin, A.: The dimensions of integral points and holomorphic curves on the complements of hyperplanes. Acta Arith. **134**(3), 259–270 (2008). ISSN 0065-1036
51. Levin, A.: Generalizations of Siegel's and Picard's theorems. Ann. Math. (2) **170**(2),609–655 (2009). ISSN 0003-486X
52. Levin, A., McKinnon, D., Winkelmann, J.: On the error terms and exceptional sets in conjectural second main theorems. Q. J. Math. **59**(4), 487–498 (2008). ISSN 0033-5606, doi:10.1093/qmath/ham052, http://dx.doi.org/10.1093/qmath/ham052
53. Lu, S.S.-Y.: On meromorphic maps into varieties of log-general type. In: Several complex variables and complex geometry, Part 2 (Santa Cruz, CA, 1989), Proc. Sympos. Pure Math., vol. 52, pp. 305–333. AMS, RI (1991)
54. Mahler, K.: On a theorem of Liouville in fields of positive characteristic. Canadian J. Math. **1**, 397–400 (1949). ISSN 0008-414X
55. McQuillan, M.: Diophantine approximations and foliations. Inst. Hautes Études Sci. Publ. Math. **87**, 121–174 (1998). ISSN 0073-8301
56. McQuillan, M.: Non-commutative Mori theory. IHES preprint IHES/M/00/15 (2000)
57. McQuillan, M.: Canonical models of foliations. Pure Appl. Math. Q. **4**(3, part 2), 877–1012 (2008). ISSN 1558-8599
58. McQuillan, M.: Old and new techniques in function field arithmetic. (2009, submitted)
59. Neukirch, J.: Algebraic number theory, Grundlehren der Mathematischen Wissenschaften, vol. 322 [Fundamental Principles of Mathematical Sciences]. Springer, Berlin (1999). ISBN 3-540-65399-6; Translated from the 1992 German original and with a note by Norbert Schappacher, With a foreword by G. Harder
60. Nevanlinna, R.: Analytic functions. Translated from the second German edition by Phillip Emig. Die Grundlehren der mathematischen Wissenschaften, Band 162. Springer, New York (1970)
61. Noguchi, J.: Lemma on logarithmic derivatives and holomorphic curves in algebraic varieties. Nagoya Math. J. **83**, 213–233 (1981) ISSN 0027-7630
62. Noguchi, J.: On Nevanlinna's second main theorem. In: Geometric complex analysis (Hayama, 1995), pp. 489–503. World Scientific, NJ (1996)
63. Noguchi, J.: On holomorphic curves in semi-abelian varieties. Math. Z. **228**(4), 713–721 (1998). ISSN 0025-5874
64. Osgood, C.F.: Effective bounds on the Diophantine approximation of algebraic functions over fields of arbitrary characteristic and applications to differential equations. Nederl. Akad. Wetensch. Proc. Ser. A **78** Indag. Math. **37**, 105–119 (1975)
65. Osgood, C.F.: A number theoretic-differential equations approach to generalizing Nevanlinna theory. Indian J. Math. **23**(1–3), 1–15 (1981). ISSN 0019-5324
66. Osgood, C.F.: Sometimes effective Thue-Siegel-Roth-Schmidt-Nevanlinna bounds, or better. J. Number Theor. **21**(3), 347–389 (1985). ISSN 0022-314X
67. Roth, K.F.: Rational approximations to algebraic numbers. Mathematika **2**,1–20; corrigendum, 168 (1955). ISSN 0025-5793
68. Ru, M.: On a general form of the second main theorem. Trans. Amer. Math. Soc. **349**(12), 5093–5105 (1997). ISSN 0002-9947
69. Ru, M.: Nevanlinna theory and its relation to Diophantine approximation. World Scientific, NJ (2001). ISBN 981-02-4402-9
70. Ru, M.: Holomorphic curves into algebraic varieties. Ann. Math. (2) **169** (1), 255–267 (2009). ISSN 0003-486X, doi:10.4007/annals.2009.169.255, http://dx.doi.org/10.4007/annals.2009.169.255
71. Schmidt, W.M.: Diophantine approximations and Diophantine equations, Lecture Notes in Mathematics, vol. 1467. Springer, Berlin (1991). ISBN 3-540-54058-X

72. Serre, J.-P.: Géométrie algébrique et géométrie analytique. Ann. Inst. Fourier, Grenoble **6**, 1–42 (1955/1956). ISSN 0373-0956
73. Serre, J.-P.: Corps locaux. Hermann, Paris (1968). Deuxième édition, Publications de l'Université de Nancago, No. VIII
74. Serre, J.-P.: Lectures on the Mordell-Weil theorem. Aspects of Mathematics, E15. Friedr. Vieweg, Braunschweig (1989). ISBN 3-528-08968-7; Translated from the French and edited by Martin Brown from notes by Michel Waldschmidt
75. Shabat, B.V.: Distribution of values of holomorphic mappings, Translations of mathematical monographs, vol. 61. American Mathematical Society, RI (1985). ISBN 0-8218-4514-4; Translated from the Russian by J. R. King, Translation edited by Lev J. Leifman
76. Shiffman, B.: Introduction to the Carlson-Griffiths equidistribution theory. In: Value distribution theory (Joensuu, 1981), Lecture Notes in Math., vol. 981, pp. 44–89. Springer, Berlin (1983)
77. Silverman, J.H.: The theory of height functions. In: Arithmetic geometry (Storrs, Conn., 1984), pp. 151–166. Springer, New York (1986)
78. Siu, Y.-T.: Hyperbolicity problems in function theory. In: Five decades as a mathematician and educator, pp. 409–513. World Scientific, NJ (1995)
79. Siu, Y.-T., Yeung, S.-K.: A generalized Bloch's theorem and the hyperbolicity of the complement of an ample divisor in an abelian variety. Math. Ann. **306**(4), 743–758 (1996). ISSN 0025-5831
80. Siu, Y.-T., Yeung, S.-K.: Defects for ample divisors of abelian varieties, Schwarz lemma, and hyperbolic hypersurfaces of low degrees. Amer. J. Math. **119**(5), 1139–1172 (1997). ISSN 0002-9327
81. Stewart, C.L., Tijdeman, R.: On the Oesterlé-Masser conjecture. Monatsh. Math. **102**(3), 251–257 (1986). ISSN 0026-9255
82. Stoll, W.: Value distribution of holomorphic maps into compact complex manifolds. Lecture Notes in Mathematics, vol. 135. Springer, Berlin (1970)
83. Szpiro, L., Ullmo, E., Zhang, S.: Équirépartition des petits points. Invent. Math. **127**(2), 337–347 (1997). ISSN 0020-9910
84. van der Poorten, A.J., Schlickewei, H.P.: The growth conditions for recurrence sequences. Macquarie Math. Reports, 82-0041 (1982)
85. van Frankenhuijsen, M.: The *ABC* conjecture implies Vojta's height inequality for curves. J. Number Theor. **95**(2), 289–302 (2002). ISSN 0022-314X
86. van Frankenhuijsen, M.: *ABC* implies the radicalized Vojta height inequality for curves. J. Number Theor. **127**(2), 292–300 (2007). ISSN 0022-314X
87. Vojta, P.: Diophantine approximations and value distribution theory, Lecture Notes in Mathematics, vol. 1239. Springer, Berlin (1987). ISBN 3-540-17551-2
88. Vojta, P.: A refinement of Schmidt's subspace theorem. Amer. J. Math. **111**(3), 489–518 (1989). ISSN 0002-9327
89. Vojta, P.: Integral points on subvarieties of semiabelian varieties. I. Invent. Math. **126**(1), 133–181 (1996). ISSN 0020-9910
90. Vojta, P.: On Cartan's theorem and Cartan's conjecture. Amer. J. Math. **119**(1),1–17 (1997). ISSN 0002-9327
91. Vojta, P.: A more general *abc* conjecture. Int. Math. Res. Notices, **1998**(21), 1103–1116 (1998). ISSN 1073-7928
92. Vojta, P.: Integral points on subvarieties of semiabelian varieties. II. Amer. J. Math. **121**(2), 283–313 (1999). ISSN 0002-9327
93. Vojta, P.: On the *ABC* conjecture and Diophantine approximation by rational points. Amer. J. Math. **122**(4), 843–872 (2000). ISSN 0002-9327
94. Vojta, P.: Correction to: On the *ABC* conjecture and Diophantine approximation by rational points [Amer. J. Math. **122**(4), 843–872 (2000); MR1771576 (2001i:11094)]. Amer. J. Math. **123**(2), 383–384 (2001). ISSN 0002-9327
95. Vojta, P.: Jets via Hasse-Schmidt derivations. In: Diophantine geometry, CRM Series, vol. 4, pp. 335–361. Ed. Norm., Pisa (2007). arXiv:math.AG/0407113

96. Vojta, P.: Nagata's embedding theorem. arXiv:0706.1907 (2007, to appear)
97. Vojta, P.: On McQuillan's tautological inequality and the Weyl-Ahlfors theory of associated curves. arXiv:0706.3044 (2008, to appear)
98. Weil, A.: L'arithmétique sur les courbes algébriques. Acta Math. **52**, 281–315 (1928)
99. Weil, A.: Arithmetic on algebraic varieties. Ann. Math. **53**(2), 412–444 (1951). ISSN 0003-486X
100. Weyl, H., Weyl, J.: Meromorphic curves. Ann. Math. (2) **39**(3), 516–538 (1938). ISSN 0003-486X
101. Wong, P.-M.: On the second main theorem of Nevanlinna theory. Amer. J. Math. **111**(4), 549–583 (1989). ISSN 0002-9327
102. Yamanoi, K.: Algebro-geometric version of Nevanlinna's lemma on logarithmic derivative and applications. Nagoya Math. J. **173**, 23–63 (2004). ISSN 0027-7630
103. Yamanoi, K.: The second main theorem for small functions and related problems. Acta Math. **192**(2), 225–294 (2004). ISSN 0001-5962