

# Diophantine Approximation and Nevanlinna Theory

Paul Vojta

**Abstract** As was originally observed by C. F. Osgood and further developed by the author, there is a formal analogy between Nevanlinna theory in complex analysis and certain results in diophantine approximation. These notes describe this analogy, after briefly introducing the theory of heights and Weil function in number theory and the main concepts of Nevanlinna theory. Parallel conjectures are then presented, in Nevanlinna theory (“Griffiths’ conjecture”) and the author’s conjecture on rational points. Following this, recent work is described, highlighting work of Corvaja, Zannier, Evertse, and Ferretti on combining Schmidt’s Subspace Theorem with geometrical constructions to obtain partial results on this conjecture. Counterparts of these results in Nevanlinna theory are also given (due to Ru). These notes also describe parallel extensions of the conjectures in Nevanlinna theory and diophantine approximation, to involve finite ramified coverings and algebraic points, respectively. Variants of these conjectures involving truncated counting functions are also introduced, and the relations of these various conjectures with the abc conjecture of Masser and Oesterlé are also described.

## 0 Introduction

Beginning with the work of C. F. Osgood [1981], 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 [Vojta, 1987] 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 projec-

---

Paul Vojta

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

Partially supported by NSF grant DMS-0500512

tive variety should have only finitely many rational points over any given number field (i.e., is *Mordellian*) 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 Section 14 for details) led to proofs of the corresponding results in number theory (again, see Section 14). 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 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, Sections 1–3 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 [1970] and Hartshorne [1977], respectively. Some proofs are omitted, however; for those the interested reader may refer to Lang [1983].

Sections 4–6 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 [1985], Nevanlinna [1970], and Goldberg and Ostrovskii [2008] for pure Nevanlinna theory, and Vojta [1987] and Ru [2001] for the analogy.

Sections 7–15 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 14 in particular introduces the main conjectures being discussed here: Conjecture 14.2 in Nevanlinna theory (“Griffiths’ conjecture”) and its counterpart in number theory, the author’s Conjecture 14.6 on rational points on varieties.

Sections 16 and 17 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 Sections 18–22.

Sections 23–27 present generalizations of Conjectures 14.2 and 14.6. Conjecture 14.2, 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 14.6 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 22.5 can be regarded as a generalization of the abc conjecture as well as of Conjecture 14.6. One can also generalize Conjecture 14.6 to treat algebraic points; this corresponds to finite ramified coverings in Nevanlinna theory. This is Conjecture 24.1, which can also be posed using truncated counting functions (Conjecture 24.3).

Sections 28 and 29 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 R. Kobayashi, M. McQuillan, and P.-M. 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 29.1), 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 24.1, on algebraic points, implies the abc conjecture, even if known only in dimension 1, and Conjecture 14.6, on rational points, also implies abc if known in high dimensions. This latter implication is the subject of Section 30. Finally, implications in the other direction are explored in Section 31.

## 1 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 [Lang, 1983] and [Vojta, 1987]. 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  $\mathcal{O}_k$  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  $\mathcal{O}_k$ , the set of real embeddings  $\sigma: k \hookrightarrow \mathbb{R}$ , and the set of unordered complex conjugate pairs  $(\sigma, \bar{\sigma})$  of complex embeddings  $\sigma: k \hookrightarrow \mathbb{C}$  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  $S_\infty$ . 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

$$(1.1) \quad \|x\|_v = \begin{cases} (\mathcal{O}_k : \mathfrak{p})^{\text{ord}_{\mathfrak{p}}(x)} & \text{if } v \text{ corresponds to } \mathfrak{p} \subseteq \mathcal{O}_k; \\ |\sigma(x)| & \text{if } v \text{ corresponds to } \sigma: k \hookrightarrow \mathbb{R}; \text{ and} \\ |\sigma(x)|^2 & \text{if } v \text{ is a complex place, corresponding to } \sigma: k \hookrightarrow \mathbb{C} \end{cases}$$

if  $x \neq 0$ . Here  $\text{ord}_{\mathfrak{p}}(x)$  means the exponent of  $\mathfrak{p}$  in the factorization of the fractional ideal  $(x)$ . If we use the convention that  $\text{ord}_{\mathfrak{p}}(0) = \infty$ , then (1.1) 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

$$(1.2) \quad N_v = \begin{cases} 0 & \text{if } v \text{ is non-archimedean;} \\ 1 & \text{if } v \text{ is real; and} \\ 2 & \text{if } v \text{ is complex.} \end{cases}$$

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

**(1.3.1)**  $\|x\|_v \geq 0$ , with equality if and only if  $x = 0$ ;

**(1.3.2)**  $\|xy\|_v = \|x\|_v \|y\|_v$  for all  $x, y \in k$ ; and

**(1.3.3)**  $\|x_1 + \cdots + x_n\|_v \leq n^{N_v} \max\{\|x_1\|_v, \dots, \|x_n\|_v\}$  for all  $x_1, \dots, x_n \in k$ ,  $n \in \mathbb{N}$ .

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

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  $\mathfrak{q} \subseteq \mathcal{O}_L$ , then  $\mathfrak{p} := \mathfrak{q} \cap \mathcal{O}_k$  is a nonzero prime of  $\mathcal{O}_k$ , 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  $\tau: L \hookrightarrow \mathbb{C}$ , and its restriction  $\tau|_k: k \hookrightarrow \mathbb{C}$  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

$$(1.4) \quad \|x\|_w = \|x\|_v^{[L_w:k_v]} \quad \text{for all } x \in k,$$

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

$$(1.5) \quad \prod_{\substack{w \in M_L \\ w \mid v}} \|y\|_w = \|N_k^L y\|_v \quad \text{for all } v \in M_k \text{ and all } y \in L.$$

This is proved by using the isomorphism  $L \otimes_k k_v \cong \prod_{w \mid v} L_w$ ; see for example [Neukirch, 1999, Ch. 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**

$$(1.6) \quad \prod_{v \in M_k} \|x\|_v = 1 \quad \text{for all } x \in k^*.$$

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 (1.5) to pass to an arbitrary number field.

## 2 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

$$(2.1) \quad H_k(x) = \prod_{v \in M_k} \max\{\|x\|_v, 1\}.$$

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 (2.1) over all finite places  $v$  is just  $|b|$ . At the infinite place, we have  $\|x\|_\infty = |a/b|$ , so this gives

$$(2.2) \quad H_{\mathbb{Q}}(x) = \max\{|a|, |b|\}.$$

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  $[x_0 : \cdots : x_n]$  be homogeneous coordinates for  $P$  (with the  $x_i$  always assumed to lie in  $k$ ). Then we define

$$(2.3) \quad H_k(P) = \prod_{v \in M_k} \max\{\|x_0\|_v, \dots, \|x_n\|_v\}.$$

By the Product Formula (1.6), 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  $i: \mathbb{A}^1 \hookrightarrow \mathbb{P}^1$ , 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**:

$$(2.4) \quad h_k(x) = \log H_k(x) = \sum_{v \in M_k} \log^+ \|x\|_v$$

and

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

Here

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

The equation (1.5) tells us how heights change when the number field  $k$  is extended to a larger number field  $L$ :

$$(2.5) \quad h_L(x) = [L : k]h_k(x)$$

and

$$(2.6) \quad h_L([x_0 : \cdots : x_n]) = [L : k]h_k([x_0 : \cdots : x_n])$$

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  $[x_0 : \cdots : x_n] \in \mathbb{P}^n(\overline{\mathbb{Q}})$ , we define

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

for any number field  $L \supseteq k(x_0, \dots, x_n)$ . These expressions are independent of the choice of  $L$  by (2.5) and (2.6), respectively.

Following EGA, if  $x$  is a point on  $\mathbb{P}_k^n$ , 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, (2.5) and (2.6) remain valid without the conditions  $x \in k$  and  $[x_0 : \cdots : x_n] \in \mathbb{P}^n(k)$ , 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 (2.1) that  $h_k(x) \geq 0$  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  $\|x\|_v \leq 1$  for all  $v$ ; if  $x \neq 0$  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 2.7.** (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  $[\mathbb{Q}(x) : \mathbb{Q}] \leq r$  and  $h(x) \leq C$ . Moreover, given any  $n \in \mathbb{N}$  there are only finitely many  $x \in \mathbb{P}^n(\overline{\mathbb{Q}})$  such that  $[\kappa(x) : \mathbb{Q}] \leq r$  and  $h(x) \leq C$ .*

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 [Lang, 1991, Ch. 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.

### 3 Roth's Theorem

K. F. Roth [1955] 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 3.1.** (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*

$$(3.1.1) \quad \left| \frac{a}{b} - \alpha \right| \leq \frac{C}{|b|^{2+\varepsilon}}.$$

*Example 3.2.* As a diophantine application of Roth's theorem, consider the diophantine equation

$$(3.2.1) \quad x^3 - 2y^3 = 11, \quad x, y \in \mathbb{Z}.$$

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 (3.2.1) has only finitely many solutions.

More generally, if  $f \in \mathbb{Z}[x, y]$  is homogeneous of degree  $\geq 3$  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 (3.1.1) 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 [Lang and Cherry, 1990, Intro. to Ch. I].

If  $a/b$  is close to  $\alpha$ , then after adjusting  $C$  one can replace  $|b|$  in the right-hand side of (3.1.1) with  $H_{\mathbb{Q}}(a/b)$  (see (2.2)). Moreover, the theorem has been generalized to allow a finite set of places (possibly non-archimedean) and to work over a number field:



**Theorem 3.3.** *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*

$$(3.3.1) \quad \prod_{v \in S} \min\{1, \|x - \alpha_v\|_v\} \leq \frac{C}{H_k(x)^{2+\varepsilon}}.$$

(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 Section 5. See, for example, (5.3).)

Taking  $-\log$  of both sides of (3.3.1), 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

$$(3.4) \quad \frac{1}{[k : \mathbb{Q}]} \sum_{v \in S} \log^+ \left\| \frac{1}{x - \alpha_v} \right\|_v \leq (2 + \varepsilon)h(x) + c$$

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

In writing (3.3.1), we assume that one has chosen an embedding  $i_v : \bar{k} \hookrightarrow \bar{k}_v$  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  $k, S, \varepsilon$ , 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  $(i_v)|_L : L \rightarrow \bar{k}_v$  induces a place  $w_0$  of  $L$  over  $v$ , and all other places  $w$  of  $L$  over  $v$  are conjugates by elements  $\sigma_w \in \text{Gal}(L/k)$ :  $\|x\|_w = \|\sigma_w^{-1}(x)\|_{w_0}$  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 (1.4), and therefore

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

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 (3.4) for almost all  $x \in k$ .

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

**Theorem 3.5.** *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*

$$(3.5.1) \quad \frac{1}{[k:\mathbb{Q}]} \sum_{v \in S} \sum_{i=1}^q \log^+ \left\| \frac{1}{x - \alpha_i} \right\|_v \leq (2 + \varepsilon)h(x) + c$$

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

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

so Theorem 3.5 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  $\alpha_i$  at each place  $v$  (where the value of  $i$  may depend on  $v$ ). Therefore, for each  $v$ ,

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

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  $\alpha_v$  for each  $v \in S$  such that

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

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 3.5.

## 4 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  $\infty$ . When  $r$  is large, the circle  $|z| = r$  is mapped to many values close to  $\infty$  (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

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

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

**Definition 4.1.** Let  $f$  be a meromorphic function on  $\mathbb{C}$ . We define the **proximity function** of  $f$  by

$$(4.1.1) \quad m_f(r) = \int_0^{2\pi} \log^+ |f(re^{i\theta})| \frac{d\theta}{2\pi}$$

for all  $r > 0$ . We also define

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

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

The integral in (4.1.1) 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 4.2.** 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

$$(4.2.1) \quad N_f(r) = \int_0^r (n_f(s) - n_f(0)) \frac{ds}{s} + n_f(0) \log r .$$

As before, we also define

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

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

The counting function can also be written

$$(4.3) \quad N_f(a, r) = \sum_{0 < |z| < r} \text{ord}_z^+(f - a) \cdot \log \frac{r}{|z|} + \text{ord}_0^+(f - a) \cdot \log r .$$

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 4.4.** Let  $f$  be as in Definition 4.1. Then the **height function** of  $f$  is the function  $T_f: (0, \infty) \rightarrow \mathbb{R}$  given by

$$(4.4.1) \quad T_f(r) = m_f(r) + N_f(r) .$$

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,

$$(4.5) \quad \liminf_{r \rightarrow \infty} \frac{T_f(r)}{\log r} > 0 .$$

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 non-constant. By [Hayman, 1964, Thm. 1.8], if  $f$  is entire and nonconstant and  $K > 1$ , then

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

This implies that  $f(z)/z^n$  has a removable singularity at  $\infty$  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  $\infty$ .

**Theorem 4.6.** (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**

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

where  $c_f$  is the leading coefficient in the Laurent expansion of  $f$  at  $z = 0$ . For details, see [Nevanlinna, 1970, Ch. VI, (1.2')] or [Ru, 2001, 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  $\pi i$ , 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 4.7.** (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*

$$(4.7.1) \quad \sum_{j=1}^q m_f(a_j, r) \leq_{\text{exc}} 2T_f(r) + O(\log^+ T_f(r)) + o(\log r),$$

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

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

By the First Main Theorem, (4.7.1) can be rewritten as a lower bound on the counting functions:

$$(4.8) \quad \sum_{j=1}^q N_f(a_j, r) \geq_{\text{exc}} (q-2)T_f(r) - O(\log^+ T_f(r)) - o(\log r).$$

As another variation, (4.7.1) can be written with a weaker error term:

$$(4.9) \quad \sum_{j=1}^q m_f(a_j, r) \leq_{\text{exc}} (2 + \varepsilon)T_f(r) + c$$

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

**Corollary 4.10.** (Picard's "little" theorem) *If  $a_1, a_2, a_3 \in \mathbb{P}^1(\mathbb{C})$  are distinct, then any holomorphic function  $f: \mathbb{C} \rightarrow \mathbb{P}^1(\mathbb{C}) \setminus \{a_1, a_2, a_3\}$  must be constant.*

*Proof.* Assume that  $f: \mathbb{C} \rightarrow \mathbb{P}^1(\mathbb{C}) \setminus \{a_1, a_2, a_3\}$  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 (4.8) vanishes. Since  $f$  is nonconstant, the right-hand side approaches  $+\infty$  by (4.5). This is a contradiction.  $\square$

As we have seen, (4.8) has some advantages over (4.7.1). 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.

## 5 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 5.1.** 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  $a \neq x$ ,

$$(5.1.1) \quad m_S(a, x) = m_S\left(\frac{1}{x-a}\right) = \sum_{v \in S} \log^+ \left\| \frac{1}{x-a} \right\|_v.$$

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

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

and

$$(5.1.2) \quad N_S(a, x) = N_S\left(\frac{1}{x-a}\right) = \sum_{v \notin S} \log^+ \left\| \frac{1}{x-a} \right\|_v.$$

By (2.4) it then follows that

$$(5.2) \quad m_S(x) + N_S(x) = \sum_{v \in M_k} \log^+ \|x\|_v = h_k(x)$$

for all  $x \in k$ . This corresponds to (4.4.1).

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  $\mathfrak{p} \subseteq \mathcal{O}_k$ . Thus, by (1.1), (5.1.2) can be rewritten

$$(5.3) \quad N_S(a, x) = \sum_{v \notin S} \text{ord}_{\mathfrak{p}}^+(x - a) \cdot \log(\mathcal{O}_k : \mathfrak{p}) ,$$

where  $\mathfrak{p}$  in the summand is the prime ideal corresponding to  $v$ . This corresponds to (4.3).

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

**Lemma 5.4.** *Let  $v$  be a place of a number field  $k$ , and let  $a, x \in k$ . Then*

$$(5.4.1) \quad \left| \log^+ \|x\|_v - \log^+ \|x - a\|_v \right| \leq \log^+ \|a\|_v + N_v \log 2 .$$

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

We first claim that

$$(5.4.2) \quad \log^+(s + t) \leq \log^+ s + \log^+ t + \log 2$$

for all real  $s, t \geq 0$ . 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  $f(s, t) \leq f(1, 1) = \log 2$ .

Now let  $z, b \in \mathbb{C}$ . Since  $|z| \leq |z - b| + |b|$ , (5.4.2) with  $s = |z - b|$  and  $t = |b|$  gives

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

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

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

These two inequalities together imply (5.4.1).

**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

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

since  $0 \leq \log^+ \|x - a\|_v \leq \log^+ \|x\|_v$ . If  $\|a\|_v = \|x - a\|_v$  then (5.4.1) follows by a similar argument.  $\square$

Corresponding to Theorem 4.6, we then have:

**Theorem 5.5.** *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 5.4 termwise to the sums in the two height functions.  $\square$

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 5.1, Roth's theorem can be made to look very similar to Nevanlinna's Second Main Theorem. Indeed, multiplying (3.5.1) by  $[k : \mathbb{Q}]$  and substituting the definition (5.1.1) of the proximity function gives the inequality

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

which corresponds to (4.9). As has been mentioned earlier, it has been conjectured that Roth's theorem should hold with sharper error terms, corresponding to (4.7.1). 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, [Wong, 1989], [Lang and Cherry, 1990], and [Cherry and Ye, 2001].

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

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

Similarly, let  $S \supseteq S_\infty$  be a finite set of places of a number field  $k$ , let  $a \in k$ , and let  $\Sigma$  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 4.6 and 5.5), we then have

$$0 \leq \delta_f(a) \leq 1 \quad \text{and} \quad 0 \leq \delta_S(a) \leq 1 ,$$

respectively. The Second Main Theorems (Theorems 4.7 and 3.5) then give

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

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 its 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 5.1 can be extended to  $x \in \bar{k}$ . Indeed, let  $k$  and  $S$  be as in Definition 5.1, 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 (1.4) gives

$$(5.7) \quad m_{T'}(x) = [L' : L]m_T(x) \quad \text{and} \quad N_{T'}(x) = [L' : L]N_T(x) .$$

This allows us to make the following definition.

**Definition 5.8.** Let  $k, S, x, L$ , and  $T$  be as above. Then we define

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

These expressions are independent of  $L \supseteq k(x)$  by (5.7). As in (5.1.1) and (5.1.2), we also let

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

Likewise, Theorem 5.5 (the number-theoretic First Main Theorem) extends to  $x \in \bar{k}$ , by (2.5), (5.7), and (5.8). The expression (5.2) 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.

## 6 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 C. F. Osgood [1981, 1985], but he did not provide an explicit dictionary for comparing the two theories. This was provided by Vojta [1987]. 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 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 Section 5, 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  $\mathbb{D}_r$  of radius  $r$ . Of course  $f|_{\mathbb{D}_r}$  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  $f|_{\mathbb{D}_r}$  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.

**Table 1** The Dictionary in the One-Dimensional Case

Nevanlinna Theory	Number Theory
$f: \mathbb{C} \rightarrow \mathbb{C}$ , non-constant	$\{x\} \subseteq k$ , infinite
$r$	$x$
$\theta$	$v \in S$
$ f(re^{i\theta}) $	$\ x\ _v, v \in S$
$\text{ord}_z f$	$\text{ord}_v x, v \notin S$
$\log \frac{r}{ z }$	$\log(\mathcal{O}_k : \mathfrak{p})$
Height function	Logarithmic height
$T_f(r) = \int_0^{2\pi} \log^+  f(re^{i\theta})  \frac{d\theta}{2\pi} + N_f(\infty, r)$	$h_k(x) = \sum_{v \in M_k} \log^+ \ x\ _v$
Proximity function	
$m_f(a, r) = \int_0^{2\pi} \log^+ \left  \frac{1}{f(re^{i\theta}) - a} \right  \frac{d\theta}{2\pi}$	$m_S(a, x) = \sum_{v \in S} \log^+ \left\  \frac{1}{x - a} \right\ _v$
Counting function	
$N_f(a, r) = \sum_{ z  < r} \text{ord}_z^+(f - a) \log \frac{r}{ z }$	$N_S(a, x) = \sum_{v \notin S} \text{ord}_v^+(x - a) \log(\mathcal{O}_k : \mathfrak{p})$
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
$\sum_{i=1}^m m_f(a_i, r) \leq_{\text{exc}} 2T_f(r) - N_{1,f}(r) + O(r \log T_f(r))$	$\sum_{i=1}^m m_S(a_i, x) \leq 2h_k(x) + O(\log h_k(x))$
Defect	
$\delta(a) = \liminf_{r \rightarrow \infty} \frac{m_f(a, r)}{T_f(r)}$	$\delta(a) = \liminf_x \frac{m_S(a, x)}{h_k(x)}$
Defect Relation	Roth's theorem
$\sum_{a \in \mathbb{C}} \delta(a) \leq 2$	$\sum_{a \in k} \delta(a) \leq 2$
Jensen's formula	Artin-Whaples Product Formula
$\log  c_f  = \int_0^{2\pi} \log  f(re^{i\theta})  \frac{d\theta}{2\pi} + N_f(\infty, r) - N_f(0, r)$	$\sum_{v \in M_k} \log \ x\ _v = 0$

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  $\pi: B \rightarrow \mathbb{C}$  be a proper surjective holomorphic map, and let  $f: B \rightarrow \mathbb{C}$  be a meromorphic function. In place of  $\mathbb{D}_r$  in the above discussion, one can work with  $\pi^{-1}(\mathbb{D}_r)$  and define the proximity, counting, and height functions accordingly. For detailed definitions, see Section 26.

When working with a finite ramified covering, though, the Second Main Theorem requires an additional term  $N_{\text{Ram}(\pi)}(r)$ , which is a counting function for ramification points of  $\pi$  (Definition 26.3c). The main inequality (4.7.1) of the Second Main Theorem then becomes

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

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 6.1.** 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

$$(6.2) \quad \sum_{j=1}^q m_S(a_j, x) \leq (2 + \varepsilon)h_k(x) + d_k(x) + C.$$

For further discussion of this situation, including its relation to the abc conjecture, see Sections 24–25.

## 7 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 7.1.** A collection of hyperplanes in  $\mathbb{P}^n$  is in **general position** if for all  $j \leq n$  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 [1933]. Before stating it, we need to define the proximity, counting, and height functions.

**Definition 7.2.** 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  $[x_0 : \cdots : x_n]$  be homogeneous coordinates for  $P$ . We then define

$$(7.2.1) \quad \lambda_H(P) = -\frac{1}{2} \log \frac{|a_0x_0 + \cdots + a_nx_n|^2}{|x_0|^2 + \cdots + |x_n|^2}$$

(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

$$(7.3) \quad \lambda_H(x) = \log^+ \left| \frac{1}{x-a} \right| + O(1).$$

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

**Definition 7.4.** Let  $n, H$ , and  $\lambda_H$  be as in Definition 7.2, and let  $f: \mathbb{C} \rightarrow \mathbb{P}^n$  be a holomorphic curve whose image is not contained in  $H$ . Then the **proximity function** for  $H$  is

$$(7.4.1) \quad m_f(H, r) = \int_0^{2\pi} \lambda_H(f(re^{i\theta})) \frac{d\theta}{2\pi}.$$

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 7.5.** 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  $\text{ord}_z f^*H$  denote its multiplicity at the point  $z$ . Then the **counting function** for  $H$  is defined to be

$$(7.5.1) \quad N_f(H, r) = \sum_{0 < |z| < r} \text{ord}_z f^*H \cdot \log \frac{r}{|z|} + \text{ord}_0 f^*H \cdot \log r$$

**Definition 7.6.** Let  $f: \mathbb{C} \rightarrow \mathbb{P}^n(\mathbb{C})$  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 7.7.** (Cartan) *Let  $n > 0$  and let  $H_1, \dots, H_q$  be hyperplanes in  $\mathbb{P}^n$  in general position. Let  $f: \mathbb{C} \rightarrow \mathbb{P}^n(\mathbb{C})$  be a holomorphic curve whose image is not contained in any hyperplane. Then*

$$(7.7.1) \quad \sum_{j=1}^q m_f(H_j, r) \leq_{\text{exc}} (n+1)T_f(r) + O(\log^+ T_f(r)) + o(\log r).$$

If  $n = 1$  then by (7.3) this reduces to the classical Second Main Theorem (Theorem 4.7).

Inequality (7.7.1) can also be expressed using counting functions as

$$(7.8) \quad \sum_{j=1}^q N_f(H_j, r) \geq_{\text{exc}} (q-n-1)T_f(r) - O(\log^+ T_f(r)) - o(\log r)$$

(cf. (4.8)).

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 7.9.** Let  $H$  be a hyperplane in  $\mathbb{P}_k^n$  and let  $a_0x_0 + \dots + a_nx_n = 0$  be a linear form defining it. (Since  $\mathbb{P}_k^n$  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

$$(7.9.1) \quad \lambda_{H,v}(P) = -\log \frac{\|a_0x_0 + \dots + a_nx_n\|_v}{\max\{\|x_0\|_v, \dots, \|x_n\|_v\}},$$

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 8.6), with domain restricted to  $\mathbb{P}^n(k) \setminus H$ .

**Definition 7.10.** For  $H$  and  $P$  as above, the **proximity function** for  $H$  is defined to be

$$(7.10.1) \quad m_S(H, P) = \sum_{v \in S} \lambda_{H,v}(P),$$

and the **counting function** is defined by

$$(7.10.2) \quad N_S(H, P) = \sum_{v \notin S} \lambda_{H,v}(P).$$

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 7.7 (with, of course, a weaker error term) is a slightly weaker form of Schmidt's Subspace Theorem.

**Theorem 7.11.** (Schmidt) *Let  $k$ ,  $S$ , and  $n$  be as above, let  $H_1, \dots, H_q$  be hyperplanes in  $\mathbb{P}_k^n$  in general position, let  $\varepsilon > 0$ , and let  $c \in \mathbb{R}$ . Then*

$$(7.11.1) \quad \sum_{j=1}^q m_S(H_j, x) \leq (n + 1 + \varepsilon)h_k(x) + c$$

for all  $x \in \mathbb{P}^n(k)$  outside of a finite union of proper linear subspaces of  $\mathbb{P}_k^n$ . This latter set depends on  $k$ ,  $S$ ,  $H_1, \dots, H_q$ ,  $\varepsilon$ ,  $c$ , 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 3.5).

Note, in particular, that the  $H_i$  are hyperplanes in the  $k$ -scheme  $\mathbb{P}_k^n$ . This automatically implies that they can be defined by linear forms with coefficients in  $k$ . This corresponds to requiring the  $\alpha_j$  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  $\mathbb{P}_k^n$  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  $\mathbb{A}_k^{n+1}$  passing through the origin and points in  $\mathbb{A}_k^{n+1}$  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 [Vojta, 1987, Ch. 2 § 2].

Theorem 7.11 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, (7.11.1) would need to be replaced by

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

where for each  $v \in S$ ,  $H_{v,1}, \dots, H_{v,q_v}$  are hyperplanes in general position (but in totality the set  $\{H_{v,j} : v \in S, 1 \leq j \leq q_v\}$  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 Section 5, is as follows. It has been stated in a form that most readily carries over to Nevanlinna theory.

**Theorem 7.12.** (Schmidt's Subspace Theorem [Schmidt, 1991, Ch. VIII, Thm. 7A]) *Let  $k$ ,  $S$ , and  $n$  be as above, and let  $H_1, \dots, H_q$  be distinct hyperplanes in  $\mathbb{P}_k^n$ . Then for all  $\varepsilon > 0$  and all  $c \in \mathbb{R}$  the inequality*

$$(7.12.1) \quad \sum_{v \in S} \max_J \sum_{j \in J} \lambda_{H_j, v}(x) \leq (n+1+\varepsilon)h_k(x) + c$$

*holds for all  $x \in \mathbb{P}^n(k)$  outside of a finite union of proper linear subspaces depending only on  $k$ ,  $S$ ,  $H_1, \dots, H_q$ ,  $\varepsilon$ ,  $c$ , 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  $\theta$ , so if one allowed the collection of hyperplanes to vary with  $\theta$  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 7.12 was stated in the way that it was.

Cartan's theorem itself can be generalized as follows.

**Theorem 7.13.** [Vojta, 1997] *Let  $n \in \mathbb{Z}_{>0}$ , let  $H_1, \dots, H_q$  be hyperplanes in  $\mathbb{P}_{\mathbb{C}}^n$ , and let  $f: \mathbb{C} \rightarrow \mathbb{P}^n(\mathbb{C})$  be a holomorphic curve whose image is not contained in a hyperplane. Then*

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

*where  $J$  varies over the same collection of sets as in Theorem 7.12.*

This has proved to be a useful formulation for applications; see [Vojta, 1997] and [Ru, 1997]. The latter reference also improves the error term in Theorem 7.13.



*Remark 7.14.* It has been further shown that in Theorem 7.12, 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 [Vojta, 1989]. In other words, the higher-dimensional part of the exceptional set depends only on the geometric data. Correspondingly, Theorem 7.13 holds for all nonconstant holomorphic curves whose image is not contained in the union of this latter set [Vojta, 1997]. For an example of the collection of higher dimensional subspaces for a specific set of lines in  $\mathbb{P}^2$ , see Example 13.3.

## 8 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 [Hartshorne, 1977], except that varieties (when discussing number theory at least) may be defined over a field that is not necessarily algebraically closed.

**Definition 8.1.** A **variety** over a field  $k$ , or a  $k$ -variety, is an integral separated scheme of finite type over  $k$  (i.e., over  $\text{Spec } k$ ). 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  $\text{Spec } k$ ).

As an example,  $X := \text{Spec } \mathbb{Q}[x, y]/(y^2 - 2x^2)$  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  $X := \text{Spec } \mathbb{Q}[x, y]/(y^2 - 2x^2)$  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  $\mathbb{A}_{\mathbb{Q}}^2$ , which does not come from any subvariety of  $\mathbb{A}_{\mathbb{Q}}^2$  (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  $\xi$  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 [Weil, 1928] and further developed in a later paper [Weil, 1951]. 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 [Lang, 1983, Ch. 10].

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

Weil functions are best described using Cartier divisors.

**Definition 8.2.** Let  $D$  be a Cartier divisor on a complex variety  $X$ . Then a **Weil function** for  $D$  is a function  $\lambda_D: (X \setminus \text{Supp } D)(\mathbb{C}) \rightarrow \mathbb{R}$  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  $D|_U = (f)$ , and a continuous function  $\alpha: U(\mathbb{C}) \rightarrow \mathbb{R}$  such that

$$(8.2.1) \quad \lambda_D(x) = -\log |f(x)| + \alpha(x)$$

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

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

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  $\mathcal{L}$  is the corresponding line sheaf, then there is a nonzero rational section  $s$  of  $\mathcal{L}$  whose vanishing describes  $D$ :  $D = (s)$ . As was noted by Néron, Weil functions on  $D$  correspond to metrics on  $\mathcal{L}$ .

Recall that if  $X$  is a complex variety and  $\mathcal{L}$  is a line sheaf on  $X$ , then a **metric** on  $\mathcal{L}$  is a collection of norms on the fibers of the complex line bundle corresponding to the sheaf  $\mathcal{L}$ , 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  $C^\infty$ . If  $X$  is singular, then we say that a function  $f: X(\mathbb{C}) \rightarrow \mathbb{C}$  is  $C^\infty$  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  $\phi: U \rightarrow \mathbb{C}^n$  for some  $n$ , and a  $C^\infty$  function  $g: \mathbb{C}^n \rightarrow \mathbb{C}$  such that  $f = g \circ \phi$ . This reduces to the usual concept of  $C^\infty$  function at smooth points of  $X$ .

To describe a metric on  $\mathcal{L}$  in concrete terms, let  $U$  be an open subset of  $X$  and let  $\phi_U: \mathcal{O}_U \xrightarrow{\sim} \mathcal{L}|_U$  be a local trivialization. Then the function  $\rho_U: U(\mathbb{C}) \rightarrow \mathbb{R}_{>0}$  given by  $\rho_U(x) = |\phi_U(1)(x)|$  is smooth (resp. continuous), and for any section  $s \in \mathcal{L}(U)$  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  $\phi_V : \mathcal{O}_V \xrightarrow{\sim} \mathcal{L}|_V$  is a local trivialization on  $V$ , then  $\phi_U^{-1} \circ \phi_V$  (appropriately restricted) is an automorphism of  $\mathcal{O}_{U \cap V}$  corresponding to multiplication by a function  $\alpha_{UV} \in \mathcal{O}_{U \cap V}^*$ . Again letting  $\rho_V(x) = |\phi_V(1)(x)|$ , we see that  $\rho_U$  and  $\rho_V$  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  $\mathcal{U}$  of  $X$  and  $\alpha_{UV} \in \mathcal{O}_{U \cap V}^*$  for all  $U, V \in \mathcal{U}$  satisfying  $\alpha_{UU} = 1$  and  $\alpha_{UV} = \alpha_{UV}\alpha_{VW}$  on  $U \cap V \cap W$  for all  $U, V, W \in \mathcal{U}$ . Moreover, with these data, one can specify a metric on the associated line sheaf by giving smooth or continuous functions  $\rho_U : U(\mathbb{C}) \rightarrow \mathbb{R}_{>0}$  for each  $U \in \mathcal{U}$  that satisfy  $\rho_V = |\alpha_{UV}|\rho_U$  on  $U \cap V$  for all  $U, V \in \mathcal{U}$ .

A continuous metric on a line sheaf  $\mathcal{L}$  determines a Weil function for any associated Cartier divisor  $D$ . Indeed, if  $s$  is a nonzero rational section of  $\mathcal{L}$  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  $\mathcal{L}$ .

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  $\alpha$  in (8.2.1) are all smooth).

An example of a Weil function in Nevanlinna theory (and perhaps the primary example) is the function  $\lambda_H$  of Definition 7.2 used in Cartan’s theorem.

Likewise, the function  $\lambda_{H,v}$  of Definition 7.9 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 5.4 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 5.5.

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

**Definition 8.3.** 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 [Koblitz, 1984, Ch. III, § 3–4] 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  $f \in \mathcal{O}(U)$ ,  $U(\mathbb{C}_v)$  is open and  $f : U(\mathbb{C}_v) \rightarrow \mathbb{C}_v$  is continuous.

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

**Definition 8.4.** 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 8.5.** 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  $\alpha: X(M_k) \rightarrow \mathbb{R}$  is said to be  $O_{M_k}(1)$  if there is an  $M_k$ -constant  $(c_v)$  such that  $|\alpha(x)| \leq c_v$  for all  $x \in X(\mathbb{C}_v)$  and all  $v \in M_k$ .

We may then define Weil functions as follows.

**Definition 8.6.** 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  $\lambda_D: (X \setminus \text{Supp } D)(M_k) \rightarrow \mathbb{R}$  that satisfies the following condition. For each  $x \in X$  there is an open neighborhood  $U$  of  $x$ , a nonzero function  $f \in \mathcal{O}(U)$  such that  $D|_U = (f)$ , and a continuous locally  $M_k$ -bounded function  $\alpha: U(M_k) \rightarrow \mathbb{R}$  satisfying

$$(8.6.1) \quad \lambda_D(x) = -\log \|f(x)\|_v + \alpha(x)$$

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

For the definition of locally  $M_k$ -bounded function, see [Lang, 1983, Ch. 10, § 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 8.2, if  $\lambda$  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  $f \in \mathcal{O}(U)$  for which  $D|_U = (f)$ .

If  $\lambda_D$  is a Weil function for  $D$ , then we write

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

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 8.2 (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, (8.6.1) could be shortened to  $\lambda(x) = -\log \|f(x)\| + \alpha(x)$  for all  $x \in (U \setminus \text{Supp } D)(M_k)$ .

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

**Theorem 8.7.** *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 [Lang, 1983, Ch. 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 8.8.** *Let  $X$  be a complete variety over a number field  $k$ . Then*

- (a). **Additivity:** If  $\lambda_1$  and  $\lambda_2$  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  $\lambda$  is a Weil function for a Cartier divisor  $D$  on  $X$ , and if  $f: X' \rightarrow X$  is a morphism of  $k$ -varieties such that  $f(X') \not\subseteq \text{Supp} D$ , then  $x \mapsto \lambda(f(x))$  is a Weil function for the Cartier divisor  $f^*D$  on  $X'$ .
- (c). **Normalization:** If  $X = \mathbb{P}_k^n$ , and if  $D$  is the hyperplane at infinity, then the function

$$(8.8.1) \quad \lambda_{D,v}([x_0 : \cdots : x_n]) := -\log \frac{\|x_0\|_v}{\max\{\|x_0\|_v, \dots, \|x_n\|_v\}}$$

is a Weil function for  $D$ .

- (d). **Uniqueness:** If both  $\lambda_1$  and  $\lambda_2$  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  $\lambda$  is a Weil function for  $D$ , then  $\lambda$  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 8.9.** *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  $\text{Supp} E_1 \cap \cdots \cap \text{Supp} E_n = \emptyset$ . Then the function*

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

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

*Proof.* See [Lang, 1983, Ch. 10, Prop. 3.2]. □

**Theorem 8.10.** (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*

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

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  $\bigcap \text{Supp} E_i = \emptyset$  (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  $\bigcap \text{Supp } F_j = \emptyset$ . 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 8.9 to  $-\log \|f_{ij}\|$  then implies that  $\min_{1 \leq j \leq m} (-\log \|f_{ij}\|)$  is a Weil function for  $D - E_i$  for all  $i$ . Applying Lemma 8.9 again to the negatives of these Weil functions then gives the theorem.  $\square$

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

**Definition 8.11.** Let  $X$  be a variety over a number field  $k$ , let  $D$  be a Cartier divisor on  $X$ , and let  $\lambda_D$  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

$$(8.11.1) \quad \lambda_{D,w} = [L_w : k_v] \lambda_{D,v} .$$

(Recall that  $\|x\|_w = \|x\|_v^{[L_w:k_v]}$  for all  $x \in \mathbb{C}_v$ , by (1.4).) 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

$$(8.11.2) \quad \lambda_{D,w}(x) = \lambda_{D,w}(x_w)$$

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

Note that, if  $x \in (X \setminus \text{Supp } D)(L)$ , 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

$$(8.12) \quad \lambda_{D,w'}(x) = [L_{w'} : L_w] \lambda_{D,w}(x) ,$$

regardless of whether the left-hand side is defined using (8.11.1) or (8.11.2) (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

$$(8.13) \quad c_w = [L_w : k_v] c_v ,$$

so that the condition  $\lambda_{D,w} \leq c_w$  will be equivalent to  $\lambda_{D,v} \leq c_v$ , by (8.11.1).

## 9 Height Functions on Varieties in Number Theory

Weil functions can be used to generalize the height  $h_k$  (defined in Section 2) to arbitrary complete varieties over  $k$ . This can also be done by working directly with heights; see [Lang, 1983, Ch. 3] or [Silverman, 1986].

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

$$(9.1.1) \quad h_{\lambda,k}(x) = \frac{1}{[L:k]} \sum_{w \in M_L} \lambda_w(x)$$

for any number field  $L \supseteq \kappa(x)$ . It is independent of the choice of  $L$  by (8.12).

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  $X = \mathbb{P}_k^n$ , if  $D$  is the hyperplane at infinity, and if  $\lambda$  is the Weil function (8.8.1), then

$$(9.2) \quad \begin{aligned} h_{\lambda,k}([x_0 : \cdots : x_n]) &= -\frac{1}{[L:k]} \sum_{w \in M_L} \log \frac{\|x_0\|_w}{\max\{\|x_0\|_w, \dots, \|x_n\|_w\}} \\ &= \frac{1}{[L:k]} \sum_{w \in M_L} \log \max\{\|x_0\|_w, \dots, \|x_n\|_w\} \\ &= h_k([x_0 : \cdots : x_n]) \end{aligned}$$

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

The restriction  $x \notin \text{Supp } D$  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  $\text{Supp } D \cup \text{Supp } D'$ , and if  $L$  is a number field containing  $\kappa(x)$ , then

$$(9.3) \quad \begin{aligned} h_{\lambda',k}(x) &= \frac{1}{[L:k]} \sum_{w \in M_L} \lambda'_w(x) \\ &= \frac{1}{[L:k]} \sum_{w \in M_L} \lambda_w(x) - \frac{1}{[L:k]} \sum_{w \in M_L} \log \|f(x)\|_w \\ &= h_{\lambda,k}(x) \end{aligned}$$

by the Product Formula (1.6). Thus we have:

**Definition 9.4.** Let  $\lambda$  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 9.1. This expression is independent of the choice of  $f$  by (9.3), and agrees with Definition 9.1 when  $x \notin \text{Supp } D$  since we can take  $f = 1$  in that case.

With this definition, (9.2) holds without the restriction  $x_0 \neq 0$ .

**Proposition 9.5.** *If both  $\lambda$  and  $\lambda'$  are Weil functions for  $D$ , then*

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

*Proof.* Indeed, this is immediate from Theorem 8.8d. □

Thus, the height function defined above depends only on the divisor; moreover, by (9.3) it depends only on the linear equivalence class of the divisor.

**Definition 9.6.** 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  $\lambda$  for  $D$ . If  $\mathcal{L}$  is a line sheaf on  $X$ , then we define

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

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

By (9.2), we then have

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

on  $\mathbb{P}_k^n$  for all  $n > 0$ . Since the automorphism group of  $\mathbb{P}_k^n$  is transitive on the set of rational points, and since automorphisms preserve the line sheaf  $\mathcal{O}(1)$ , the term  $O(1)$  in the above formula cannot be eliminated without additional structure. Thus, Definition 9.6 cannot give an exact definition for the height without additional structure. (This additional structure can be given using Arakelov theory.)

Theorem 8.8 and (8.12) also immediately imply the following properties of heights:

**Theorem 9.7.** (a). **Functoriality:** *if  $f: X' \rightarrow X$  is a morphism of  $k$ -varieties, and if  $\mathcal{L}$  is a line sheaf on  $X$ , then*

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

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



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

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

for all  $x \in X(\bar{k})$ , where the implied constant depends only on  $\mathcal{L}_1, \mathcal{L}_2$ , 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  $\mathcal{L}$  is a line sheaf on  $X$ , and is generated by its global sections, then  $h_{\mathcal{L}, k}(x)$  is bounded from below for all  $x \in X(\bar{k})$ , by a bound depending only on  $\mathcal{L}$  and the choice of height function.
- (e). **Change of number field:** If  $L \supseteq k$  then

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

for all line sheaves  $\mathcal{L}$  on  $X$  and all  $x \in X(\bar{k})$ . (Strictly speaking, the left-hand side should be  $h_{\mathcal{L}', L}(x')$ , where  $\mathcal{L}'$  is the pull-back of  $\mathcal{L}$  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 9.8.** If  $\mathcal{L}$  is an ample line sheaf on  $X$ , then  $h_{\mathcal{L}, k}(x)$  is bounded from below for all  $x \in X(\bar{k})$ , by a bound depending only on  $\mathcal{L}$  and the choice of height function.

*Proof.* By Theorem 9.7b, we may replace  $\mathcal{L}$  with  $\mathcal{L}^{\otimes n}$  for any positive integer  $n$ , and therefore may assume that  $\mathcal{L}$  is very ample. Then the result follows from Theorem 9.7d. □

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

**Proposition 9.9.** Let  $\mathcal{L}$  and  $\mathcal{M}$  be line sheaves on  $X$ , with  $\mathcal{L}$  ample. Then there is a constant  $C$ , depending only on  $\mathcal{L}$  and  $\mathcal{M}$ , such that

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

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

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

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

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

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

**Theorem 9.10.** (Northcott) Assume that  $X$  is projective, and let  $\mathcal{L}$  be an ample line sheaf on  $X$ . Then, for all integers  $d$  and all  $c \in \mathbb{R}$ , the set

$$(9.10.1) \quad \{x \in X(\bar{k}) : [\kappa(x) : k] \leq d \text{ and } h_{\mathcal{L},k}(x) \leq c\}$$

is finite.

*Proof.* First, if  $X = \mathbb{P}_k^n$  and  $\mathcal{L} = \mathcal{O}(1)$ , then the result follows by bounding the heights of  $[x_i : x_0]$  (if  $x_0 \neq 0$ , which we assume without loss of generality), and applying Theorem 2.7 to these points for each  $i$ . The general case then follows by replacing  $\mathcal{L}$  with a very ample positive multiple and using an associated projective embedding and functoriality of heights.  $\square$

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

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

for all sufficiently large and divisible  $n$ . A Cartier divisor  $D$  on  $X$  is big if  $\mathcal{O}(D)$  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  $\pi: X' \rightarrow X$ . If  $\mathcal{L}$  is a big line sheaf on  $X$ , then  $\pi^*\mathcal{L}$  is big on  $X'$ . Therefore, it makes some sense to compare big line sheaves with ample ones.

**Proposition 9.12.** (Kodaira's lemma) *Let  $X$  be a projective variety over an arbitrary field, and let  $\mathcal{L}$  and  $\mathcal{A}$  be line sheaves on  $X$ , with  $\mathcal{A}$  ample. Then  $\mathcal{L}$  is big if and only if there is a positive integer  $n$  such that  $H^0(X, \mathcal{L}^{\otimes n} \otimes \mathcal{A}^{\vee}) \neq 0$ . 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 [Vojta, 1987, Prop. 1.2.7].  $\square$

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

**Proposition 9.13.** *Let  $X$  be a complete variety over a number field. Let  $\mathcal{L}$  and  $\mathcal{M}$  be line sheaves on  $X$ , with  $\mathcal{L}$  big. Then there is a constant  $C$  and a proper Zariski-closed subset  $Z$  of  $X$ , depending only on  $\mathcal{L}$  and  $\mathcal{M}$ , such that*

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

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

*Proof.* After applying Chow’s lemma and pulling back  $\mathcal{L}$  and  $\mathcal{M}$ , we may assume that  $X$  is projective. We may also replace  $\mathcal{L}$  with a positive multiple, and hence may assume that  $\mathcal{L}$  is isomorphic to  $\mathcal{A} \otimes \mathcal{O}(D)$ , where  $\mathcal{A}$  is an ample line sheaf and  $D$  is an effective Cartier divisor. Then the result follows from Proposition 9.9, with  $Z = \text{Supp} D$ , by Theorem 9.7.  $\square$

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 9.14.* For a general complete variety  $X$  over  $k$ , we can define a **big height** to be a function  $h: X(\bar{k}) \rightarrow \mathbb{R}$  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  $U_i \hookrightarrow \bar{U}_i$ , an ample line sheaf  $\mathcal{L}_i$  on  $\bar{U}_i$ , and real constants  $c_i > 0$  and  $C_i$  such that  $h(x) \geq c_i h_{\mathcal{L}_i, k}(x) + C_i$  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  $\mathcal{L}$  is an ample line sheaf on  $X$  then  $h_{\mathcal{L}, k}$  is a big height on  $X$ ;
- if  $\mathcal{L}$  is a line sheaf on  $X$  then there are real constants  $c$  and  $C$  such that  $h_{\mathcal{L}, k}(x) \leq ch(x) + C$  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 9.15 (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 29.3.

**Proposition 9.15.** *Let  $f: X_1 \rightarrow X_2$  be a morphism of projective varieties over a number field, and let  $\mathcal{A}_1$  and  $\mathcal{A}_2$  be ample line sheaves on  $X_1$  and  $X_2$ , respectively. Then there is a constant  $C$ , depending only on  $f$ ,  $\mathcal{A}_1$ , and  $\mathcal{A}_2$  such that*

$$(9.15.1) \quad h_{\mathcal{A}_1, k}(x) \leq Ch_{\mathcal{A}_2, k}(f(x)) + O(1)$$

for all points  $x \in X_1(\bar{k})$  that are isolated in their fibers of  $f$ , where the implied constant depends only on  $f$ ,  $\mathcal{A}_1$ ,  $\mathcal{A}_2$ , 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  $f^* \mathcal{A}_2$  is big. The result then follows by Proposition 9.13 and noetherian induction applied to the irreducible components of the set  $Z$  in that proposition  $\square$

Note that if any fiber component of  $f$  has dimension  $> 0$ , then it contains algebraic points of arbitrarily large height, so (9.15.1) cannot possibly hold for all such points.

## 10 Proximity and Counting Functions on Varieties in Number Theory

The definitions of proximity and counting functions given in Sections 5 and 7 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  $S_\infty$ , and  $X$  is a complete variety over  $k$ .

**Definition 10.1.** Let  $D$  be a Cartier divisor on  $X$ , let  $\lambda_D$  be a Weil function for  $D$ , let  $x \in X(\bar{k})$  with  $x \notin \text{Supp } D$ , 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

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

These expressions are independent of the choice of  $L$ , by (8.11.1). They depend on the choice of  $\lambda_D$  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  $x \notin \text{Supp } D$  cannot be eliminated. By (9.1.1) and Definition 9.6, 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 14.6).

Theorem 8.8 immediately implies the following properties of proximity and counting functions.

**Proposition 10.2.** *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  $f: X' \rightarrow X$  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 9.7e).

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.

## 11 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  $f: \mathbb{C} \rightarrow X$  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 [Hartshorne, 1977, App. B].

We begin with the proximity and counting functions.

**Definition 11.1.** Let  $\lambda$  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  $\lambda'$  is another Weil function for  $D$ , then  $|\lambda - \lambda'|$  is bounded, so the proximity function is independent of  $D$  (up to  $O(1)$ ).

**Definition 11.2.** The **counting function** for  $f$  relative to  $D$  is the function

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

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

Corresponding to Proposition 10.2, we then have

**Proposition 11.3.** *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  $\phi: X \rightarrow X'$  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*

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

(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 11.4.** 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 11.5.** *The height function  $T_{D,f}$  is additive in  $D$ , is functorial, and is bounded from below if  $D$  is effective, as in Proposition 11.3.*

*Proof.* Immediate from Proposition 11.3. □

**Proposition 11.6.** (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*

$$(11.6.1) \quad T_{D',f}(r) = T_{D,f}(r) + O(1) .$$

*Proof.* We first consider the special case  $X = \mathbb{P}_{\mathbb{C}}^1$ ,  $D = [0]$  (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 Section 4). The result then follows by Theorem 4.6 (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  $X \dashrightarrow \mathbb{P}_{\mathbb{C}}^1$ . Let  $X'$  be the closure of the graph, with projections  $p: X' \rightarrow X$  and  $q: X' \rightarrow \mathbb{P}_{\mathbb{C}}^1$ . By the additivity property of heights, (11.6.1) 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  $f: \mathbb{C} \rightarrow X$  lifts to a function  $f': \mathbb{C} \rightarrow X'$  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 11.7.** The **height function** of  $f$  relative to a line sheaf  $\mathcal{L}$  on  $X$  is defined to be  $T_{\mathcal{L},f}(r) = T_{D,f}(r) + O(1)$  for any divisor  $D$  such that  $\mathcal{O}(D) \cong \mathcal{L}$  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  $\mathcal{L}$ . 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 9.7.

**Theorem 11.8.** (a). **Functoriality:** if  $\phi: X \rightarrow X'$  is a morphism of complete complex varieties and if  $\mathcal{L}$  is a line sheaf on  $X'$ , then

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

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

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

(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  $\mathcal{L}$  is a line sheaf on  $X$ , and is generated by its global sections, then  $T_{\mathcal{L},f}(r)$  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 Section 9.

**Corollary 11.9.** If  $\mathcal{L}$  is an ample line sheaf on  $X$ , then  $T_{\mathcal{L},f}(r)$  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 [Goldberg and Ostrovskii, 2008, Ch 1, Thm. 6.4] for the second assertion. The general case is left as an exercise for the reader.  $\square$

**Proposition 11.10.** Let  $\mathcal{L}$  and  $\mathcal{M}$  be line sheaves on  $X$ , with  $\mathcal{L}$  ample. Then there is a constant  $C$ , depending only on  $\mathcal{L}$  and  $\mathcal{M}$ , such that

$$T_{\mathcal{M},f}(r) \leq CT_{\mathcal{L},f}(r) + O(1).$$

*Proof.* This is true for essentially the same reasons as Proposition 9.9. The details are left to the reader.  $\square$

**Proposition 11.11.** Let  $\mathcal{L}$  and  $\mathcal{M}$  be line sheaves on  $X$ , with  $\mathcal{L}$  big. Then there is a constant  $C$  and a proper Zariski-closed subset  $Z$  of  $X$ , depending only on  $\mathcal{L}$  and  $\mathcal{M}$ , such that

$$T_{\mathcal{M},f}(r) \leq CT_{\mathcal{L},f}(r) + O(1),$$

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

*Proof.* Similar to the proof of Proposition 9.13; details are again left to the reader.  $\square$



*Remark 11.12.* For an arbitrary complete variety  $X$  over  $\mathbb{C}$  and a holomorphic curve  $f: \mathbb{C} \rightarrow X$ , one can define a **big height** to be a real-valued function  $T_{\text{big},f}(r)$  with the property that if  $Z$  is the Zariski closure of the image of  $f$ , if  $\tilde{Z} \rightarrow Z$  is a proper birational morphism with  $\tilde{Z}$  projective, if  $\mathcal{L}$  is an ample line sheaf on  $\tilde{Z}$ , and if  $\tilde{f}: \mathbb{C} \rightarrow \tilde{Z}$  is a lifting of  $f$ , then there are constants  $c > 0$  and  $C$  such that  $T_{\text{big},f}(r) \geq cT_{\mathcal{L},\tilde{f}}(r) + C$  for all  $r > 0$ . This condition is independent of the choices of  $\mathcal{L}$  and  $\tilde{Z}$ . This height satisfies the same properties as in Remark 9.14. (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  $U_1, \dots, U_n, \bar{U}_1, \dots, \bar{U}_n$ , and  $\mathcal{L}_1, \dots, \mathcal{L}_n$  as in Remark 9.14; then extend  $f^{-1}(U_i) \rightarrow U_i$  to a map  $\mathbb{C} \rightarrow \bar{U}_i$  for  $i$  such that  $U_i$  contains the generic point of  $Z$ .)

## 12 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  $\mathcal{O}_k$ . 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

$$(12.1) \quad \mathcal{O}_{k,S} := \{x \in k : \|x\|_v \leq 1 \text{ for all } v \notin S\}$$

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  $\mathcal{O}_k$  (resp.  $\mathcal{O}_{k,S}$ ). (Of course,  $\mathcal{O}_k = \mathcal{O}_{k,S_\infty}$ , so only one definition is really needed.) These definitions are inherited by points on a closed subvariety  $X$  of  $\mathbb{A}_k^n$ .

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  $\mathbb{A}_k^n$  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 [Serre, 1989, § 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  $i: X \hookrightarrow \mathbb{A}_k^n$  for some  $n$  and a nonzero element  $a \in k$  such that, for all  $x \in \Sigma$ , all coordinates of  $i(x)$  lie in  $(1/a)\mathcal{O}_{k,S}$  (resp.  $a$  times all coordinates are integral over  $\mathcal{O}_{k,S}$ ).

As noted above, this definition is meaningful only if  $\Sigma$  is an infinite set.

This definition can be phrased in geometric terms using Weil functions. Indeed, we identify  $\mathbb{A}_k^n$  with the complement of the hyperplane  $x_0 = 0$  in  $\mathbb{P}_k^n$ , by identifying  $(x_1, \dots, x_n) \in \mathbb{A}_k^n$  with the point  $[1 : x_1 : \dots : x_n] \in \mathbb{P}_k^n$ . Let  $H$  denote the hyperplane  $x_0 = 0$  at infinity, and let  $\lambda_H$  be the Weil function (8.8.1):

$$(12.2) \quad \begin{aligned} \lambda_{H,v}([1 : x_1 : \cdots : x_n]) &= -\log \frac{\|1\|_v}{\max\{\|1\|_v, \|x_1\|_v, \dots, \|x_n\|_v\}} \\ &= \log \max\{\|1\|_v, \|x_1\|_v, \dots, \|x_n\|_v\}. \end{aligned}$$

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  $\mathcal{O}_{k,S}$  if and only if  $\|ax\|_w \leq 1$  for all places  $w$  of  $M_L$  lying over places in  $M_k \setminus S$ , which holds if and only if  $\|x\|_w \leq \|a\|_w^{-1}$  for all such  $w$ . Thus, by (12.2),  $\Sigma \subseteq X(\bar{k})$  is  $S$ -integral if and only if there is a closed immersion  $i: X \hookrightarrow \mathbb{A}_k^n$  for some  $n$  and an  $M_k$ -constant ( $c_v$ ) with the following property. For all  $x \in \Sigma$ ,  $\lambda_{H,w}(x) \leq c_w$  for all places  $w$  of  $M_{k(x)}$  lying over places not in  $S$ . (Here, as above, we identify  $\mathbb{A}_k^n$  with  $\mathbb{P}_k^n \setminus H$ .)

By functoriality of Weil functions (Theorem 8.8b), this leads to the following definition.

**Definition 12.3.** 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  $k$ , 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  $\lambda_D$  for  $D$  and an  $M_k$ -constant ( $c_v$ ) such that

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

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  $\mathbb{A}_k^n$ , then it holds for all such closed immersions (with varying  $n$ ).

Similarly, by Theorem 8.8d, one can use a fixed Weil function  $\lambda_D$  in Definition 12.3 (after adjusting ( $c_v$ )). One can also vary the divisor, as follows.

**Proposition 12.4.** *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 Theorems 8.8a and 8.8e (additivity and boundedness of Weil functions). Details are left to the reader.  $\square$

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

**Proposition 12.5.** *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*

$$\phi: X_1 \setminus \text{Supp } D_1 \xrightarrow{\sim} X_2 \setminus \text{Supp } D_2$$

*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  $\phi$  extends to a morphism from  $X_1$  to  $X_2$ . In that case, it follows from Theorems 8.8a and 8.8e.  $\square$

**Definition 12.6.** 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  $i: U \rightarrow X$  of  $U$  into a complete variety  $X$  over  $k$  and an effective Cartier divisor  $D$  on  $X$  such that  $i(U) = X \setminus \text{Supp } D$  and  $i(\Sigma)$  is a  $D$ -integral set on  $X$ .

**Proposition 12.7.** Let  $\phi: X_1 \rightarrow X_2$  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  $\phi$ , and that the support of  $D_1$  contains the support of  $\phi^*D_2$ . If  $\Sigma$  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 12.4, we may assume that  $D_1 - \phi^*D_2$  is effective, and then the result follows by Theorems 8.8a and 8.8e.  $\square$

If we let  $U_1 = X_1 \setminus \text{Supp } D_1$  and  $U_2 = X_2 \setminus \text{Supp } D_2$ , then the above conditions on the supports of  $D_1$  and  $D_2$  are equivalent to  $\phi(U_1) \subseteq U_2$ . Therefore Proposition 12.7 says that integral sets of points on varieties are preserved by morphisms of those varieties. This phenomenon is more obvious when using models over  $\text{Spec } \mathcal{O}_k$  to work with integral points, but this will not be explored in these notes.

We also note that Definition 12.6 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 12.3, and the integrality condition is therefore vacuous in that case.

When working with rational points, Definition 12.3 can be stated using counting functions instead:  $\Sigma \subseteq X(k)$  is integral if and only if  $N_\Sigma(D, x)$  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 12.5: Let  $\phi: X_1 \rightarrow X_2$  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

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

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

Now consider the situation where  $\phi: X' \rightarrow X$  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  $\text{Supp} D' = \phi^{-1}(\text{Supp} D)$ . Assume that  $\phi$  is étale outside of  $\text{Supp} D'$ . Then any holomorphic curve  $f: \mathbb{C} \rightarrow X \setminus \text{Supp} D$  lifts to a holomorphic curve  $f': \mathbb{C} \rightarrow X' \setminus \text{Supp} D'$  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  $\phi: X' \rightarrow X$  be a morphism of complete  $k$ -varieties, and let  $D$  and  $D'$  be as above, with  $\phi$  étale outside of  $\text{Supp} D'$ . If  $\Sigma$  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 [1989, § 4.2] or Vojta [1987, § 5.1], and implies that although the points of  $\Sigma'$  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  $\Sigma' \subseteq X'(L)$ .

### 13 Units and the Borel Lemma

Units in a number field  $k$  can be related to integral points on the affine variety  $xy = 1$  in  $\mathbb{P}_k^2$ :  $u$  is a unit if and only if there is a point  $(u, v)$  on this variety with  $u, v \in \mathcal{O}_k$ . This variety is isomorphic to  $\mathbb{P}^1$  minus two points, which we may take to be 0 and  $\infty$ . 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 13.1.** [Borel, 1897] *If  $g_1, \dots, g_n$  are entire functions such that*

$$(13.1.1) \quad e^{g_1} + \dots + e^{g_n} = 1,$$

*then some  $g_j$  is constant.*

*Proof.* We may assume that  $n \geq 2$ . The homogeneous coordinates  $[e^{g_1} : \dots : e^{g_n}]$  define a holomorphic curve  $f: \mathbb{C} \rightarrow \mathbb{P}^{n-1}(\mathbb{C})$ . The image of this map omits the  $n$  coordinate hyperplanes, and also omits the hyperplane  $x_1 + \dots + x_n = 0$  (expressed in homogeneous coordinates  $[x_1 : \dots : x_n]$ ). Therefore  $N_f(H_j, r) = 0$ , as  $H_j$  varies over these  $n + 1$  hyperplanes. This contradicts (7.8) unless the image of  $f$  is contained in a hyperplane (note that  $n$  is different in (7.8)). 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.  $\square$

In fact, by induction, it can be shown that some nontrivial subsum of the terms on the left-hand side of (13.1.1) 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 [1982], and independently to Evertse [1984].

**Theorem 13.2.** (Unit Theorem) *Let  $k$  be a number field and let  $S \supseteq S_\infty$  be a finite set of places of  $k$ . Let  $\mathcal{U}$  be a collection of  $n$ -tuples  $(u_1, \dots, u_n)$  of  $S$ -units in  $k$  that satisfy the equation*

$$(13.2.1) \quad a_1 u_1 + \dots + a_n u_n = 1,$$

where  $a_1, \dots, a_n$  are fixed nonzero elements of  $k$ . Then all but finitely many elements of  $\mathcal{U}$  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  $\mathcal{U}'$  be the set of all  $(u_1, \dots, u_n)$  for which there is no such  $J$  as above. Then  $\mathcal{U}'$  is infinite.

If we regard each  $(u_1, \dots, u_n) \in \mathcal{U}'$  as a point  $[u_1 : \dots : u_n] \in \mathbb{P}_k^{n-1}$ , then by looking directly at the formula (7.9.1) for Weil functions, we see that  $N_S(H_j, x)$  is bounded as  $x$  varies over  $\mathcal{U}'$ , 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  $x \in \mathcal{U}'$  and all  $j$ , contradicting Theorem 7.11 unless all points in  $\mathbb{P}_k^{n-1}$  corresponding to points in  $\mathcal{U}'$  lie in a finite union of proper linear subspaces.

Consider one of those linear subspaces containing infinitely many points of  $\mathcal{U}'$ . Combining the equation of some hyperplane containing that subspace with (13.2.1) 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).  $\square$

*Example 13.3.* The condition with the set  $J$  is essential because, for example, the unit equation (13.2.1) 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 13.1 and 13.2, 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  $\infty$  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 13.1 by elementary geometric arguments can be readily translated to theorems on units, by replacing the use of Theorem 13.1 with Theorem 13.2. For example, Theorem 13.4 below leads directly to Theorem 13.5.

**Theorem 13.4.** [Dufresnoy, 1944, Théorème XVI]; see also [Fujimoto, 1972] and [Green, 1972]. *Let  $f: \mathbb{C} \rightarrow \mathbb{P}^n$  be a holomorphic curve that omits  $n + m$  hyperplanes in general position,  $m \geq 1$ . Then the image of  $f$  is contained in a linear subspace of dimension  $\leq [n/m]$ , where  $[\cdot]$  denotes the greatest integer function.*

**Theorem 13.5.** [Levin, 2008, 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,  $m \geq 1$ . Then  $\Sigma$  is contained in a finite union of linear subspaces of dimension  $\leq [n/m]$ .*

## 14 Conjectures in Nevanlinna Theory and Number Theory

Since the canonical line sheaf  $\mathcal{K}$  of  $\mathbb{P}^1$  is  $\mathcal{O}(-2)$ , the main inequality of Theorem 4.7 can be stated in the form

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

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

**Definition 14.1.** 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  $z_1 = \dots = z_r = 0$  for some  $r$  ( $0 \leq r \leq n$ ). 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 14.2.* Let  $X$  be a smooth complex projective variety, let  $D$  be a normal crossings divisor on  $X$ , let  $\mathcal{K}$  be the canonical line sheaf on  $X$ , and let  $\mathcal{A}$  be an ample line sheaf on  $X$ . Then:

(a). The inequality

$$(14.2.1) \quad m_f(D, r) + T_{\mathcal{K}, f}(r) \leq_{\text{exc}} O(\log^+ T_{\mathcal{A}, f}(r)) + o(\log r)$$

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

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

$$(14.2.2) \quad m_f(D, r) + T_{\mathcal{K}, f}(r) \leq_{\text{exc}} \varepsilon T_{\mathcal{A}, f}(r) + C$$

holds for all nonconstant holomorphic curves  $f: \mathbb{C} \rightarrow X$  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  $\mathbb{C}^d \rightarrow X$  if  $d = \dim X$  and the jacobian determinant of the map is not identically zero; see [Stoll, 1970] and [Carlson and Griffiths, 1972]. The conjecture itself is attributed to Griffiths, although he seems not to have put it in print anywhere.

Conjecture 14.2 has been proved for curves (Theorem 22.2 and Corollary 28.7), 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 7.7). The second part is also known in this case [Vojta, 1997].

A consequence of Conjecture 14.2 concerns holomorphic curves in varieties of general type, or of log general type.

**Proposition 14.3.** *Assume that either part of Conjecture 14.2 is true. If  $X$  is a smooth variety of general type, then a holomorphic curve  $f: \mathbb{C} \rightarrow X$  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  $f: \mathbb{C} \rightarrow X \setminus D$  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 (4.8) and (7.8), (14.2.1) can be rephrased as

$$(14.3.1) \quad N_f(D, r) \geq_{\text{exc}} T_{\mathcal{H}(D), f}(r) - O(\log^+ T_{\mathcal{A}, f}(r)) - o(\log r).$$

In this case, since  $f$  misses  $D$ , the left-hand side is zero. By the definition of log general type, the line sheaf  $\mathcal{H}(D) := \mathcal{H} \otimes \mathcal{O}(D)$  is big. Therefore, this inequality contradicts Proposition 11.11.  $\square$

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  $f: \mathbb{C} \rightarrow X$  cannot have Zariski-dense image. See [Kawamata, 1980], [Green and Griffiths, 1980], and [Siu, 1995] for the case of abelian varieties, and [Noguchi, 1981] for the more general case of semiabelian varieties. All of these references build on work of A. Bloch [1926].

Conjecture 14.2b is also known if  $X$  is an abelian variety and  $D$  is ample [Siu and Yeung, 1996]. The theorem has been extended to semiabelian varieties again by Noguchi [1998], but only applies to holomorphic curves whose image does not meet the divisor at infinity. Again, these proofs build on work of Bloch [1926].

Conjecture 14.2 will be discussed further once its counterpart in number theory has been introduced. This, in turn, requires some definitions.

**Definition 14.4.** 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  $\phi: X' \rightarrow X$  with image containing  $P$  such that  $\phi^*D$  has strict normal crossings.

This definition is discussed more in [Vojta, 2007a, § 7].

**Definition 14.5.** [Szpiro, Ullmo, and Zhang, 1997] 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 14.2 is then the following.

*Conjecture 14.6.* 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  $\mathcal{K}$  be the canonical line sheaf on  $X$ , and let  $\mathcal{A}$  be an ample line sheaf on  $X$ . Then:

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

$$(14.6.1) \quad m_S(D, x) + h_{\mathcal{K}, k}(x) \leq O(\log^+ h_{\mathcal{A}, k}(x))$$

holds for all  $x \in \Sigma$ .

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

$$(14.6.2) \quad m_S(D, x) + h_{\mathcal{K}, k}(x) \leq \varepsilon h_{\mathcal{A}, k}(x) + C$$

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

By Remark 9.14, one can replace  $h_{\mathcal{K}, k}$  in this conjecture with a big height (after possibly adjusting  $Z$  and  $\varepsilon$  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 14.6 by Chow's lemma, resolution of singularities, and Proposition 24.2 (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 14.2.

Except for error terms, the cases in which Conjecture 14.6 is known correspond closely to those cases for which Conjecture 14.2 is known. Indeed, Conjecture 14.6b is known for curves by Roth's theorem, by a theorem of Lang [1960, Thm. 2], and by Faltings's theorem on the Mordell conjecture [Faltings, 1983, 1986], 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 7.12) proves Conjecture 14.6 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  $\mathcal{A}$  and  $\varepsilon$ ).

*Remark 14.7.* Conjecture 14.6 (and also Conjecture 14.2) 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  $p_i: X \rightarrow X_i$  ( $i = 1, 2$ ). Let  $D_1$  and  $D_2$  be normal crossings divisors on  $X_1$  and  $X_2$ , respectively, and let  $\mathcal{K}, \mathcal{K}_1$  and  $\mathcal{K}_2$  be the canonical line sheaves on  $X, X_1$ , and  $X_2$ , respectively. We have  $\mathcal{K} \cong p_1^* \mathcal{K}_1 \otimes p_2^* \mathcal{K}_2$ . 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 14.8.* One may ask whether one can make the same change to this conjecture as was done in going from Theorem 7.11 to 7.12 (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  $\pi: X' \rightarrow X$  such that the support of the divisor  $\sum \pi^* D_i$  has normal crossings. Let  $D'$  be the reduced divisor on  $X'$  for which  $\text{Supp } D' = \text{Supp } \sum \pi^* D_i$ , and let  $\mathcal{H}'$  and  $\mathcal{H}$  be the canonical line sheaves of  $X'$  and  $X$ , respectively. By Proposition 24.2, we have

$$(14.8.1) \quad \sum_{v \in S} \max\{\lambda_{D_i, v}(x) : i = 1, \dots, q\} + h_{\mathcal{H}, k}(x) \leq m_S(D', x') + h_{\mathcal{H}', k}(x') + O(1)$$

for all  $x' \in X'(k)$ , where  $x = \pi(x')$ . Therefore, if the left-hand side of (14.6.1) or (14.6.2) were replaced by the left-hand side of (14.8.1), then the resulting conjecture would be a consequence of Conjecture 14.6 applied to  $D'$  on  $X'$ .

Theorems 7.12 and 7.13 are still needed, though, because Conjectures 14.6 and 14.2 have not been proved for blowings-up of  $\mathbb{P}^n$ .

Corresponding to Proposition 14.3, we also have

**Proposition 14.9.** *Assume that either part of Conjecture 14.6 is true. Let  $k$  and  $S$  be as in Conjecture 14.6, 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, (14.6.1) is equivalent to

$$N_S(D, x) \geq h_{\mathcal{H}(D), k}(x) - O(\log^+ h_{\mathcal{H}, k}(x)),$$

and (14.6.2) 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 9.13 since  $\mathcal{H}(D)$  is big.  $\square$

This proof shows how Conjecture 14.6 is tied to the Mordell conjecture.

As was the case in Nevanlinna theory, the conclusion of Proposition 14.9 has been shown to hold for closed subvarieties of semiabelian varieties, by Faltings [1994] in the abelian case and Vojta [1996] in the semiabelian case.

In addition, Conjecture 14.6b has been proved when  $X$  is an abelian variety and  $D$  is ample [Faltings, 1991]. This has been extended to semiabelian varieties [Vojta, 1999], but in that case (14.6.2) was shown only to hold for sets of integral points on the semiabelian variety.

In parts (b) of Conjectures 14.2 and 14.6, the exceptional set  $Z$  must depend on  $\varepsilon$ ; this is because of the following theorem.

**Theorem 14.10.** [Levin, McKinnon, and Winkelmann, 2008] *There are examples of smooth projective surfaces  $X$  containing infinitely many rational curves  $Z_i$  for which the restrictions of (14.2.1) and (14.6.1) 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 [1991, Ch. I, § 3] has an extensive conjectural framework concerning how the exceptional set in part (b) of Conjectures 14.2 and 14.6 may behave, especially for varieties of general type (which would not include the examples of Theorem 14.10). See also Section 16. 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 14.2a in such a setting as well. Such a change would not be meaningful for Conjecture 14.6a, however, since the set of rational (and even algebraic) points on a variety is at most countable.

The formulation of Conjecture 14.6 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 14.11.** 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 14.6 has also been posed for algebraic points. See Conjecture 24.1.

## 15 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 [1949] and Osgood [1975] 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  $\mathcal{O}_{B,b}$  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 (1.3) hold.

Let  $L$  be a finite extension of  $k$ . Then it, too, is a function field over  $F$ , and equations (1.4) and (1.5) 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  $B' \rightarrow B$  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

$$(15.1) \quad \deg \sum n_b \cdot b = \sum n_b [\kappa(b) : F].$$

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  $F$ ,<sup>1</sup> 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 (1.6) 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 [Artin, 1967, Ch. 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 15.2.** A **model** for a variety  $X$  over  $k$  is an integral scheme  $\mathcal{X}$ , given with a flat morphism  $\pi: \mathcal{X} \rightarrow B$  of finite type and an isomorphism  $\mathcal{X} \times_B \text{Spec } k \cong X$  of schemes over  $k$ . The model is said to be **projective** (resp. **proper**) if the morphism  $\pi$  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  $\mathbb{P}_B^N$ . Likewise, a proper model for a complete variety exists, by Nagata's embedding theorem. In either case the model may be constructed

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

so that any given finite collection of Cartier divisors and line sheaves extends to the same sorts of objects on the model [Vojta, 2007b].

If  $\mathcal{X}$  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  $i: B \rightarrow \mathcal{X}$  of  $\pi: \mathcal{X} \rightarrow B$ . Indeed, if  $i$  is such a section then it takes the generic point of  $B$  to a point on the generic fiber of  $\pi$ , which is  $X$ . Conversely, given a point in  $X(k)$ , one can take its closure to get a curve in  $\mathcal{X}$ ; it is then possible to show that the restriction of  $\pi$  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  $B' \rightarrow \mathcal{X}$  over  $B$ . This follows by applying the above argument to  $\mathcal{X} \times_B B'$ , 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.

**Definition 15.3.** Let  $X$  be a complete variety over  $k$ , let  $D$  be a Cartier divisor on  $X$ , and let  $\pi: \mathcal{X} \rightarrow B$  be a proper model for  $X$ . Assume that  $D$  extends to a Cartier divisor on  $\mathcal{X}$ , also denoted by  $D$ . Let  $L$  be a finite extension of  $k$ , let  $x \in X(L)$  be a point not lying on  $\text{Supp } D$ , let  $i: B' \rightarrow \mathcal{X}$  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  $\text{Supp } D$  on  $\mathcal{X}$ , 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 8.6 (where now  $k$  is a function field). Consequently, Theorems 8.8 and 8.10 hold in this context. Moreover, this definition is compatible with (8.12) (corresponding to changing  $B'$ ).

In the case of Theorem 8.8, 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  $\mathcal{X}$  of a  $k$ -variety  $X$ , then (in the notation of Definition 15.3)  $i^*D$  is an effective divisor on  $B'$ , so  $n_b \geq 0$  for all  $b$ , hence  $\lambda_{D,w}(x) \geq 0$  for all  $w$  and all  $x \notin \text{Supp } D$ . Similarly, suppose that  $D$  is a Cartier divisor on  $X$ , and that  $\lambda_D$  and  $\lambda'_D$  are two Weil functions obtained from extensions  $D$  and  $D'$  of  $D$  to models  $\mathcal{X}$  and  $\mathcal{X}'$ , respectively, using Definition 15.3. We may reduce to the situation where the two models are the same: let  $\mathcal{X}''$  be the closure of the graph of the birational map between  $\mathcal{X}$  and  $\mathcal{X}'$ , and pull back  $D$  and  $D'$  to  $\mathcal{X}''$ . But now the difference  $D - D'$  is a divisor on  $\mathcal{X}$  which does not meet the generic fiber, so it is supported only on a finite sum of closed fibers of  $\pi: \mathcal{X} \rightarrow B$ . Therefore the Weil function associated to  $D - D'$ , and hence the difference between  $\lambda_D$  and  $\lambda'_D$ , is bounded by an  $M_k$ -constant.

Following Section 9, one can then define the height of points  $x \notin \text{Supp } D$ , starting from a model  $\mathcal{X}$  for  $X$  over  $B$  and a Cartier divisor  $D$  on  $\mathcal{X}$ : with notation as in Definition 15.3, we have

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

by (15.1).

Therefore, heights can be expressed using intersection numbers. It is this observation that led to the development of Arakelov theory, which defines models over  $\mathcal{O}_k$  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 (15.4) are defined exactly (given a model of the variety and an extension of the Cartier divisor to that model). Except for Theorem 9.10 (Northcott’s theorem), all of the results of Section 9 extend to the case of heights defined as in (15.4). In particular, if  $\mathcal{L}$  is a line sheaf on a model  $\mathcal{X}$  for  $X$ , then

$$(15.5) \qquad h_{\mathcal{L},k}(x) = \frac{\text{deg } i^* \mathcal{L}}{[L:k]} .$$

Northcott’s theorem is false over function fields (unless  $F$  is finite). Instead, however, it is true that the set (9.10.1) 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  $X$ ,  $\pi: \mathcal{X} \rightarrow B$ ,  $x \in X(k)$ , and  $i: B \rightarrow \mathcal{X}$  be as in Definition 15.3 (with  $L = k$ ), let  $D$  be an effective Cartier divisor on  $\mathcal{X}$ , and let  $\lambda_D$  be the corresponding Weil function as in Definition 15.3. 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 12.3 with  $\lambda_D$  as above and  $c_v = 0$  if and only if it corresponds to a section of the map  $\pi^{-1}(B \setminus S) \rightarrow B \setminus S$ . A similar situation holds with algebraic points, which then correspond to multisections of  $\pi^{-1}(B \setminus S) \rightarrow B \setminus S$ .

Conversely, given a set of integral points as in Definition 12.3, 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.

## 16 The Exceptional Set

The exceptional set mentioned in Conjectures 14.2b and 14.6b leads to interesting questions of its own, even when working only with rational points (or integral points) in the contexts of Propositions 14.3 or 14.9. This question has been explored in more detail by S. Lang; this is the main topic of this section. For references, see [Lang, 1986], [Lang, 1991, Ch. I § 3], and [Levin, 2009].

**Definition 16.1.** Let  $X$  be a complete variety over a field  $k$ .

- (a). The **exceptional set**  $\text{Exc}(X)$  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**  $\text{Exc}_{\text{dio}}(X)$  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**  $\text{Exc}_{\text{hol}}(X)$  is the Zariski closure of the union of the images of all nonconstant holomorphic curves  $\mathbb{C} \rightarrow X$ .

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  $\text{Exc}_{\text{dio}}(X) = \text{Exc}(X)$  if  $k$  is a finitely-generated extension of  $\mathbb{Q}$ , and that  $\text{Exc}_{\text{hol}}(X) = \text{Exc}(X)$  if  $k = \mathbb{C}$ . And, finally, if  $k$  is finitely generated over  $\mathbb{Q}$ , then he conjectured that

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

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 [Kawamata, 1980]:

**Theorem 16.2.** (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 [Noguchi, 1981] that the image of a nonconstant holomorphic curve  $\mathbb{C} \rightarrow X$  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 7.14 are the same in Theorems 7.12 and 7.13.

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  $\text{Exc}(X \setminus D)$ : 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** [Iitaka, 1982, §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 [2004]. 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  $\geq 2$ . 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  $X \rightarrow C$ . 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  $\leq B$ , as  $B$  varies. This is a very active area of number theory, but is beyond the scope of these notes.

## 17 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  $\text{Exc}(X \setminus D)$ .

- B: For each  $\varepsilon > 0$ , find the exceptional subset  $Z$  in Conjectures 14.2 and 14.6. 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  $D|_Y$ , the main inequality (14.2.2) or (14.6.2) 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  $Y \setminus S \rightarrow X \setminus D$  whose image is not contained in the exceptional set, is parametrized by a finite union of varieties.
- B: Prove Conjecture 14.6 in the split function field case of characteristic zero.
- 3A: Prove that all holomorphic curves  $\mathbb{C} \rightarrow X \setminus D$  must lie in the exceptional set.
- B: Prove Conjecture 14.2 for holomorphic curves  $\mathbb{C} \rightarrow X$ .
- 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 14.6 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 14.6 (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 28.9, 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 Sections 24 and 26.

## 18 Embeddings

A major goal of these lectures is to describe recent work on partial proofs of Conjecture 14.6 (as well as Conjecture 14.2). 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  $X = \mathbb{P}_k^2$  and  $D$  is a normal crossings divisor of degree  $\geq 4$ . 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  $\mathbb{C} \rightarrow X \setminus D$  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 Section 13.

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 [Green, 1975] and [Vojta, 1987, 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 3 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 then give an inequality of the form

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

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

$$(18.1) \quad \frac{\binom{r+n-d}{n}}{\binom{r+n}{n} \cdot r} = \frac{(r-d+n) \cdots (r-d+1)}{(r+n) \cdots (r+1)r} = \frac{r^n + O(r^{n-1})}{r(r^n + O(r^{n-1}))} \rightarrow \frac{1}{r} \rightarrow 0.$$

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

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

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

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

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

$$(18.2) \quad \frac{\sum_{k=1}^{\lfloor r/d \rfloor} \binom{r-kd+n}{n}}{r \binom{r+n}{n}} \approx \frac{\frac{r^{n+1}}{(n+1)d \cdot n!}}{\frac{r^{n+1}}{n!}} \rightarrow \frac{1}{(n+1)d}.$$

This indeed gives a nontrivial inequality

$$(18.3) \quad \frac{1}{d} m_S(D, x) + \cdots \leq (n+1 + \varepsilon) h_k(x) + O(1).$$

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 (18.3) 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  $\rho_1$  and  $\rho_2$ , respectively. We have

$$(18.4) \quad \text{codim}(L_1 \cap L_2) \leq \rho_1 + \rho_2$$

(assuming  $L_1 \cap L_2 \neq \emptyset$ ). 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  $\rho_1$  generic hyperplanes containing  $L_1$  and  $\rho_2$  generic hyperplanes containing  $L_2$ . If these  $\rho_1 + \rho_2$  hyperplanes are in general position then this implies

$$\text{codim}(L_1 \cap L_2) \geq \rho_1 + \rho_2.$$

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 (7.12.1) will not be as large as one would hope.

So, in order to apply the reasoning leading up to (18.3) independently for each irreducible component of the divisor, equality must hold in (18.4) 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 (18.1) to (18.2).

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.

## 19 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 Section 18. This was done by Corvaja and Zannier [2002], and gave a new proof of Siegel's theorem.

**Theorem 19.1.** (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  $\bar{C}$  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  $\bar{C}$  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  $h^0(\bar{C}, \mathcal{O}(ND - \ell Q_j)) \geq Nr - \ell + 1 - g$ , so we can choose  $Nr - \ell + 1 - g$  linearly independent hyperplanes in  $\mathbb{P}^{M-1}$  vanishing to order  $\geq \ell$  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

$$(19.1.1) \quad \sum_{v \in S} \sum_{\ell=0}^{Nr-g} \ell \lambda_{Q_{j(v),v}}(P_i) \leq (M + \varepsilon) h_k(P_i) + O(1)$$

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  $\lambda_{Q_{j(v),v}}(P_i) = O(1)$  for all  $j \neq j(v)$  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 (19.1.1) becomes

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

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

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

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

Of course, if  $g \geq 1$  then Siegel only required  $r > 0$ . This can be proved by reducing to the above case. Indeed, embed  $\bar{C}$  in its Jacobian and pull back by multiplication by 2. This gives an étale cover of  $\bar{C}$  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 Section 12).

## 20 The Corvaja-Zannier Method in Higher Dimensions

Corvaja and Zannier further developed their method to higher dimensions; see for example [Corvaja and Zannier, 2004]. It did not provide the full strength of Conjecture 14.6, 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 A. Levin [2009].

**Definition 20.1.** 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

$$(20.1.1) \quad \sum_{f \in B} \text{ord}_E f > 0$$

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

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

*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*

$$(20.2.1) \quad \alpha: H^0(X, \mathcal{O}(D)) \otimes_k L \rightarrow H^0(X', \mathcal{O}(D'))$$

*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 (20.1.1) for  $D$  at  $\phi(P') \in X$ . We have a commutative diagram

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

in which  $\beta$  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

$$(20.2.2) \quad \text{ord}_{E'} \beta(f \otimes 1) \geq \sum_E n_E \text{ord}_E f,$$

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 (20.2.2), note that if  $s$  is a nonzero element of  $H^0(X, \mathcal{O}(D))$ , then  $\text{ord}_{E'} \alpha(s \otimes 1) \geq \sum_E n_E \text{ord}_E s$ , with equality if  $(s) = D$ . (Strictness may arise if  $E'$  is

exceptional for  $\phi$  and  $s$  vanishes along prime divisors containing  $\phi(E')$  that do not occur in  $D$ .)

By (20.2.2), we then have

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

(since at least one  $n_E$  is strictly positive). Thus  $B'$  satisfies (20.1.1) for  $E'$ .  $\square$

**Corollary 20.3.** *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  $\phi: X_L \rightarrow X$ , 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 (20.2.1) is an isomorphism because  $L$  is flat over  $k$  [Hartshorne, 1977, III Prop. 9.3].  $\square$

**Corollary 20.4.** *Let  $\phi: X' \rightarrow X$  be a proper birational morphism<sup>2</sup> of nonsingular complete varieties over a field, and let  $D$  be a divisor on  $X$ . If  $D$  is large (resp. very large), then  $\phi^*D$  is a large (resp. very large) divisor on  $X'$ .*

*Proof.* Again, we may assume that  $D$  is very large. In this case (20.2.1) is an isomorphism because  $\phi_*\mathcal{O}_{X'} = \mathcal{O}_X$ ; see [Hartshorne, 1977, proof of III Cor. 11.4 and III Remark 8.8.1].  $\square$

*Remark 20.5.* More generally, for a linear subspace  $V$  of  $L(D)$ , one may define  $V$ -very large. For this definition, Proposition 20.2 holds with the weaker assumption that (20.2.1) 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 20.6.** (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 12.4, we may assume that  $D$  is very large.

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

We may assume that  $\Phi$  is a morphism. Indeed, let  $X'$  be a desingularization of the closure of the graph of  $\Phi$ . Replace  $X$  with  $X'$  and  $D$  with its pull-back. By Corollary 20.4 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  $X' \rightarrow X$  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

<sup>2</sup> This requires  $\phi$  to be a morphism, not just a rational map

$P_v \in X(k_v)$  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 (20.1.1) 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 20.3 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  $\text{Spec } k_v$  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

$$(20.6.1) \quad \sum_{v \in S} \sum_{j=1}^{\ell(D)} \lambda_{H_{v,j},v}(\Phi(P_i)) \geq (\ell(D) + \varepsilon)h_k(\Phi(P_i)) + O(1)$$

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  $\mu$  be the largest multiplicity of a component of  $D$ , and let  $\varepsilon = 1/\mu$ , so that  $\text{ord}_E D \leq 1$  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  $\sum \text{ord}_E \Phi^* H_{v,j} > \ell(D) \text{ord}_E D$  (by (20.1.1)), so

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

and therefore

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

since  $\lambda_{E,v}(P_i) \geq O(1)$ . 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

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

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

Summing over  $v \in S$  then gives

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

since the  $P_i$  are  $(D, S)$ -integral. This is equivalent to (20.6.1) by functoriality of heights.  $\square$

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 A. Levin [2009], works around this problem. The key idea is that it suffices to use only finitely many bases in Definition 20.1.

**Lemma 20.7.** *Let  $X$  be a nonsingular complete variety over  $\mathbb{C}$ , and let  $D$  be an effective divisor on  $X$ . Let  $\sigma_0$  be the set of prime divisors occurring in  $D$ , and let  $\Sigma$  be the set of subsets  $\sigma$  of  $\sigma_0$  for which  $\bigcap_{E \in \sigma} E$  is nonempty. For each  $\sigma \in \Sigma$  let  $D_\sigma$  be the sum of those components of  $D$  not lying in  $\sigma$ , with the same multiplicities as they have in  $D$ . Pick a Weil function for each such  $D_\sigma$ . Then there is a constant  $C$ , depending only on  $X$  and  $D$ , such that*

$$(20.7.1) \quad \min_{\sigma \in \Sigma} \lambda_{D_\sigma}(P) \leq C$$

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

*Proof.* The conditions imply that

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

since for all  $P \in X$  the set  $\sigma := \{E \in \sigma_0 : E \ni P\}$  is an element of  $\Sigma$ , and then  $P \notin \text{Supp } D_\sigma$ . The lemma then follows from Lemma 8.9, since  $\Sigma$  is a finite set.  $\square$

**Lemma 20.8.** *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  $\Phi: X \rightarrow \mathbb{P}^{\ell(D)-1}$  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  $\lambda_D$  of Weil functions on  $\mathbb{P}^{\ell(D)-1}$  and  $X$ , respectively, we have*

$$(20.8.1) \quad \max_J \sum_{j \in J} \lambda_{H_j}(\Phi(P)) \geq (\ell(D) + \varepsilon) \lambda_D(P) + O(1),$$

where the implicit constant in  $O(1)$  is independent of  $P \in X(\mathbb{C})$ . Here, as in Theorem 7.13,  $J$  varies over all subsets of  $\{1, \dots, q\}$  corresponding to subsets of  $\{H_1, \dots, H_q\}$  that lie in general position.

*Proof.* Let  $\Sigma$  be the (finite) set of Lemma 20.7, and for each  $\sigma \in \Sigma$  let  $B_\sigma$  be a basis for  $L(D)$  that satisfies (20.1.1) 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  $\lambda_{H_j}$  for them. For each  $\sigma \in \Sigma$  let  $D_\sigma$  be as in Lemma 20.7, and choose a Weil function  $\lambda_{D_\sigma}$  for it. Let  $C$  be a constant



that satisfies (20.7.1). Finally, choose Weil functions  $\lambda_E$  for each prime divisor  $E$  occurring in  $D$ .

Let  $\mu$  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

$$(20.8.2) \quad \lambda_{D_\sigma} \leq C ,$$

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  $B_\sigma$ . As before, (20.1.1) applied to  $B_\sigma$  implies that  $\sum_{j \in J} \text{ord}_E \Phi^* H_j > \ell(D) \text{ord}_E D$  for all  $E \in \sigma$ ; hence

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

and therefore

$$(20.8.3) \quad \sum_{j \in J} (\text{ord}_E \Phi^* H_j) \lambda_E(P) \geq (\ell(D) + \varepsilon) (\text{ord}_E D) \lambda_E(P) + O(1)$$

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

$$(20.8.4) \quad D = D_\sigma + \sum_{E \in \sigma} (\text{ord}_E D) \cdot E .$$

By (20.8.3), (20.8.2), and (20.8.4), we then have

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

In the above, the constants in  $O(1)$  depend only on the choices of  $B_\sigma$  and the choices of Weil functions, and on  $\sigma$  (which has only finitely many choices). Since  $J$  is one of the sets in (20.8.1), the lemma then follows.  $\square$

This then leads to the theorem in Nevanlinna theory corresponding to Theorem 20.6:

**Theorem 20.9.** [Levin, 2009] *Let  $X$  be a nonsingular complete variety over  $\mathbb{C}$ , let  $D$  be a nonzero large divisor on  $X$ , and let  $f: \mathbb{C} \rightarrow X$  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 20.6, we may assume that  $D$  is very large and base point free. Let  $\Phi: X \rightarrow \mathbb{P}_{\mathbb{C}}^{\ell(D)-1}$  be a morphism corresponding to a complete linear system of  $D$ . Let  $f: \mathbb{C} \rightarrow X$  be a holomorphic curve whose image does not meet  $D$ . Let  $H_1, \dots, H_q$  and  $\varepsilon$  be as in Lemma 20.8. 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 7.13 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.  $\square$

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 19.1 is a consequence of Theorem 20.6, we have the following.

**Proposition 20.10.** *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  $r \geq 3$  then  $D$  is large.*

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

Assume  $r \geq 3$ . 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 19.1, we have  $h^0(C, \mathcal{O}(D - \ell Q_j)) \geq Nr - \ell + 1 - g$  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  $H^0(C, \mathcal{O}(D))$  such that  $s_\ell$  vanishes to order  $\geq \ell - 1$  at  $Q_j$ . Dividing each such  $s_\ell$  by the canonical section  $1_D$  then gives a basis  $(f_1, \dots, f_{Nr+1-g})$  of  $L(D)$  such that  $\text{ord}_{Q_j} f_\ell \geq \ell - 1 - N$  for all  $\ell$ . Thus

$$\sum_{\ell=1}^{Nr+1-g} \text{ord}_{Q_j} f_\ell \geq (Nr+1-g) \left( \frac{Nr-g}{2} - N \right)$$

if  $N > (g-1)/r$ , and is strictly positive if also  $N > g/(r-2)$ .  $\square$

**Proposition 20.11.** *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  $p_i: X_1 \times_k X_2 \rightarrow X_i$  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

$$(20.11.1) \quad H^0(X_1, \mathcal{O}(D_1)) \otimes_k H^0(X_2, \mathcal{O}(D_2)) \longrightarrow H^0(X_1 \times_k X_2, \mathcal{O}(D))$$

is an isomorphism. Indeed, the projection formula [Hartshorne, 1977, II Ex. 5.1d] gives an isomorphism

$$\mathcal{O}(D_1) \otimes_k H^0(X_2, \mathcal{O}(D_2)) \xrightarrow{\sim} (p_1)_* \mathcal{O}(p_1^*D_1 + p_2^*D_2),$$

of sheaves on  $X_1$ , and taking global sections gives (20.11.1).

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 (20.1.1) with respect to the points  $p_1(P)$  and  $p_2(P)$ , respectively. By (20.11.1),  $\{p_1^*f_1 \cdot p_2^*f_2 : f_1 \in B_1, f_2 \in B_2\}$  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

$$\sum_{h \in B} \text{ord}_E h = \ell(D_2) \sum_{f \in B_1} \text{ord}_{E_1} f > 0,$$

so (20.1.1) is satisfied for  $E$ . Otherwise, we must have  $E = p_2^*E_2$  for an irreducible component  $E_2$  of  $D_2$ , and (20.1.1) is satisfied for a symmetrical reason. Thus  $D$  is very large.  $\square$

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  $\deg j^* \mathcal{O}(D) \geq 0$  for all maps  $j: C \rightarrow X$  from a curve  $C$  over  $k$  to  $X$ .

**Theorem 20.12.** [Levin, 2009, 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  $P \in \text{Supp } D$ , where  $D_P = \sum_{\{i: D_i \ni P\}} D_i$ . Then  $D$  is large.*

For the proof, see [Levin, 2009]. Note that this generalizes Proposition 20.10. As another example of this method, we note another theorem of Levin.

**Definition 20.13.** 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 20.14.** [Levin, 2009, 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 [Levin, 2009].

Again, we note that if  $X$  is a nonsingular curve, then this reduces to the combination of Theorem 20.6 and Proposition 20.10.

## 21 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 [2000], and was further developed jointly with Evertse; see for example [Evertse and Ferretti, 2008]. This work was translated into Nevanlinna theory by Ru [2009], solving a conjecture of Shiffman.

Throughout this section,  $k$  is a field of characteristic 0 and  $X \subseteq \mathbb{P}_k^N$  is a projective variety over  $k$  of dimension  $n$  and degree  $\Delta$ .

**Definition 21.1.** 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  $\Delta$  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  $\mathbb{P}_k^N$  corresponding to  $\mathbf{u}_i \in (\mathbb{P}_k^N)^*$ .

For more details on Chow forms, see Hodge and Pedoe [1952, Vol. II, Ch. X, §§ 6–8].

**Definition 21.2.** 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  $X \subseteq \mathbb{P}_k^N$ , 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 21.3.** 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  $\ell$ , 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  $a_{\ell 0} + \cdots + a_{\ell N} = m$  for all  $\ell$ .)

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

$$(21.4) \quad \frac{S_X(m, \mathbf{c})}{H_X(m)} \geq \frac{m}{(n+1)\Delta} e_X(\mathbf{c}) - (2n+1)\Delta \max_{0 \leq j \leq N} c_j$$

[Evertse and Ferretti, 2008, 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 [2008].

**Theorem 21.5.** 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  $\mathbb{P}_k^N$  satisfying

$$X \cap \bigcap_{j=0}^n \text{Supp } D_j^{(v)} = \emptyset.$$

Then for all  $\varepsilon$  with  $0 < \varepsilon \leq 1$ , there are hypersurfaces  $G_1, \dots, G_u$  in  $\mathbb{P}^N$ , not containing  $X$  and of degree

$$(21.5.1) \quad \deg G_i \leq 2(n+1)(2n+1)(n+2)\Delta d^{n+1} \varepsilon^{-1},$$

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

$$(21.5.2) \quad \sum_{v \in S} \sum_{j=0}^n \frac{\lambda_{D_j^{(v)}, v}(x)}{\deg D_j^{(v)}} \geq (n+1+\varepsilon)h_k(x) + O(1)$$

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 (21.5.2) of height be-

low a given explicit bound. They obtain a weaker bound than (21.5.1), because of this added strength.)

Note that this result is weaker than Conjecture 14.6, 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 [2008], 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 [2009], even though he is working in Nevanlinna theory.

**Theorem 21.6.** *Assume that  $n = \dim X > 0$ . Let  $D_0, \dots, D_q$  be effective divisors on  $\mathbb{P}_k^N$ , whose supports do not contain  $X$ . Let  $\mathcal{J}$  be the set of all  $(n + 1)$ -element subsets  $J$  of  $\{0, \dots, q\}$  for which*

$$(21.6.1) \quad X \cap \bigcap_{j \in J} \text{Supp } D_j = \emptyset,$$

and assume that  $\mathcal{J}$  is not empty. Then for all  $\varepsilon$  with  $0 < \varepsilon \leq 1$ , 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

$$(21.6.2) \quad \sum_{v \in S} \max_{J \in \mathcal{J}} \sum_{j \in J} \frac{\lambda_{D_j, v}(x)}{\deg D_j} \geq (n + 1 + \varepsilon)h_k(x) + C$$

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  $\phi: \mathbb{P}_k^N \rightarrow \mathbb{P}_k^M$  be the  $d$ -uple embedding, where  $M = \binom{N+d}{N} - 1$ , and let  $Y = \phi(X)$ . Then  $Y$  has dimension  $n$ , degree  $\Delta d^n$ , and  $\phi$  multiplies the projective height by  $d$ . Moreover, there are hyperplanes  $E_0, \dots, E_q$  on  $\mathbb{P}_k^M$  such that  $\phi^* E_j = D_j$  for all  $j$ . Thus if  $y = \phi(x)$  then (21.6.2) is equivalent to

$$(21.6.3) \quad \sum_{v \in S} \max_J \sum_{j \in J} \lambda_{E_j, v}(y) \geq (n + 1 + \varepsilon)h_k(y) + C'$$

for a suitable constant  $C'$  independent of  $x$ . Applying Theorem 21.6 with  $d = 1$  to  $Y$  and  $E_0, \dots, E_q$  then gives hypersurfaces  $G'_1, \dots, G'_u$  in  $\mathbb{P}_k^M$ , 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 (21.6.3). Pulling these hypersurfaces back to  $\mathbb{P}_k^N$  multiplies their degrees by  $d$ , so these pull-backs satisfy (21.6.3).

By a further linear embedding of  $\mathbb{P}_k^N$ , 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 (21.6.2) are nonnegative.

Now assume, by way of contradiction, that the set of solutions of (21.6.2) is not contained in a finite union of hypersurfaces  $G_i$  satisfying (21.5.1). By a partitioning argument [Evertse and Ferretti, 2008, Lemma 5.3] there is a subset  $\Sigma$  of  $X(k)$ , not contained in a finite union of hypersurfaces  $G_i$  as before,  $(n + 1)$ -element subsets  $J_\nu$  of  $\{0, \dots, q\}$  satisfying (21.6.1) for each  $\nu \in S$ , and nonnegative real constants  $c_{j,\nu}$  for all  $\nu \in S$  and all  $j \in J_\nu$ , such that

$$(21.6.4) \quad \sum_{\nu \in S} \sum_{j \in J_\nu} c_{j,\nu} = 1$$

and such that the inequality

$$(21.6.5) \quad \lambda_{D_{j,\nu}}(x) \geq c_{j,\nu} \left( n + 1 + \frac{\epsilon}{2} \right) h_k(x)$$

holds for all  $\nu \in S$ , all  $j \in J_\nu$ , and all  $x \in \Sigma$ . Also let  $c_{j,\nu} = 0$  if  $\nu \in S$  and  $j \notin J_\nu$ , so that (21.6.5) holds for all  $\nu \in S$  and all  $j = 0, \dots, q$ .

Now for some  $m > \Delta$  (its exact value is given by (21.6.15)), let

$$\phi_m : \mathbb{P}_k^N \rightarrow \mathbb{P}_k^{R_m}$$

be the  $m$ -uple embedding; here  $R_m = \binom{N+m}{m} - 1$ . Let  $X_m$  be the linear subspace of  $\mathbb{P}_k^{R_m}$  spanned by  $\phi_m(X)$ . We have

$$\dim X_m = H_X(m) - 1 .$$

For each  $\nu \in S$  let

$$\mathbf{c}_\nu = (c_{0,\nu}, \dots, c_{q,\nu}, 0, \dots, 0) \in \mathbb{R}_{\geq 0}^{N+1} ,$$

and let  $\mathbf{a}_{1,\nu}, \dots, \mathbf{a}_{H_X(m),\nu}$  be the elements of  $\mathbb{N}^{N+1}$  for which the monomials  $\mathbf{x}^{\mathbf{a}_{\ell,\nu}}$ ,  $\ell = 1, \dots, H_X(m)$ , give a basis for  $k[x_0, \dots, x_N]_m / I_m$  satisfying

$$(21.6.6) \quad S_X(m, \mathbf{c}_\nu) = \sum_{\ell=1}^{H_X(m)} \mathbf{a}_{\ell,\nu} \cdot \mathbf{c}_\nu .$$

For each  $\nu$  the monomials  $\mathbf{x}^{\mathbf{a}_{\ell,\nu}}$ ,  $\ell = 1, \dots, H_X(m)$ , define linear forms  $L_{\ell,\nu}$  in the homogeneous coordinates on  $\mathbb{P}_k^{R_m}$  which are linearly independent on  $\phi_m(X)$ , and therefore on  $X_m$ . For each  $\nu$  and  $\ell$  choose Weil functions  $\lambda_{L_{\ell,\nu}}$  on  $\mathbb{P}_k^{R_m}$ . We have

$$\lambda_{L_{\ell,\nu},\nu}(\phi_m(x)) \geq \sum_{j=0}^q a_{\ell,\nu,j} \lambda_{D_{j,\nu}}(x) + O(1)$$

for all  $\nu$  and  $\ell$ . After adjusting the Weil function, we may assume that the  $O(1)$  term is not necessary.

By (21.6.5) and (21.6.6), we then have

$$(21.6.7) \quad \sum_{\ell=1}^{H_X(m)} \lambda_{L_{\ell,v},v}(\phi_m(x)) \geq S_X(m, \mathbf{c}_v) \left( n+1 + \frac{\varepsilon}{2} \right) h_k(x)$$

for all  $x \in \Sigma$  and all  $v \in S$ . Assume for now that there is an integer  $m \geq \Delta$  such that

$$(21.6.8) \quad m \leq \frac{2(n+1)(2n+1)(n+2)\Delta}{\varepsilon}$$

and such that

$$(21.6.9) \quad \left( n+1 + \frac{\varepsilon}{2} \right) \sum_{v \in S} S_X(m, \mathbf{c}_v) > m H_X(m).$$

Then, for sufficiently small  $\varepsilon' > 0$ , (21.6.7) will imply

$$(21.6.10) \quad \sum_{v \in S} \sum_{\ell=1}^{H_X(m)} \lambda_{L_{\ell,v},v}(\phi_m(x)) \geq (H_X(m) + \varepsilon') m h_k(x) + O(1)$$

for all  $x \in \Sigma$ . Note that  $h_k(\phi_m(x)) = m h_k(x) + O(1)$ . 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  $\mathbb{P}_k^{H_X(m)}$ , containing  $\phi_m(\Sigma)$ . These pull back to give homogeneous polynomials  $G_i$  of degree  $m$  on  $\mathbb{P}_k^N$ ; they satisfy (21.5.1) by (21.6.8).

We now show that (21.6.9) holds for some  $m$  satisfying (21.6.8).

By (21.4), we have

$$(21.6.11) \quad \sum_{v \in S} S_X(m, \mathbf{c}_v) \geq H_X(m) \sum_{v \in S} \left( \frac{m}{(n+1)\Delta} e_X(\mathbf{c}_v) - (2n+1)\Delta \max_{j \in J_v} c_{j,v} \right).$$

By [Evertse and Ferretti, 2008, Lemma 5.1], we have

$$(21.6.12) \quad e_X(\mathbf{c}_v) \geq \Delta \sum_{j \in J_v} c_{j,v},$$

and therefore

$$(21.6.13) \quad \sum_{v \in S} e_X(\mathbf{c}_v) \geq \Delta \sum_{v \in S} \sum_{j \in J_v} c_{j,v} = \Delta$$

by (21.6.4). Thus (21.6.11) becomes

$$(21.6.14) \quad \begin{aligned} \sum_{v \in S} S_X(m, \mathbf{c}_v) &\geq H_X(m) \left( \frac{m}{n+1} - (2n+1)\Delta \sum_{v \in S} \max_{j \in J_v} c_{j,v} \right) \\ &\geq H_X(m) \left( \frac{m}{n+1} - (2n+1)\Delta \right). \end{aligned}$$

Now let



$$(21.6.15) \quad m = \left\lfloor \frac{2(n+1)(2n+1)(n+2)\Delta}{\varepsilon} \right\rfloor.$$

This clearly satisfies (21.6.8); in addition, we have

$$(21.6.16) \quad m < \left( n + 1 + \frac{\varepsilon}{2} \right) \left( \frac{m}{n+1} - (2n+1)\Delta \right).$$

Thus (21.6.14) becomes

$$\left( n + 1 + \frac{\varepsilon}{2} \right) \sum_{v \in S} S_X(m, \mathbf{c}_v) > mH_X(m),$$

which is (21.6.9). □

In Nevanlinna theory, the counterpart to Theorem 21.6 was proved by Ru [2009]. Here we give a slightly stronger version of his theorem:  $X$  is not required to be nonsingular, and we incorporate the set  $\mathcal{J}$ . This stronger version still follows from his proof without essential changes, though.

**Theorem 21.7.** *Let  $k = \mathbb{C}$  and assume that  $n = \dim X > 0$ . Let  $D_0, \dots, D_q$  be effective divisors on  $\mathbb{P}_{\mathbb{C}}^N$ , whose supports do not contain  $X$ . Let  $\mathcal{J}$  be the set of all  $(n+1)$ -element subsets  $J$  of  $\{0, \dots, q\}$  for which*

$$X \cap \bigcap_{j \in J} \text{Supp } D_j = \emptyset,$$

*and assume that  $\mathcal{J}$  is not empty. Fix  $\varepsilon \in \mathbb{R}$  with  $0 < \varepsilon \leq 1$ , fix  $C \in \mathbb{R}$ , and choose Weil functions  $\lambda_{D_j}$  for all  $j$ . Let  $f: \mathbb{C} \rightarrow X(\mathbb{C})$  be a holomorphic curve whose image is not contained in any hypersurface in  $\mathbb{P}^N$  not containing  $X$  of degree  $\leq 2(n+1)(2n+1)(n+2)\Delta d^{n+1} \varepsilon^{-1}$ . Then*

$$(21.7.1) \quad \int_0^{2\pi} \max_{J \in \mathcal{J}} \sum_{j \in J} \frac{\lambda_{D_j}(f(re^{i\theta}))}{\deg D_j} \frac{d\theta}{2\pi} \leq_{\text{exc}} (n+1+\varepsilon)T_f(r) + C.$$

*Proof (sketch).* This proof uses the same general outline as the proof of Theorem 21.6, 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 (21.6.4); then (21.6.13) is no longer valid. However, (21.6.12) 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  $x_0 = 0, \dots, x_q = 0$ , and that the Weil functions  $\lambda_{D_j}$  are nonnegative.

For each  $r > 0$  and  $\theta \in [0, 2\pi]$  let  $J_{r,\theta}$  be an element of  $\mathcal{J}$  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

$$\mathbf{c}_{r,\theta} = (c_{0,r,\theta}, \dots, c_{N,r,\theta}) \in \mathbb{R}_{\geq 0}^{N+1}.$$

Then, as before, [Evertse and Ferretti, 2008, Lemma 5.1] gives

$$(21.7.2) \quad e_X(\mathbf{c}_{r,\theta}) \geq \Delta \sum_{j \in J_{r,\theta}} c_{j,r,\theta}.$$

Let  $m$  be as in (21.6.15). By (21.4), (21.7.2), and nonnegativity of  $c_{j,r,\theta}$ , we have

$$(21.7.3) \quad \begin{aligned} \frac{1}{H_X(m)} \int_0^{2\pi} S_X(m, \mathbf{c}_{r,\theta}) \frac{d\theta}{2\pi} &\geq \int_0^{2\pi} \left( \frac{m}{(n+1)\Delta} e_X(\mathbf{c}_{r,\theta}) - (2n+1)\Delta \max_{j \in J_{r,\theta}} c_{j,r,\theta} \right) \frac{d\theta}{2\pi} \\ &\geq \int_0^{2\pi} \left( \frac{m}{n+1} \sum_{j \in J_{r,\theta}} c_{j,r,\theta} - (2n+1)\Delta \max_{j \in J_{r,\theta}} c_{j,r,\theta} \right) \frac{d\theta}{2\pi} \\ &\geq \left( \frac{m}{n+1} - (2n+1)\Delta \right) \int_0^{2\pi} \left( \sum_{j \in J_{r,\theta}} c_{j,r,\theta} \right) \frac{d\theta}{2\pi}. \end{aligned}$$

By (21.6.16) there is an  $\varepsilon' > 0$  such that

$$\frac{m}{n+1} - (2n+1)\Delta \geq \frac{m}{n+1 + \varepsilon/2} \cdot \frac{H_X(m) + \varepsilon'}{H_X(m)};$$

hence (21.7.3) becomes

$$(21.7.4) \quad \left( n+1 + \frac{\varepsilon}{2} \right) \int_0^{2\pi} S_X(m, \mathbf{c}_{r,\theta}) \frac{d\theta}{2\pi} \geq m(H_X(m) + \varepsilon') \int_0^{2\pi} \left( \sum_{j \in J_{r,\theta}} c_{j,r,\theta} \right) \frac{d\theta}{2\pi}.$$

This corresponds to (21.6.9) in the earlier proof.

Now let  $\phi_m: \mathbb{P}_{\mathbb{C}}^N \rightarrow \mathbb{P}_{\mathbb{C}}^{R_m}$  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  $\mathbb{P}_{\mathbb{C}}^{R_m}$  spanned by  $\phi_m(X)$ .

As before, for each  $r$  and  $\theta$  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

$$(21.7.5) \quad S_X(m, \mathbf{c}_{r,\theta}) = \sum_{\ell=1}^{H_X(m)} \mathbf{a}_{\ell,r,\theta} \cdot \mathbf{c}_{r,\theta}.$$

These correspond to linear forms  $L_{\ell,r,\theta}$  on  $\mathbb{P}_{\mathbb{C}}^{R_m}$ ,  $\ell = 1, \dots, H_X(m)$ , which are linearly independent on  $X_m$  for each  $r$  and  $\theta$ , and which satisfy

$$(21.7.6) \quad \lambda_{L_{\ell,r,\theta}}(\phi_m(f(re^{i\theta}))) \geq \sum_{j=0}^q a_{\ell,r,\theta,j} \lambda_{D_j}(f(re^{i\theta}))$$

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 (21.7.4), by (21.7.5), by (21.7.6), and by applying Cartan's Theorem 7.13 to  $X_m$ , we then have

$$\begin{aligned} & \int_0^{2\pi} \max_{J \in \mathcal{J}} \sum_{j \in J} \lambda_{D_j}(f(re^{i\theta})) \frac{d\theta}{2\pi} \\ &= \int_0^{2\pi} \sum_{j \in J_{r,\theta}} \lambda_{D_j}(f(re^{i\theta})) \frac{d\theta}{2\pi} \\ &= \int_0^{2\pi} \left( \sum_{j \in J_{r,\theta}} c_{j,r,\theta} \right) \frac{d\theta}{2\pi} \\ &\leq \frac{n+1+\varepsilon/2}{m(H_X(m)+\varepsilon')} \int_0^{2\pi} S_X(m, \mathbf{c}_{r,\theta}) \frac{d\theta}{2\pi} \\ &= \frac{n+1+\varepsilon/2}{m(H_X(m)+\varepsilon')} \int_0^{2\pi} \left( \sum_{\ell=1}^{H_X(m)} \mathbf{a}_{\ell,r,\theta} \cdot \mathbf{c}_{r,\theta} \right) \frac{d\theta}{2\pi} \\ &\leq \frac{n+1+\varepsilon/2}{m(H_X(m)+\varepsilon')} \int_0^{2\pi} \left( \sum_{\ell=1}^{H_X(m)} \lambda_{L_{\ell,r,\theta}}(\phi_m(f(re^{i\theta}))) \right) \frac{d\theta}{2\pi} \\ &\leq_{\text{exc}} \frac{n+1+\varepsilon/2}{m} T_{\phi_m \circ f}(r) + C' \\ &\leq \left( n+1 + \frac{\varepsilon}{2} \right) T_f(r) + C, \end{aligned}$$

where  $C'$  is chosen so that the last inequality holds. □

(This is better than (21.7.1) by  $\varepsilon/2$ , since we have removed the partitioning argument. Thus, the bound (21.5.1) 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.)

## 22 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 22.1.** Let  $X$  be a complete complex variety, let  $D$  be an effective Cartier divisor on  $X$ , let  $f: \mathbb{C} \rightarrow X$  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

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

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

$$N_f^{(n)}(D_1 + D_2, r) \leq N_f^{(n)}(D_1, r) + N_f^{(n)}(D_2, r) .$$

In Nevanlinna theory, the Second Main Theorem for curves has been extended to truncated counting functions:

**Theorem 22.2.** *Let  $X$  be a smooth complex projective curve, let  $D$  be a reduced effective divisor on  $X$ , let  $\mathcal{K}$  be the canonical line sheaf on  $X$ , and let  $\mathcal{A}$  be an ample line sheaf on  $X$ . Then the inequality*

$$N_f^{(1)}(D, r) \geq_{\text{exc}} T_{\mathcal{K}(D), f}(r) - O(\log^+ T_{\mathcal{A}, f}(r)) - o(\log r)$$

holds for all nonconstant holomorphic curves  $f: \mathbb{C} \rightarrow X$ .

Of course,  $X$  needs to be a curve of genus  $\leq 1$  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 26.7 and the discussion following it.

Also, in Theorem 7.7, the counting functions in (7.8) can be replaced by  $n$ -truncated counting functions:

$$(22.3) \quad \sum_{j=1}^q N_f^{(n)}(H_j, r) \geq_{\text{exc}} (q - n - 1)T_f(r) - O(\log^+ T_f(r)) - o(\log r)$$

Theorem 7.13 is not suitable for using truncated counting functions, though, due to its emphasis on the proximity function.

Conjecture 14.2, though, should also be true with counting functions. The question arises, however: truncation to what? Note that (22.3) 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 22.8). I do believe that Conjecture 14.2 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 22.2 and (22.3) are only conjectural).

**Definition 22.4.** 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  $\mathcal{O}_k$ . Then the  **$n$ -truncated counting function** with respect to  $D$  is

$$N_S^{(n)}(D, x) = \sum_{v \notin S} \min\{\lambda_{D,v}(x), n \log(\mathcal{O}_k : \mathfrak{p}_v)\}$$

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

$$(22.4.1) \quad N_S^{(n)}(D, x) = \frac{1}{[L : k]} \sum_{w \notin T} \min\{\lambda_{D,v}(x), n \log(\mathcal{O}_L : \mathfrak{p}_w)\}.$$

Truncation does not respect (8.11.1) at ramified places, so (22.4.1) 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

$$N_S^{(n)}(D_1 + D_2, x) \leq N_S^{(n)}(D_1, x) + N_S^{(n)}(D_2, x)$$

if  $D_1$  and  $D_2$  are effective.

We conjecture that a counterpart to Conjecture 14.6 holds with truncated counting functions:

*Conjecture 22.5.* 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  $\mathcal{K}$  be the canonical line sheaf on  $X$ , and let  $\mathcal{A}$  be an ample line sheaf on  $X$ . Then:

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

$$(22.5.1) \quad N_S^{(1)}(D, x) \geq h_{\mathcal{K}(D), k}(x) - O\left(\sqrt{h_{\mathcal{A}, k}(x)}\right)$$

holds for all  $x \in \Sigma$ .

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

$$(22.5.2) \quad N_S^{(1)}(D, x) \geq h_{\mathcal{K}(D), k}(x) - \varepsilon h_{\mathcal{A}, k}(x) - C$$

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

Note that the error term in (22.5.1) is weaker than in (14.6.1); see Stewart and Tijdeman [1986] and van Frankenhuijsen [2007].

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  $X = \mathbb{P}_k^1$  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 [Noguchi, 1996, (9.5)]. This conjecture can be stated (over  $\mathbb{Q}$  for simplicity, and with a weaker error term) as follows.

*Conjecture 22.6.* Fix  $\varepsilon > 0$ . Then there is a constant  $C_\varepsilon$  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

$$(22.6.1) \quad \log \max\{|a|, |b|, |c|\} \leq (1 + \varepsilon) \sum_{p|abc} \log p + C_\varepsilon .$$

To see the equivalence with the above-mentioned special case of Conjecture 22.5, let  $(a, b, c)$  be a triple of relatively prime rational integers satisfying  $a + b + c = 0$ , and let  $x \in \mathbb{P}_{\mathbb{Q}}^2$  be the corresponding point with homogeneous coordinates  $[a : b : c]$ . Then the left-hand side of (22.6.1) 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  $\mathbb{P}_{\mathbb{Q}}^2$  (defined respectively by  $x_0 = 0$ ,  $x_1 = 0$ , and  $x_2 = 0$ ). Since  $\gcd(a, b, c) = 1$ , we have

$$\lambda_{H_0, p}(x) = -\log \frac{\|a\|_p}{\max\{\|a\|_p, \|b\|_p, \|c\|_p\}} = \text{ord}_p(a) \log p ,$$

for all (finite) rational primes  $p$ , where  $\text{ord}_p(a)$  denotes the largest integer  $m$  for which  $p^m \mid a$ . Thus

$$N_{\{\infty\}}^{(1)}(H_0, x) = \sum_{p|a} \log p .$$

Similarly

$$N_{\{\infty\}}^{(1)}(H_1, x) = \sum_{p|b} \log p \quad \text{and} \quad N_{\{\infty\}}^{(1)}(H_2, x) = \sum_{p|c} \log p .$$

Therefore, by relative primeness,

$$N_{\{\infty\}}^{(1)}(D, x) = \sum_{p|abc} \log p ,$$

so (22.6.1) is equivalent to

$$h_{\mathbb{Q}}(x) \leq (1 + \varepsilon) N_{\{\infty\}}^{(1)}(D, x) + C_\varepsilon .$$

Since  $a + b + c = 0$ , the points  $[a : b : c]$  all lie on the line  $x_0 + x_1 + x_2 = 0$  in  $\mathbb{P}_{\mathbb{Q}}^2$ . Choosing an isomorphism of this line with  $\mathbb{P}_{\mathbb{Q}}^1$  such that the restriction of  $D$  corresponds to the divisor  $[0] + [1] + [\infty]$  on  $\mathbb{P}_{\mathbb{Q}}^1$ , it follows by functoriality of  $h_{\mathbb{Q}}$  and

$N_{\{\infty\}}^{(1)}(D, x)$  that Conjecture 22.6 is equivalent to the special case of Conjecture 22.5 with  $k = \mathbb{Q}$ ,  $S = \{\infty\}$ ,  $X = \mathbb{P}_{\mathbb{Q}}^1$ ,  $D = [0] + [1] + [\infty]$ , 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  $x := [a_0 : \dots : a_n] \in \mathbb{P}_{\mathbb{Q}}^n$  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

$$N_{\{\infty\}}^{(1)}(D, x) = \sum_{p \mid a_0 \dots a_{n+1}} p + O(1).$$

Since the canonical line sheaf  $\mathcal{H}$  on  $\mathbb{P}^n$  is  $\mathcal{O}(-n - 1)$  and  $D$  has degree  $n + 2$ , we have  $\mathcal{H}(D) \cong \mathcal{O}(1)$  and therefore (22.5.2) would (if true) give

$$(22.7) \quad \sum_{p \mid a_0 \dots a_{n+1}} \log p \geq (1 - \varepsilon) \log \max\{|a_0|, \dots, |a_{n+1}|\} - C$$

for all  $\varepsilon > 0$  and all  $C$ , for almost all rational points  $x = [a_0 : \dots : a_n]$  outside of some proper Zariski-closed subset of  $\mathbb{P}_{\mathbb{Q}}^n$  depending on  $\varepsilon$ .

The following example, due to Brownawell and Masser [1986, 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 22.8.* Let  $n \in \mathbb{Z}_{>0}$ , and consider the map  $\phi : \mathbb{P}_{\mathbb{Q}}^1 \rightarrow \mathbb{P}_{\mathbb{Q}}^n$  given by

$$\phi([x_0 : x_1]) = \left[ x_0^n : \binom{n}{1} x_0^{n-1} x_1 : \binom{n}{2} x_0^{n-2} x_1^2 : \dots : x_1^n \right].$$

Let  $D$  be the divisor  $(x_0 x_1 (x_0 + x_1)) = [0] + [-1] + [\infty]$  on  $\mathbb{P}_{\mathbb{Q}}^1$  and let  $D'$  be the (similar) divisor  $(y_0 \dots y_n (y_0 + \dots + y_n))$  on  $\mathbb{P}_{\mathbb{Q}}^n$ . Note that  $\text{Supp } \phi^* D' = \text{Supp } D$ , so that

$$N_{\{\infty\}}^{(1)}(D', \phi([x_0 : x_1])) = N_{\{\infty\}}^{(1)}(D, [x_0 : x_1])$$

if  $x_0 \neq 0$  and  $x_1 \neq 0$ , 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

$$N_{\{\infty\}}^{(1)}(D, [a : b]) \leq h_{\mathbb{Q}}([a : b]) ;$$

therefore we have

$$N_{\{\infty\}}^{(1)}(D', \phi([a : b])) \leq \frac{1}{n} h_{\mathbb{Q}}(\phi([a : b])) + O(1) ,$$

contrary to (22.7). The points  $\phi([a : b])$  are not contained in any hyperplane in  $\mathbb{P}_{\mathbb{Q}}^n$ . They are, of course, contained in the image of  $\phi$ , 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 14.6 that all have the property of implying the abc conjecture.

## 23 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 Section 24.

**Definition 23.1.** 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  $\mathcal{D}_{k/\mathbb{Q}}$  of  $k$  over  $\mathbb{Q}$  by the formulas

$$|D_k| = (\mathbb{Z} : N_{\mathbb{Q}}^k \mathcal{D}_{k/\mathbb{Q}}) = (\mathcal{O}_k : \mathcal{D}_{k/\mathbb{Q}}) .$$

By multiplicativity of the different in towers, we therefore have

$$\begin{aligned} d_k(L) &= \frac{1}{[L : k]} \log(\mathcal{O}_L : \mathcal{D}_{L/k}) \\ (23.2) \quad &= \frac{1}{[L : k]} \sum_{\substack{\mathfrak{q} \in \text{Spec } \mathcal{O}_L \\ \mathfrak{q} \neq (0)}} \text{ord}_{\mathfrak{q}} \mathcal{D}_{L/k} \cdot \log(\mathcal{O}_L : \mathfrak{q}) \end{aligned}$$

for number fields  $L \supseteq k$ .

This expression can be used to define a “localized” log discriminant term:

**Definition 23.3.** Let  $L \supseteq k$  be number fields, and let  $S \supseteq S_{\infty}$  be a finite set of places of  $k$ . Let  $\mathcal{O}_{L,S}$  denote the localization of  $\mathcal{O}_L$  away from (finite) places of  $L$  lying over



places in  $S$ ; note that  $\mathcal{O}_{L,S} = \mathcal{O}_L \otimes_{\mathcal{O}_k} \mathcal{O}_{k,S}$  (cf. (12.1)). Then we define

$$(23.3.1) \quad d_S(L) = \frac{1}{[L:k]} \sum_{\substack{\mathfrak{q} \in \text{Spec } \mathcal{O}_{L,S} \\ \mathfrak{q} \neq (0)}} \text{ord}_{\mathfrak{q}} \mathcal{D}_{L/k} \cdot \log(\mathcal{O}_{L,S} : \mathfrak{q}) .$$

Also, if  $X$  is a variety over  $k$  and  $x \in X(\bar{k})$ , then let

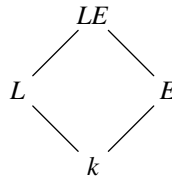
$$d_S(x) = d_S(\kappa(x)) .$$

By (23.2), 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  $(1 + \log_p [k:\mathbb{Q}]) \log p$  [Serre, 1968, Ch. 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 23.4.** *Let*



*be a diagram of number fields. Then*

$$d_S(LE) - d_S(E) \leq d_S(L) .$$

*Proof.* We first show that

$$(23.4.1) \quad \mathcal{D}_{LE/E} \mid \mathcal{D}_{L/k} \cdot \mathcal{O}_{LE} ,$$

which really amounts to showing that  $\mathcal{D}_{L/k} \subseteq \mathcal{D}_{LE/E}$ . Recall that  $\mathcal{D}_{L/k}$  is the ideal in  $\mathcal{O}_L$  generated by all elements  $f'(\alpha)$ , as  $\alpha$  varies over the set  $\{\alpha \in \mathcal{O}_L : L = k(\alpha)\}$  and  $f$  is the (monic) irreducible polynomial for  $\alpha$  over  $k$ . Such  $\alpha$  also lie in  $\mathcal{O}_{LE}$ , and generate  $LE$  over  $E$ . Let  $g$  be the irreducible polynomial for  $\alpha$  over  $E$ ; we note that  $g \mid f$  and therefore  $f = gh$  for a monic polynomial  $h \in \mathcal{O}_E[t]$ . We also have

$f'(\alpha) = g'(\alpha)h(\alpha)$ . Since  $h(\alpha) \in \mathcal{O}_{LE}$ , it then follows that  $f'(\alpha) \in \mathcal{D}_{LE/E}$ , which then implies (23.4.1).

It then suffices to show that

$$\text{ord}_{\mathfrak{q}} \mathcal{D}_{L/k} \cdot \log(\mathcal{O}_L : \mathfrak{q}) = \frac{1}{[LE : L]} \sum_{\substack{Q \in \text{Spec } \mathcal{O}_{LE} \\ Q|\mathfrak{q}}} \text{ord}_Q(\mathcal{D}_{L/k} \mathcal{O}_{LE}) \cdot \log(\mathcal{O}_{LE} : Q)$$

for all nonzero  $\mathfrak{q} \in \text{Spec } \mathcal{O}_L$ . But this follows immediately from the classical fact that  $[LE : L] = \sum e_{Q/\mathfrak{q}} f_{Q/\mathfrak{q}}$ .  $\square$

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 23.5.** Let  $\phi : X \rightarrow Y$  be a generically finite, dominant morphism of non-singular complete varieties over a field  $k$ . Assume that the function field extension  $K(X)/K(Y)$  is separable. Then the natural map  $\phi^* \Omega_{Y/k} \rightarrow \Omega_{X/k}$  induces a natural map  $\phi^* \mathcal{H}_Y \rightarrow \mathcal{H}_X$  of canonical sheaves, which in turn defines a natural map  $\mathcal{O}_X \rightarrow \mathcal{H}_X \otimes \phi^* \mathcal{H}_Y^\vee$ . This latter map defines a section of  $\mathcal{H}_X \otimes \phi^* \mathcal{H}_Y^\vee$ , whose divisor is the **ramification divisor** of  $X$  over  $Y$ . It is an effective divisor  $R$ , and we have  $\mathcal{H}_X \cong \mathcal{H}_Y(R)$ .

(The remainder of this section will be quite technical. Most readers will likely be interested only in the statement of Theorem 23.15, and should now skip to the statement of that theorem and to Theorem 23.17.)

The following definition will also be needed to generalize Chevalley-Weil.

**Definition 23.6.** Let  $M$  be a finitely generated module over a noetherian ring  $R$ , and let

$$(23.6.1) \quad R^m \xrightarrow{f} R^n \longrightarrow M \longrightarrow 0$$

be a presentation of  $M$ . Then the  $0^{\text{th}}$  **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 [Eisenbud, 1995, 20.4]. This globalizes: if  $\mathcal{F}$  is a coherent sheaf on a noetherian scheme  $X$ , then locally one can form presentations (23.6.1) and glue them to give a sheaf of ideals  $F_0(\mathcal{F})$ .

**Lemma 23.7.** Let  $\mathcal{F}$  be a coherent sheaf on a noetherian scheme  $X$ .

(a). *Forming the Fitting ideal commutes with pull-back: Let  $\phi : X' \rightarrow X$  be a morphism of noetherian schemes. Then*

$$(23.7.1) \quad F_0(\phi^* \mathcal{F}) = \phi^* F_0(\mathcal{F}) \cdot \mathcal{O}_{X'}.$$

(b). *If  $\mathcal{F} \twoheadrightarrow \mathcal{F}'$  is a surjection, then*

$$F_0(\mathcal{F}') \supseteq F_0(\mathcal{F}).$$

*Proof.* Let  $\phi: X' \rightarrow X$  be as in (a). Since tensoring is right exact, a presentation of  $\mathcal{F}$  pulls back to give a presentation of  $\phi^*\mathcal{F}$  on  $X'$ , and (23.7.1) follows directly.

If  $\mathcal{F} \twoheadrightarrow \mathcal{F}'$  is a surjection, then one can use the same middle terms in the local presentations of  $\mathcal{F}$  and of  $\mathcal{F}'$ , and the first term in the presentation of  $\mathcal{F}$  can be a direct summand of the first term in the presentation of  $\mathcal{F}'$ . In this case, the resulting generators of  $F_0(\mathcal{F})$  are a subset of the generators of  $F_0(\mathcal{F}')$ . This gives (b).  $\square$

The ramification divisor can be described using a Fitting ideal.

**Lemma 23.8.** *Let  $\phi: X \rightarrow Y$  be as in Definition 23.5. 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.  $\square$

The proof of Theorem 23.15 will also need the notion of a *model* of a variety over a number field (corresponding to Definition 15.2 in the function field case).

**Definition 23.9.** Let  $X$  be a variety over a number field  $k$ . A **model** for  $X$  over  $\mathcal{O}_k$  is an integral scheme  $\mathcal{X}$ , flat over  $\mathcal{O}_k$ , together with an isomorphism  $X \cong \mathcal{X} \times_{\mathcal{O}_k} k$ .

If  $X$  is a complete variety, then a model  $\mathcal{X}$  can be constructed using Nagata's embedding theorem. Moreover,  $\mathcal{X}$  can be chosen to be proper over  $\text{Spec } \mathcal{O}_k$ . On such a model, rational points correspond naturally and bijectively to sections  $\mathcal{O}_k \rightarrow \mathcal{X}$  by the valuative criterion of properness, and algebraic points  $\text{Spec } L \rightarrow X$  correspond naturally and bijectively to morphisms  $\text{Spec } \mathcal{O}_L \rightarrow \mathcal{X}$  over  $\mathcal{O}_k$ .

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 23.10.** Let  $X$  be a complete variety over a number field  $k$ , let  $\mathcal{X}$  be a proper model for  $X$  over  $\mathcal{O}_k$ , let  $\mathcal{I}$  be a sheaf of ideals on  $\mathcal{X}$ , 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  $\mathcal{X}$  defined by  $\mathcal{I}$ . Then the **counting function**  $N_S(\mathcal{I}, x)$  is defined as follows. Let  $L$  be some number field containing  $\kappa(x)$ , let  $i: \text{Spec } \mathcal{O}_L \rightarrow \mathcal{X}$  be the morphism over  $\text{Spec } \mathcal{O}_k$  corresponding to  $x$ , and let  $I$  be the ideal in  $\mathcal{O}_L$  corresponding to the ideal sheaf  $i^*\mathcal{I} \cdot \mathcal{O}_{\text{Spec } \mathcal{O}_L}$  on  $\text{Spec } \mathcal{O}_L$ . Then we define

$$\begin{aligned} N_S(\mathcal{I}, x) &= \frac{1}{[L:k]} \log(\mathcal{O}_{L,S} : I\mathcal{O}_{L,S}) \\ &= \frac{1}{[L:k]} \sum_{\substack{\mathfrak{q} \in \text{Spec } \mathcal{O}_{L,S} \\ \mathfrak{q} \neq (0)}} \text{ord}_{\mathfrak{q}} I \cdot \log(\mathcal{O}_{L,S} : \mathfrak{q}). \end{aligned}$$

It can be shown (although we will not do so here) that if  $\mathcal{X}$  is a proper model over  $\text{Spec } \mathcal{O}_k$  for a complete variety  $X$  over  $k$ , if  $\mathcal{I}$  is an ideal sheaf on  $\mathcal{X}$ , and if the restriction of  $\mathcal{I}$  to  $X$  corresponds to a Cartier divisor  $D$ , then

$$(23.11) \quad N_S(\mathcal{I}, x) = N_S(D, x) + O(1)$$

for all  $x \in X(\bar{k})$  not lying in  $\text{Supp } D$ .

Counting (and proximity and height) functions relative to sheaves of ideals were first introduced by Yamanoi [2004a], in the context of Nevanlinna theory.

The different can be described via differentials as well. Indeed,  $\mathcal{D}_{L/k}$  is the annihilator of the sheaf of relative differentials:

$$(23.12) \quad \mathcal{D}_{L/k} = \text{Ann } \Omega_{\mathcal{O}_L/\mathcal{O}_k}.$$

In this case  $\Omega_{\mathcal{O}_L/\mathcal{O}_k}$  is a torsion sheaf locally generated by one element; hence (23.3.1) can be rewritten as

$$\begin{aligned} d_S(L) &= \frac{1}{[L:k]} \sum_{\substack{\mathfrak{q} \in \text{Spec } \mathcal{O}_{L,S} \\ \mathfrak{q} \neq (0)}} \text{length}_{\mathfrak{q}} \Omega_{\mathcal{O}_L/\mathcal{O}_k} \cdot \log(\mathcal{O}_{L,S} : \mathfrak{q}) \\ (23.13) \quad &= \frac{1}{[L:k]} \log \#H^0(\mathcal{O}_{L,S}, \Omega_{\mathcal{O}_L/\mathcal{O}_k}). \end{aligned}$$

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 23.14.** *Let  $A \rightarrow B$  be a local homomorphism of discrete valuation rings, with  $B$  finite over  $A$ , let  $\phi: X \rightarrow Y$  be a generically finite morphism of schemes, and let*

$$(23.14.1) \quad \begin{array}{ccc} \text{Spec } B & \xrightarrow{j} & X \\ \downarrow & & \downarrow \phi \\ \text{Spec } A & \longrightarrow & Y \end{array}$$

*be a commutative diagram. Assume also that the image of  $j$  is not contained in the support of  $\Omega_{X/Y}$ , that  $j^* \mathcal{O}_X$  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*

$$(23.14.2) \quad \mathcal{D}_{B/A} \supseteq j^* F_0(\Omega_{X/Y}) \cdot B.$$

*Proof.* The map  $j$  factors through the product  $\text{Spec } A \times_Y X$ :

$$\begin{array}{ccccc} \text{Spec } B & \xrightarrow{j''} & \text{Spec } A \times_Y X & \xrightarrow{q} & X \\ & & \downarrow & & \downarrow \\ & & \text{Spec } A & \longrightarrow & Y ; \end{array}$$

here  $j = q \circ j''$ . We may replace  $\text{Spec } A \times_Y X$  in this diagram with an open affine neighborhood  $\text{Spec } B''$  of  $j''(b)$ , where  $b$  denotes the closed point of  $\text{Spec } B$ . By Lemma 23.7a, we have

$$(23.14.3) \quad j^* F_0(\Omega_{X/Y}) \cdot B = (j'')^* F_0(\Omega_{B''/A}) \cdot B .$$

The map  $\text{Spec } B \rightarrow \text{Spec } B''$  corresponds to a ring homomorphism  $B'' \rightarrow B$ ; let  $B'$  denote its image. Then  $j''$  factors through  $j' : \text{Spec } B \rightarrow \text{Spec } B'$  and a closed immersion  $\text{Spec } B' \rightarrow \text{Spec } B''$ . By the second exact sequence for differentials, the map  $\Omega_{B''/A}|_{B'} \rightarrow \Omega_{B'/A}$  is surjective, and by Lemma 23.7b

$$(23.14.4) \quad (j'')^* F_0(\Omega_{B''/A}) \cdot B \subseteq (j')^* F_0(\Omega_{B'/A}) \cdot B .$$

By [Kunz, 1986, 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  $A$ , [Kunz, 1986, Prop. 10.22] gives

$$(23.14.5) \quad F_0(\Omega_{B'/A}) \subseteq \mathcal{D}_{B'/A}$$

(note that  $\mathcal{D}_{B'/A}$  is  $\mathfrak{d}_D(B'/A)$  in Kunz's notation). Finally, since  $B' \subseteq B$ , it follows directly from the definition (see for example [Kunz, 1986, G.9a]) that

$$(23.14.6) \quad \mathcal{D}_{B'/A} \subseteq \mathcal{D}_{B/A} .$$

Combining (23.14.3)–(23.14.6) then gives (23.14.2). □

The generalized Chevalley-Weil theorem can now be stated as follows. This was originally proved in [Vojta, 1987, 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 23.15.** *Let  $\phi : X \rightarrow Y$  be a generically finite, dominant morphism of non-singular complete varieties over  $k$ , with ramification divisor  $R$ . Then for all  $x \in X(\bar{k})$  not lying on  $\text{Supp } R$ , we have*

$$(23.15.1) \quad d_S(x) - d_S(\phi(x)) \leq N_S(R, x) + O(1) .$$

*Proof.* Let  $\mathcal{X}$  and  $\mathcal{Y}$  be models for  $X$  and  $Y$  over  $\mathcal{O}_k$ , respectively. By replacing  $\mathcal{X}$  with the closure of the graph of the rational map  $\phi : \mathcal{X} \rightarrow \mathcal{Y}$  if necessary, we may assume that  $\phi$  extends to a morphism  $\mathcal{X} \rightarrow \mathcal{Y}$  over  $\text{Spec } \mathcal{O}_k$ . By Lemma 23.8 and

(23.11), it will then suffice to show that

$$(23.15.2) \quad d_S(x) - d_S(\phi(x)) \leq N_S(F_0(\Omega_{\mathcal{X}/\mathcal{Y}}), x).$$

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  $\mathcal{O}_w$  and  $\mathcal{O}_v$  denote the localizations of  $\mathcal{O}_E$  and  $\mathcal{O}_L$  at the primes corresponding to  $w$  and  $v$ , respectively. Let  $j: \text{Spec } \mathcal{O}_w \rightarrow \mathcal{X}$  be the restriction of the map  $\text{Spec } \mathcal{O}_E \rightarrow \mathcal{X}$  over  $\mathcal{O}_k$  corresponding to  $x$ . By (23.3.1), multiplicativity of the different in towers, Definition 23.10, and compatibility of various things with localization, it suffices to show that

$$(23.15.3) \quad \mathcal{D}_{\mathcal{O}_w/\mathcal{O}_v} \supseteq j^*F_0(\Omega_{\mathcal{X}/\mathcal{Y}}) \cdot \mathcal{O}_w.$$

The point  $\phi(x) \in Y(L)$  determines a map  $\text{Spec } \mathcal{O}_v \rightarrow \mathcal{Y}$ , so there is a diagram

$$\begin{array}{ccc} \text{Spec } \mathcal{O}_w & \xrightarrow{j} & \mathcal{X} \\ \downarrow & & \downarrow \\ \text{Spec } \mathcal{O}_v & \longrightarrow & \mathcal{Y}. \end{array}$$

This satisfies the conditions of Lemma 23.14, which implies (23.15.3). □

*Remark 23.16.* More generally, the above proof shows that if one replaces (23.15.1) with (23.15.2), then Theorem 23.15 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 23.17.** *Let  $B$  be a connected (nonempty) Riemann surface, let  $\phi: X \rightarrow Y$  be a generically finite, dominant morphism of smooth complete complex varieties, with ramification divisor  $R$ , and let  $f: B \rightarrow X$  be a holomorphic function whose image is not contained in  $\phi(\text{Supp } R)$ . Then there is a connected Riemann surface  $B'$ , a proper surjective holomorphic map  $\pi: B' \rightarrow B$  of degree bounded by  $[K(X) : K(Y)]$ , and a holomorphic function  $g: B' \rightarrow X$  such that the diagram*

$$(23.17.1) \quad \begin{array}{ccc} B' & \xrightarrow{g} & X \\ \downarrow \pi & & \downarrow \phi \\ B & \xrightarrow{f} & Y \end{array}$$

*commutes. Moreover, if  $e$  is the ramification index of  $B'$  over  $B$  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  $\pi$  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  $B_0 \rightarrow B$  have at most  $d$  points, and some fibers have exactly  $d$  points). Let  $B'$  be the normalization of  $B_0$  [Gunning, 1990, 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  $\leq d$  over  $B$ ). We then have holomorphic functions  $\pi : B' \rightarrow B$  and  $g : B' \rightarrow X$  as in (23.17.1). Also,  $B'$  is a Riemann surface [Gunning, 1990, 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  $\mathcal{O}$  of germs of holomorphic functions on  $B$  at  $b$  with the ring of germs of holomorphic functions on  $\mathbb{C}$  at 0. Also let  $\mathcal{O}'$  be the ring of germs of holomorphic functions on  $B'$  at  $b'$ , and let  $e$  be the ramification index of  $\pi$  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  $\mathcal{O}$  with a subring of  $\mathcal{O}'$  via  $\pi$ . By [Gunning, 1990, C5 and C8], there is a canonically defined monic polynomial  $P \in \mathcal{O}[t]$  of degree  $e$  such that  $P(z') = 0$ , and

$$\mathcal{O}' \cong \mathcal{O}[t]/P(t).$$

Since  $B'$  is regular at  $b'$ , the germ of the variety  $B'$  at  $b'$  is irreducible, so  $\mathcal{O}'$  is an entire ring [Gunning, 1990, B6]. Therefore  $P(t)$  is irreducible, and by the Weierstrass Preparation Theorem [Gunning, 1990, A4] it is a Weierstrass polynomial. This means that all non-leading coefficients vanish at  $b$ . A straightforward computation then gives

$$\Omega_{\mathcal{O}'/\mathcal{O}} = \mathcal{O}'/(z')^{e-1}.$$

By [Gunning, 1990, A8 and G20],  $\mathcal{O}$  and  $\mathcal{O}'$  are discrete valuation rings, hence Dedekind, and then (23.12) gives

$$\mathcal{D}_{\mathcal{O}'/\mathcal{O}} = (z')^{e-1}.$$

Moreover, we have a commutative diagram

$$\begin{array}{ccc} \text{Spec } \mathcal{O}' & \xrightarrow{j} & X \\ \downarrow & & \downarrow \phi \\ \text{Spec } \mathcal{O} & \longrightarrow & Y \end{array}$$

which satisfies the conditions of Lemma 23.14. Therefore,

$$\mathcal{D}_{\mathcal{O}'/\mathcal{O}} \supseteq j^* F_0(\Omega_{X/Y}) \cdot \mathcal{O}'.$$

This, together with Lemma 23.8, implies the theorem. □

*Remark 23.18.* As was the case in Remark 23.16, Theorem 23.17 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  $\pi$  is bounded by the analytic divisor associated to  $g^*F_0(\Omega_{X/Y})$ .

## 24 A Diophantine Conjecture for Algebraic Points

This section describes an extension of Conjecture 14.6 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 26.8.

*Conjecture 24.1.* 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  $\mathcal{H}$  be the canonical line sheaf on  $X$ , let  $\mathcal{A}$  be an ample line sheaf on  $X$ , and let  $r$  be a positive integer. Then:

- (a). Let  $\Sigma$  be a generic subset of  $X(\bar{k}) \setminus \text{Supp } D$  such that  $[\kappa(x) : k] \leq r$  for all  $x \in \Sigma$ . Then the inequality

$$(24.1.1) \quad m_S(D, x) + h_{\mathcal{H}, k}(x) \leq d_k(x) + O(\log^+ h_{\mathcal{A}, k}(x))$$

holds for all  $x \in \Sigma$ .

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

$$(24.1.2) \quad m_S(D, x) + h_{\mathcal{H}, k}(x) \leq d_k(x) + \varepsilon h_{\mathcal{A}, k}(x) + C$$

holds for almost all  $x \in (X \setminus Z)(\bar{k})$  with  $[\kappa(x) : k] \leq r$ .

When  $r = 1$ , this just reduces to Conjecture 14.6, 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  $\varepsilon$  results in a stronger statement.

As was the case with Conjecture 14.6, this Conjecture 24.1 can also be posed for smooth complete varieties  $X$ , provided that  $h_{\mathcal{A}, k}$  is replaced by a big height.

Remark 14.7 does not extend trivially to Conjecture 24.1, though, since the discriminant terms may not add up.



By [Vojta, 1987, Prop. 5.4.1],<sup>3</sup> Conjecture 24.1b with  $D = 0$  implies the full Conjecture 24.1b. This uses a covering construction.

Next, we show how Conjecture 24.1 relates to generically finite ramified covers.

**Proposition 24.2.** *Let  $k$  and  $S$  be as in Conjecture 24.1, let  $\pi: X' \rightarrow X$  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  $\pi^*D$ ), and assume that it too has normal crossings. Let  $\mathcal{H}$  and  $\mathcal{H}'$  denote the canonical line sheaves on  $X$  and  $X'$ , respectively. Then, for all  $x \in X'(\bar{k})$  not lying on  $\text{Supp} D'$  or on the support of the ramification divisor,*

$$(24.2.1) \quad m_S(D, \pi(x)) + h_{\mathcal{H},k}(\pi(x)) - d_S(\pi(x)) \leq m_S(D', x) + h_{\mathcal{H}',k}(x) - d_S(x) + O(1) .$$

*In particular, since the pull-back of any big line sheaf on  $X$  to  $X'$  remains big, either part of Conjecture 24.1 for  $D'$  on  $X'$  implies that same part for  $D$  on  $X$ .*

*Proof.* Let  $R$  be the ramification divisor for  $X'$  over  $X$ ; since  $\mathcal{H}' \cong \pi^* \mathcal{H} \otimes \mathcal{O}(R)$ , we then have

$$h_{\mathcal{H},k}(\pi(x)) - h_{\mathcal{H}',k}(x) = -m_S(R, x) - N_S(R, x) + O(1) .$$

Also, Theorem 23.15 gives

$$d_S(x) - d_S(\pi(x)) \leq N_S(R, x) + O(1) .$$

Finally, by [Vojta, 1987, Lemma 5.2.2],  $\pi^*D - (\pi^*D)_{\text{red}} \leq R$  (relative to the cone of effective divisors). Therefore

$$m_S(D, \pi(x)) - m_S(D', x) \leq m_S(R, x) + O(1) .$$

Adding this equation and the two inequalities then gives (24.2.1). □

It was this proposition that motivated the original version of Conjecture 24.1 [Vojta, 1987, p. 63].

Finally, we note that Conjecture 24.1 can also be posed with truncated counting functions.

*Conjecture 24.3.* Let  $k, S, X, D, \mathcal{H}, \mathcal{A}$ , and  $r$  be as in Conjecture 24.1. Then:

- (a). Let  $\Sigma$  be a generic subset of  $X(\bar{k}) \setminus \text{Supp} D$  such that  $[\kappa(x) : k] \leq r$  for all  $x \in \Sigma$ . Then the inequality

$$(24.3.1) \quad N_S^{(1)}(D, x) + d_k(x) \geq h_{\mathcal{H}(D),k}(x) - O(\log^+ h_{\mathcal{A},k}(x))$$

holds for all  $x \in \Sigma$ .

---

<sup>3</sup> 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  $\text{Supp} D$  has an open neighborhood  $U$  such that  $D = (f_i)$  on  $U$  for some  $i$ .

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

$$(24.3.2) \quad N_S^{(1)}(D, x) + d_k(x) \geq h_{\mathcal{H}(D), k}(x) - \varepsilon h_{\mathcal{A}, k}(x) - C$$

holds for almost all  $x \in (X \setminus Z)(\bar{k})$  with  $[\kappa(x) : k] \leq r$ .

*Remark 24.4.* Using a covering construction, it has been shown that Conjecture 24.3b would follow from Conjecture 24.1b [Vojta, 1998]. As noted earlier in this section, the latter would then follow from Conjecture 24.1b 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 24.3b has been fully proved for curves over function fields of characteristic 0.

Proposition 24.2 does not extend to the situation of truncated counting functions, though.

## 25 The $1 + \varepsilon$ Conjecture and the abc Conjecture

The special case of Conjecture 24.1 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 25.1.* Let  $k$  be a number field, let  $X$  be a smooth projective curve over  $k$ , let  $\mathcal{H}$  denote the canonical line sheaf on  $X$ , let  $r$  be a positive integer, let  $\varepsilon > 0$ , and let  $C \in \mathbb{R}$ . Then

$$h_{\mathcal{H}, k}(x) \leq (1 + \varepsilon)d_k(x) + C$$

for almost all  $x \in X(\bar{k})$  with  $[\kappa(x) : k] \leq r$ .

This conjecture was recently proved over function fields of characteristic 0 by McQuillan [2008] and (independently) by Yamanoi [2004b]. See also McQuillan [2009] and Gasbarri [2008]. Thus, Conjectures 24.1 and 24.3 hold for curves over function fields of characteristic 0, by Remark 24.4; see also [Yamanoi, 2004b, Thm. 5]

Conjecture 25.1 is known to imply the abc conjecture [Vojta, 1987, pp. 71–72].

**Proposition 25.2.** *If Conjecture 25.1 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

$$P_n = [\sqrt[n]{a} : \sqrt[n]{b} : \sqrt[n]{c}] \in X_n(\mathbb{Q}),$$

where  $X_n$  is the nonsingular curve  $x_0^n + x_1^n + x_2^n = 0$  in  $\mathbb{P}_{\mathbb{Q}}^2$ . This point has height

$$h_{\mathcal{X}, \mathbb{Q}}(P_n) = \frac{n-3}{n} \log \max\{|a|, |b|, |c|\} + O(1),$$

since the canonical line sheaf  $\mathcal{X}$  on  $X_n$  is the restriction of  $\mathcal{O}(n-3)$ . Here the implicit constant depends only on  $n$ . We also have

$$d_{\mathbb{Q}}(P_n) \leq \frac{n-1}{n} \sum_{p|abc} \log p + O(1),$$

where the implicit constant depends only on  $n$ . Therefore, applying Conjecture 25.1 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

$$\frac{n-3}{n} \log \max\{|a|, |b|, |c|\} \leq \left(\frac{n-1}{n} + \varepsilon'\right) \sum_{p|abc} \log p + C_{n, \varepsilon'}.$$

The proof concludes by taking  $n$  sufficiently large and  $\varepsilon'$  sufficiently small so that

$$\left(\frac{n-1}{n} + \varepsilon'\right) / \frac{n-3}{n} < 1 + \varepsilon,$$

and noting that the constants in the above discussion are independent of the triple  $(a, b, c)$ . □

## 26 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 [1990, Ch. III] and Yamanoi [2004b].

Throughout this section,  $B$  is a connected Riemann surface,  $\pi: B \rightarrow \mathbb{C}$  is a proper surjective holomorphic map,  $X$  is a smooth complete complex variety, and  $f: B \rightarrow X$  is a holomorphic function.

Note that  $\pi$  has a well-defined, finite degree, denoted  $\deg \pi$ .

We again refer to  $f$  as a holomorphic curve.

**Definition 26.1.** Define

$$\begin{aligned} B[r] &= \{b \in B : |\pi(b)| \leq r\}, \\ B(r) &= \{b \in B : |\pi(b)| < r\}, \quad \text{and} \\ B\langle r \rangle &= \{b \in B : |\pi(b)| = r\}. \end{aligned}$$

On  $B\langle r \rangle$ , let  $\sigma$  be the measure

$$\sigma = \frac{1}{\deg \pi} \pi^* \left( \frac{d\theta}{2\pi} \right).$$

**Definition 26.2.** Let  $D$  be an effective divisor on  $X$  whose support does not contain the image of  $f$ , and let  $\lambda_D$  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 26.3.** (a). The **counting function** for an analytic divisor  $\Delta = \sum_{b \in B} n_b \cdot b$  on  $B$  is

$$N_\Delta(r) = \frac{1}{\deg \pi} \left( \sum_{b \in B\langle r \rangle \setminus \pi^{-1}(0)} n_b \log \frac{r}{|\pi(b)|} + \sum_{b \in \pi^{-1}(0)} n_b \log r \right).$$

(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  $\pi$  is the counting function for the ramification divisor of  $\pi$ . It is denoted  $N_{\text{Ram}(\pi)}(r)$ .

If  $B = \mathbb{C}$  and  $\pi$  is the identity mapping, then the proximity function of Definition 26.2 and the counting function of Definition 26.3b extend those of Definitions 11.1 and 11.2, respectively. They also satisfy additivity, functoriality, and boundedness properties, as in Proposition 11.3.

If  $B'$  is another connected Riemann surface and  $\pi' : B' \rightarrow B$  is another proper surjective holomorphic map, then

$$(26.4) \quad m_{f \circ \pi'}(D, r) = m_f(D, r) \quad \text{and} \quad N_{f \circ \pi'}(D, r) = N_f(D, r).$$

This holds in particular if  $B = \mathbb{C}$  and  $\pi$  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 10.1.

In this situation (but without the assumption  $B = \mathbb{C}$ ), we also have

$$(26.5) \quad N_{\text{Ram}(\pi \circ \pi')}(r) = N_{\text{Ram}(\pi)}(r) + N_{\text{Ram}(\pi')}(r).$$

(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  $\text{Ram}(\pi)$  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  $\text{Ram}(\pi)$  may have infinite support.

The height is defined similarly to Definitions 11.4 and 11.7:

**Definition 26.6.** 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  $\mathcal{L}$  is a line sheaf on  $X$ , then the height  $T_{\mathcal{L},f}(r)$  is defined to be  $T_{D,f}(r)$  for any divisor  $D$  on  $X$  for which  $\mathcal{O}(D) \cong \mathcal{L}$  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 [Lang and Cherry, 1990, III Thm. 2.1]. Theorem 11.8 also holds for heights in this context, as do Propositions 11.10 and 11.11. Corollary 11.9 is not meaningful in this context, since  $B$  need not be algebraic.

If  $\pi' : B' \rightarrow B$  is as in (26.4) and  $D$  and  $\mathcal{L}$  are as in Definition 26.6, then

$$T_{D,f \circ \pi'}(r) = T_{D,f}(r) + O(1) \quad \text{and} \quad T_{\mathcal{L},f \circ \pi'}(r) = T_{\mathcal{L},f}(r) + O(1) .$$

Griffiths' conjecture (Conjecture 14.2) can be posed in this context, without any changes other than the domain of the holomorphic curve  $f$ , and adding terms  $N_{\text{Ram}(\pi)}(r)$  to the right-hand sides of (14.2.1) and (14.2.2). It will not be repeated here.

The variant with truncated counting functions reads as follows.

*Conjecture 26.7.* Let  $X$  be a smooth complex projective variety, let  $D$  be a normal crossings divisor on  $X$ , let  $\mathcal{H}$  be the canonical line sheaf on  $X$ , and let  $\mathcal{A}$  be an ample line sheaf on  $X$ . Then:

- (a). The inequality

$$(26.7.1)$$

$$N_f^{(1)}(D, r) + N_{\text{Ram}(\pi)}(r) \geq_{\text{exc}} T_{\mathcal{H}(D),f}(r) - O(\log^+ T_{\mathcal{A},f}(r)) - o(\log r)$$

holds for all holomorphic curves  $f : B \rightarrow X$  with Zariski-dense image.

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

$$(26.7.2) \quad N_f^{(1)}(D, r) + N_{\text{Ram}(\pi)}(r) \geq_{\text{exc}} T_{\mathcal{H}(D),f}(r) - \varepsilon T_{\mathcal{A},f}(r) - C$$

holds for all nonconstant holomorphic curves  $f : B \rightarrow X$  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 14.2 is proved (except possibly for the level of truncation of the counting functions). See for example [Lang and Cherry, 1990], and also Corollary 28.7 for the case when  $\dim X = 1$ .

*Remark 26.8.* This conjecture, and therefore also Conjecture 24.1, is doubted by some. For example, McQuillan [2000, Example V.1.5] notes that if  $X$  is a quotient of the unit ball in  $\mathbb{C}^2$ , if  $f: B \rightarrow X$  is a one-dimensional geodesic, and if  $\pi: B \rightarrow \mathbb{C}$  is a proper surjective holomorphic map, then

$$T_{\mathcal{K}_X, f}(r) = N_{\text{Ram}(\pi)}(r) + o(T_{\mathcal{K}_X, f}(r)).$$

However, *loc. cit.* does not address how to show that a suitable map  $\pi$  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.

## 27 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 [Lang, 1983, 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  $B' \rightarrow B$  over  $F$ , then the logarithmic discriminant term in the function field case is defined as

$$(27.1) \quad d_k(L) = \frac{\deg \mathcal{K}'}{[L:k]} - \deg \mathcal{K},$$

where  $\mathcal{K}'$  and  $\mathcal{K}$  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

$$(27.2) \quad d_k(L) = \frac{1}{[L:k]} \dim_F H^0(B', \Omega_{B'/B})$$

(cf. (23.13)).

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  $\mathcal{X}$  is a product  $X_0 \times_F B$  (so that the model *splits* into a product), and  $\pi: \mathcal{X} \rightarrow B$  is the projection morphism to the second factor.

Following early work of de Franchis [Lang, 1983, p. 223], it is fairly easy to prove the  $1 + \varepsilon$  conjecture in the split function field case of characteristic 0.

**Theorem 27.3.** *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  $\mathcal{X} = X_0 \times_F B$ . Let  $\mathcal{K}$  be the canonical line sheaf on  $X_0$ , and let  $\mathcal{A}$  be an ample line sheaf on  $X_0$ . View both of these line sheaves as line sheaves on  $X$  or on  $\mathcal{X}$  by pulling back via the projection morphisms. Then*

$$(27.3.1) \quad h_{\mathcal{K},k}(x) \leq d_k(x) + \deg \Omega_{B/F}$$

for all  $x \in X(\bar{k})$ .

*Proof.* The proof is particularly easy if  $F$  is algebraically closed.

In that case, let  $q: \mathcal{X} \rightarrow X_0$  denote the projection morphism. If  $q \circ i$  is a constant morphism, then by (15.5) the left-hand side of (27.3.1) is zero. Since the right-hand side is nonnegative by (27.2), the inequality is true in this case.

If  $q \circ i$  is nonconstant, then it is finite and surjective, and we have

$$h_{\mathcal{K},k}(x) = \frac{(2g(X_0) - 2) \deg(q \circ i)}{[K(B') : k]} \leq \frac{2g(B') - 2}{[K(B') : k]} = d_k(x) + 2g(B) - 2.$$

by (15.5), the Riemann-Hurwitz formula (twice), and by (27.1). (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 (27.3.1) 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  $i: B' \rightarrow \mathcal{X}$  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  $X_0 \times_F \bar{F}$ ,  $B$  with  $B \times_F \bar{F}$ ,  $\mathcal{X}$  with  $\mathcal{X} \times_F \bar{F}$ , and  $B'$  with  $B' \times_{F'} \bar{F}$ . This again does not affect the validity of (27.3.1), since both sides are divided by  $[F' : F]$ . Indeed, replacing  $B'$  with  $B' \times_{F'} \bar{F}$  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  $B' \times_{F'} \bar{F}$  instead of  $B' \times_F \bar{F}$  for some choice of embedding  $F' \hookrightarrow \bar{F}$ . Also, these changes do not affect the fact that  $\mathcal{X} = X_0 \times_F B$ .

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  $i: B' \rightarrow \mathcal{X}$ . It takes values in the absolute tangent bundle  $T_{\mathcal{X}/F}$ . Since  $\mathcal{X}$  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  $T_{\mathcal{K},k}(x)$ .

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  $T_{\mathcal{X}/B}$ . This is a subbundle of the absolute tangent bundle, not a quotient, and there is no canonical projection from  $T_{\mathcal{X}/F}$  to  $T_{\mathcal{X}/B}$ . McQuillan's proof works mainly because, for points of large height, the tangent vectors giving the derivative of  $i: B' \rightarrow \mathcal{X}$  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.

## 28 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 28.1). 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  $\pi: B \rightarrow \mathbb{C}$  is a proper surjective holomorphic map.

Nevanlinna’s original Lemma on the Logarithmic Derivative (LLD) is the following.

**Theorem 28.1.** (Lemma on the Logarithmic Derivative) *Let  $f$  be a meromorphic function on  $\mathbb{C}$ . Then*

$$(28.1.1) \quad \int_0^{2\pi} \log^+ \left| \frac{f'(re^{i\theta})}{f(re^{i\theta})} \right| \frac{d\theta}{2\pi} \leq_{\text{exc}} O(\log^+ T_f(r)) + o(\log r).$$

*More generally, if  $f$  is a meromorphic function on  $B$ , then*

$$(28.1.2) \quad \int_{B(r)} \log^+ \left| \frac{df/\pi^* dz}{f} \right| \sigma \leq_{\text{exc}} O(\log T_f(r) + \log r).$$

*Proof.* For the first part, see [Nevanlinna, 1970, IX 3.3], or [Shiffman, 1983, Thm. 3.11] for the error term given here. The second part follows from [Ashline, 1999, Thm. 2.2].  $\square$

A geometrical adaptation of this lemma has recently been discovered by R. Kobayashi, M. McQuillan, P.-M. Wong, and others. This first requires a definition.

**Definition 28.2.** Let  $X$  be a smooth complex projective variety, and let  $D$  be a normal crossings divisor on  $X$ . Then the sheaf  $\Omega_X^1(\log D)$  is the subsheaf of the sheaf of meromorphic sections of  $\Omega_X^1$  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  $D$  [Deligne, 1970, 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 28.3.** (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  $f: B \rightarrow X$  be a holomorphic curve whose image is not contained in  $\text{Supp } D$ . Let  $\mathcal{A}$  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  $d_D f: B \rightarrow T_X(-\log D)$  denote the canonical lifting of  $f$  (as a meromorphic function). Then*

$$(28.3.1) \quad \int_{B(r)} \log^+ |d_D f(re^{i\theta})| \sigma \leq_{\text{exc}} O(\log T_{\mathcal{A},f}(r) + \log r).$$

*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 (28.3.1) reduces to (28.1.1).

The assertion of the theorem is preserved by taking products. Indeed, if it holds for holomorphic curves  $f_1: B \rightarrow X_1$  and  $f_2: B \rightarrow X_2$  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  $(f_1, f_2): B \rightarrow X_1 \times X_2$  relative to the normal crossings divisor  $D := p_1^*D_1 + p_2^*D_2$ , where  $p_j: X_1 \times X_2 \rightarrow X_j$  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 (5.4.2); 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  $\phi: Z \dashrightarrow X'$  be a rational map. Assume that there is a nonempty Zariski-open subset  $U$  of  $Z$  and a constant  $C > 0$  such that  $|\phi_*(v)| \geq C|v|$  for all  $v \in T_X(-\log D)$  lying over  $U$ , that the holomorphic curve  $f$  meets  $U$ , and that Theorem 28.3 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 (28.3.1) 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 11.11 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  $\pi: X' \rightarrow X$ , isomorphic over  $X \setminus \text{Supp } D$ , 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  $\pi_*$  induces a holomorphic map  $T_{X'}(-\log D') \rightarrow T_X(-\log D)$ , the inverse rational map  $\pi^{-1}$  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  $\phi: Z \dashrightarrow (\mathbb{P}^1)^n$  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  $p_j: (\mathbb{P}^1)^n \rightarrow \mathbb{P}^1$  is the projection morphism to the  $j^{\text{th}}$  factor,  $1 \leq j \leq n$ .

To satisfy the conditions of the claim, it suffices to find a finite set  $\mathcal{G}$  of nonzero functions in  $K(X)$  such that at each closed point  $z \in Z$  there is a subset  $\mathcal{G}_z \subseteq \mathcal{G}$  such that for each  $g \in \mathcal{G}_z$  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  $\mathcal{G}_z$  the differentials  $dg/g$  generate  $\Omega_X^1(\log D)$  at  $z$ .

To construct  $\mathcal{G}$ , 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  $\mathcal{G}$  satisfying the condition everywhere on  $Z$ .  $\square$

*Remark 28.4.* There is no  $N_{\text{Ram}(\pi)}(r)$  term in either of these theorems; it appears subsequently. The same is true of the exceptional set (it appears later still).

*Remark 28.5.* When  $B = \mathbb{C}$  and  $\pi$  is the identity map, the error term in (28.3.1) can be sharpened to  $O(\log^+ T_{\mathcal{A},f}(r)) + o(\log r)$ . 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 [1998, 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  $\mathcal{E}$  is a quasi-coherent sheaf on a scheme  $X$ , then

$$\mathbb{P}(\mathcal{E}) = \mathbf{Proj} \bigoplus_{d \geq 0} S^d \mathcal{E}$$

(as in EGA). In particular, if  $\mathcal{E}$  is a vector sheaf, then points on  $\mathbb{P}(\mathcal{E})$  correspond bijectively to *hyperplanes* (not lines) in the fiber over the corresponding point on  $X$ . This scheme comes with a **tautological line sheaf**  $\mathcal{O}(1)$ , which gives rise to the name of McQuillan's inequality.

If  $X$  and  $D$  are as in Theorem 28.3, then  $\Omega_X^1(\log D)$  is defined as a locally free sheaf on  $X$  as an analytic space. This is a coherent sheaf, hence by GAGA [Serre, 1955–1956], 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 Section 29.

**Theorem 28.6.** (McQuillan’s “Tautological Inequality”) *Let  $X, D, f: B \rightarrow X$ , and  $\mathcal{A}$  be as in Theorem 28.3. Assume also that  $f$  is not constant. Let*

$$f': B \rightarrow \mathbb{P}(\Omega_{X/\mathbb{C}}(\log D))$$

*be the canonical lifting of  $f$ , associated to the nonzero map from  $f^*\Omega_X^1(\log D)$  to the cotangent sheaf of  $B$ . Then*

$$(28.6.1) \quad T_{\mathcal{O}(1),f'}(r) \leq_{\text{exc}} N_f^{(1)}(D, r) + N_{\text{Ram}(\pi)}(r) + O(\log T_{\mathcal{A},f}(r) + \log r).$$

*Proof.* Let

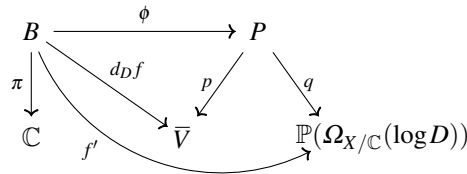
$$V = \mathbb{V}(\Omega_{X/\mathbb{C}}(\log D)) = \mathbf{Spec} \bigoplus_{d \geq 0} S^d \Omega_{X/\mathbb{C}}(\log D).$$

This is the total space of  $T_X(-\log D)$ . Also let

$$\bar{V} = \mathbb{P}(\Omega_{X/\mathbb{C}}(\log D) \oplus \mathcal{O}_X).$$

We have a natural embedding  $V \hookrightarrow \bar{V}$  that realizes  $\bar{V}$  as the projective closure on fibers of  $V$ .

Let  $[\infty]$  denote the (reduced) divisor  $\bar{V} \setminus V$ . The integrand of (28.3.1) can be viewed as a proximity function for  $[\infty]$ , and the strategy of the proof is to use this to get a bound on  $T_{\mathcal{O}(1),f'}(r)$ , via the rational map  $\bar{V} \dashrightarrow \mathbb{P}(\Omega_{X/\mathbb{C}}(\log D))$ . To compare the geometries of these two objects, we use the closure of the graph of this rational map in  $\bar{V} \times_X \mathbb{P}(\Omega_{X/\mathbb{C}}(\log D))$ , which is the blowing-up of  $\bar{V}$  along the image  $[0]$  of the zero section. Let  $p: P \rightarrow \bar{V}$  be this blowing-up, let  $E$  be its exceptional divisor. Let  $q: P \rightarrow \mathbb{P}(\Omega_{X/\mathbb{C}}(\log D))$  be the projection to the second factor. We have a diagram



There is a unique lifting  $\phi: B \rightarrow P$  that satisfies  $d_D f = p \circ \phi$  and  $f' = q \circ \phi$ . We also have

$$p^* \mathcal{O}(1) \cong q^* \mathcal{O}(1) \otimes \mathcal{O}(E)$$

(where the first  $\mathcal{O}(1)$  is on  $\bar{V}$  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  $\mathcal{O}(1)$  on  $\bar{V}$ , and also a rational section of  $\mathcal{O}(1)$  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  $\mathcal{O}([\infty]) \cong \mathcal{O}(1)$  on  $\bar{V}$ , because the divisor  $[\infty]$  is cut out by the section  $(0, 1)$  of  $\Omega_{X/\mathbb{C}}(\log D) \oplus \mathcal{O}_X$ .

Thus, we have

$$\begin{aligned} T_{\mathcal{O}(1),f'}(r) &= T_{q^*\mathcal{O}(1),\phi}(r) + O(1) \\ &= T_{\mathcal{O}(1),d_D f}(r) - T_{\mathcal{O}(E),\phi}(r) + O(1) \\ &= N_{d_D f}([\infty], r) - T_{\mathcal{O}(E),\phi}(r) + m_{d_D f}([\infty], r) + O(1) \\ &\leq_{\text{exc}} N_f^{(1)}(D, r) + N_{\text{Ram}(\pi)}(r) + O(\log T_{\mathcal{A},f}(r) + \log r) . \end{aligned}$$

To explain the last step above,  $m_{d_D f}([\infty], r)$  is bounded by Theorem 28.3. Since  $E$  is effective and does not contain the image of  $\phi$  (since  $f$  is not constant),  $T_{\mathcal{O}(E),\phi}(r)$  is bounded from below. (It can also be used to subtract a term  $N_{\text{Ram}(f)}(r)$  from the right-hand side of (28.6.1).)

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  $f(b)$ ,  $\bar{V}$  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

$$\begin{aligned} &\left[ \frac{d(z_1 \circ f)/dz}{z_1 \circ f} : \dots : \frac{d(z_r \circ f)/dz}{z_r \circ f} : \frac{d(z_{r+1} \circ f)}{dz} : \dots : \frac{d(z_n \circ f)}{dz} : 1 \right] \\ &= \left[ \frac{d(z_1 \circ f)/dw}{z_1 \circ f} : \dots : \frac{d(z_r \circ f)/dw}{z_r \circ f} : \frac{d(z_{r+1} \circ f)}{dw} : \dots : \frac{d(z_n \circ f)}{dw} : \frac{dz}{dw} \right] . \end{aligned}$$

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  $f(b) \in \text{Supp} D$ . 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  $N_{\text{Ram}(\pi)}(r)$ .  $\square$

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 26.7.

**Corollary 28.7.** *Let  $X$  be a smooth complex projective curve, let  $D$  be an effective reduced divisor on  $X$ , and let  $f: B \rightarrow X$  be a non-constant holomorphic curve. Then*

$$(28.7.1) \quad N_f^{(1)}(D, r) + N_{\text{Ram}(\pi)}(r) \geq_{\text{exc}} T_{\mathcal{K}(D),f}(r) - O(\log T_{\mathcal{A},f}(r) + \log r) .$$

*Proof.* Since  $X$  is a curve, the vector sheaf  $\Omega_{X/\mathbb{C}}(\log D)$  is isomorphic to the line sheaf  $\mathcal{K}(D)$ . Therefore the canonical projection  $p: \mathbb{P}(\Omega_{X/\mathbb{C}}(\log D)) \rightarrow X$  is an isomorphism,  $\mathcal{O}(1) \cong p^* \mathcal{K}(D)$ , and  $f' = p^{-1} \circ f$ . Thus

$$T_{\mathcal{O}(1),f'}(r) = T_{\mathcal{K}(D),f}(r) + O(1) ,$$

so (28.7.1) is equivalent to (28.6.1).  $\square$

*Remark 28.8.* In fact, when  $\dim X = 1$ , McQuillan’s inequality is directly equivalent to Conjecture 26.7, as can be seen from the above proof. This is not true in higher dimension, though (McQuillan’s inequality is proved, but Conjecture 26.7 is not).

Cartan’s theorem (Theorem 7.7) can also be proved using McQuillan’s inequality, but this requires more work than can be included here. See [Vojta, 2008]. The modified version (Theorem 7.13) 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 28.6.

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 [Siu and Yeung, 1997, 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 [Lu, 1991, § 4], [Vojta, 2000, Cor. 5.2], and [Vojta, 2001].

**Corollary 28.9.** *Let  $X$  be a smooth complex projective variety, let  $D$  be a normal crossings divisor on  $X$ , let  $f: B \rightarrow X$  be a holomorphic map, let  $\mathcal{A}$  be an ample line sheaf on  $X$ , let  $\mathcal{L}$  be a line sheaf on  $X$ , let  $d$  be a positive integer, and let  $\omega$  be a global section of  $S^d \Omega_{X/\mathbb{C}}(\log D)$ . If  $f^* \omega \neq 0$  (i.e., it does not vanish everywhere on  $B$ ), then*

$$\frac{1}{d} T_{\mathcal{L},f}(r) \leq_{\text{exc}} N_f^{(1)}(D, r) + N_{\text{Ram}(\pi)}(r) + O(\log T_{\mathcal{A},f}(r) + \log r).$$

*Proof.* Let  $f': B \rightarrow \mathbb{P}(\Omega_{X/\mathbb{C}}(\log D))$  be as in Theorem 28.6, and let

$$p: \mathbb{P}(\Omega_{X/\mathbb{C}}(\log D)) \rightarrow X$$

be the canonical projection. Then  $\omega$  corresponds to a global section

$$\omega' \in \Gamma(\mathbb{P}(\Omega_{X/\mathbb{C}}(\log D)), \mathcal{O}(d) \otimes p^* \mathcal{L}^\vee),$$

and  $(f')^* \omega' = f^* \omega$ . Thus the image of  $f'$  is not contained in the base locus of  $\mathcal{O}(d) \otimes \mathcal{L}^\vee$ , so

$$T_{\mathcal{O}(d),f'}(r) \geq T_{\mathcal{L},f}(r) + O(1)$$

by Theorem 11.8c. The result then follows immediately from (28.6.1). □

One can think of this result as a generalization of the fact that if  $f: C \rightarrow X$  is a nonconstant map from a nonsingular projective curve of genus  $g$  to a smooth complete variety  $X$ , then  $\deg f^* \mathcal{K}_X \leq 2g - 2$ , where  $\mathcal{K}_X$  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 Section 17). It is used in this manner in the proof of [Vojta, 2000, Thm. 5.3].

## 29 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 Section 27). 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 [Kato, 1989, 1.7]. Kato [1989, 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 [Grothendieck et al., 1971, VIII Thm. 1.1]. It is a vector sheaf (with non-obvious generators) by [Grothendieck and Dieudonné, 1960–1967, IV 2.5.2].

*Conjecture 29.1. (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  $\mathcal{A}$  be an ample line sheaf on  $X$ , and let  $\varepsilon > 0$ . Then, for all  $x \in X(\bar{k}) \setminus \text{Supp } D$  with  $[\kappa(x) : k] \leq r$ , there is a closed point  $x' \in \mathbb{P}(\Omega_{X/k}(\log D))$  lying over  $x$  such that

$$(29.1.1) \quad h_{\mathcal{O}(1),k}(x') \leq N_S^{(1)}(D, x) + d_k(x) + \varepsilon h_{\mathcal{A},k}(x) + O(1) .$$

Moreover, given a finite collection of rational maps  $g_i : X \dashrightarrow W_i$  to varieties  $W_i$ , there are finite sets  $\Sigma_i$  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  $k, S, X, D, r, \mathcal{A}, \varepsilon$ , 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 (29.1.1) 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 [Weyl and Weyl, 1938], further developed by L. Ahlfors [1941]. 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  $\wedge^p \mathbb{C}^{n+1}$  or  $\wedge^p k^{n+1}$ , respectively. This strongly suggested that the theory of successive minima may be related to the use of derivatives in number theory [Vojta, 1987, Ch. 6]. This has been recently refined [Vojta, 2008] to more explicitly involve a variant of Conjecture 29.1, 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  $Y = \text{Spec } \mathcal{O}_k$ , and let  $\mathcal{X} \rightarrow Y$  be a proper model for  $X$ . Then we have a relative tangent sheaf  $T_{\mathcal{X}/Y}(-\log D)$  on  $\mathcal{X}$ . This is not necessarily a vector sheaf, since  $\mathcal{X}$  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  $i: \text{Spec } Y \rightarrow \mathcal{X}$  is the section of  $\mathcal{X} \rightarrow Y$  corresponding to a given rational point, then  $i^*T_{\mathcal{X}/Y}(-\log D)$  is a vector sheaf on  $Y = \text{Spec } \mathcal{O}_k$ , 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  $i^*T_{\mathcal{X}/Y}(-\log D)$ , obeying certain upper bounds on its metric at each infinite place. See, for example, [Gillet and Soulé, 1991]. 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  $\mathcal{L}$  of  $i^*\Omega_{\mathcal{X}/Y}(\log D)$  of smallest degree. This, in turn, corresponds to giving a point  $x' \in \mathbb{P}(\Omega_{\mathcal{X}/Y}(\log D))$  lying over the rational point in question [Hartshorne, 1977, II 7.12]. The sheaf  $\mathcal{L}$  is none other than the pull-back of the tautological bundle  $\mathcal{O}(1)$  to  $Y$  via the section  $i': Y \rightarrow \mathbb{P}(\Omega_{\mathcal{X}/Y}(\log d))$  corresponding to the point  $x'$ . Therefore, the degree of  $\mathcal{L}$  is the height  $h_{\mathcal{O}(1),k}(x')$  (see (15.5)). This again leads to Conjecture 29.1.

We emphasize that specific bounds on the  $O(1)$  term in (29.1.1) are not given, so Conjecture 29.1 is meaningful only for an infinite set of points—or, better yet, for a generic or semi-generic set of points (Definitions 14.5 and 14.11).

*Remark 29.2.* The assertions of Remark 28.8 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 Tautolog-

ical Conjecture is equivalent to Conjecture 24.3b (for the same reasons as in Remark 28.8). Thus, by Remark 24.4, the Tautological Conjecture is proved for curves over function fields of characteristic 0. As in Remark 28.8, though, Conjectures 29.1 and 24.3b are not as closely related when  $\dim X > 1$ .

We also note that points of low height as in Conjecture 29.1 behave like derivatives in the following sense.

**Proposition 29.3.** (Arithmetic Chain Rule) *Let  $f: X_1 \rightarrow X_2$  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  $f_*: \mathbb{P}(\Omega_{X_1/k}) \dashrightarrow \mathbb{P}(\Omega_{X_2/k})$  takes  $x'$  to a point  $x'_2$  (lying over  $f(x)$ ) for which*

$$h_{\mathcal{O}(1),k}(x'_2) \leq h_{\mathcal{O}(1),k}(x') + O(1).$$

*Moreover, assume that  $X_1$  and  $X_2$  are projective, with ample line sheaves  $\mathcal{A}_1$  and  $\mathcal{A}_2$ , respectively, and that  $\varepsilon_1 > 0$  is a positive number for which (29.1.1) holds for all  $x$  in some set  $\Sigma \subseteq X_1(\bar{k})$  (with respect to  $\mathcal{A}_1$  and  $\varepsilon_1$ ). Then there is an  $\varepsilon_2 > 0$  such that (29.1.1) holds for  $f(x) \in X_2(\bar{k})$  for all  $x \in \Sigma$ , with respect to  $\mathcal{A}_2$  and  $\varepsilon_2$ .*

*Proof.* Let  $\mathcal{X}_1$  and  $\mathcal{X}_2$  be models for  $X_1$  and  $X_2$  over  $Y := \text{Spec } \mathcal{O}_k$ , respectively, chosen such that  $f$  extends as a morphism  $f: \mathcal{X}_1 \rightarrow \mathcal{X}_2$ , let  $Y' = \text{Spec } \mathcal{O}_{\kappa(x')}$ , and let  $i: Y' \rightarrow \mathcal{X}_1$  be the multisection corresponding to  $x$  (which factors through  $\text{Spec } \mathcal{O}_{\kappa(x)}$ ). Then  $x'$  corresponds to a surjection  $i^* \Omega_{\mathcal{X}_1/Y} \rightarrow \mathcal{L}$  for a line sheaf  $\mathcal{L}$  on  $Y'$ , and  $h_{\mathcal{O}(1),k}(x')$  is the Arakelov degree of  $\mathcal{L}$  divided by  $[\kappa(x') : k]$ . We also have a morphism  $f^* \Omega_{\mathcal{X}_2/Y} \rightarrow \Omega_{\mathcal{X}_1/Y}$ , isomorphic at  $x$ . This gives a nonzero map  $(f \circ i)^* \Omega_{\mathcal{X}_2/Y} \rightarrow \mathcal{L}$ , so  $h_{\mathcal{O}(1),k}(x'_2) \leq h_{\mathcal{O}(1),k}(x')$  (with heights defined using these models).

The second assertion is immediate from the first assertion, by Proposition 9.15.  $\square$

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.

### 30 Another Conjecture Implies abc

Conjecture 22.5, involving truncated counting functions, is of course a vast generalization of the abc conjecture, and Conjecture 24.1, which involved algebraic points, also rather easily implies abc. Actually, though, the (seemingly) weaker Conjecture 14.6 has also been shown to imply the abc conjecture [Vojta, 2000]. 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 30.1.** *For any  $\varepsilon > 0$  there is a nonsingular projective variety  $X_\varepsilon$  over  $\mathbb{Q}$ , a normal crossings divisor  $D_\varepsilon$  on  $X_\varepsilon$ , and a real number  $\varepsilon' > 0$ , such that if Conjecture 14.6b holds for  $X_\varepsilon$ ,  $D_\varepsilon$ , and  $\varepsilon'$ , then the abc conjecture (Conjecture 22.6) holds for  $\varepsilon$ .*

*Proof (sketch).* Fix an integer  $n > 3/\varepsilon + 3$ . Let  $X_n$  be the closed subvariety in  $(\mathbb{P}^2)^n$  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

$$(30.1.1) \quad ([x_1 : y_1 : z_1], \dots, [x_n : y_n : z_n]) \mapsto [\prod x_i^i : \prod y_i^i : \prod z_i^i].$$

Let  $\Gamma_n$  be the closure of the graph of this rational map in  $X_n \times \mathbb{P}^2$ , and let  $\phi : \Gamma_n \rightarrow \mathbb{P}^2$  be the projection to the second factor. The image of  $\phi$  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

$$x_n = \prod_p p^{\lfloor (\text{ord}_p a)/n \rfloor} \quad \text{and} \quad x_i = \prod_{\text{ord}_p a \equiv i \pmod n} p \quad (i < n).$$

(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  $\Gamma_n$  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  $\phi$ , 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

$$\lambda_{E,v}([a : b : c]) = \text{ord}_p(abc) \log p$$

and

$$\lambda_{D,v}(P_{a,b,c}) = \text{ord}_p(x_1 \cdots x_n y_1 \cdots y_n z_1 \cdots z_n) \log p,$$

using Weil functions suitably defined using (7.9.1). It then follows that

$$N_S(D, P_{a,b,c}) \leq \sum_{p|abc} \log p + \frac{1}{n} N_S(E, \phi(P_{a,b,c})).$$

One would like to apply Conjecture 14.6 to the divisor  $D$  on  $\Gamma_n$ , but this is not possible since  $\Gamma_n$  is singular. However, there is a nonsingular projective variety  $\Gamma'_n$ ,

a normal crossings divisor  $D'$  on  $\Gamma'_n$ , and a proper birational morphism  $\psi: \Gamma'_n \rightarrow \Gamma_n$ , such that  $\text{Supp } D' = \psi^{-1}(\text{Supp } D)$ ,  $\psi$  is an isomorphism over a suitably large set, and  $\mathcal{K}_{\Gamma'_n}(D') \geq \psi^* \phi^* \mathcal{O}(1)$  relative to the cone of effective divisors. For details see [Vojta, 2000, Lemma 3.9].

Let  $P'_{a,b,c}$  be the point on  $\Gamma'_n$  lying over  $P_{a,b,c}$ . Then one can show that

$$\begin{aligned} h_{\mathbb{Q}}([a : b : c]) &= h_{\psi^* \phi^* \mathcal{O}(1), \mathbb{Q}}(P'_{a,b,c}) + O(1) \\ &\leq h_{\mathcal{K}_{\Gamma'_n}(D'), \mathbb{Q}}(P'_{a,b,c}) + O(1) \\ &\leq N_S(D', P'_{a,b,c}) + \varepsilon' h_{\mathcal{A}, \mathbb{Q}}(P'_{a,b,c}) + O(1) \\ &\leq N_S(D, P_{a,b,c}) + \varepsilon' h_{\mathcal{A}, \mathbb{Q}}(P'_{a,b,c}) + O(1) \\ &\leq \sum_{p|abc} \log p + \frac{3}{n} h_{\mathbb{Q}}([a : b : c]) + \varepsilon'' h_{\mathbb{Q}}([a : b : c]) + O(1) \\ &\leq \sum_{p|abc} \log p + \frac{\varepsilon}{1 + \varepsilon} h_{\mathbb{Q}}([a : b : c]) + O(1), \end{aligned}$$

and therefore  $h_{\mathbb{Q}}([a : b : c]) \leq (1 + \varepsilon) \sum_{p|abc} \log p + O(1)$ . Here we have used the fact that  $N_S(D', P'_{a,b,c}) \leq N_S(D, P_{a,b,c})$ , which follows from the fact that  $\psi^* D - D'$  is effective.

This chain of inequalities holds outside of some proper Zariski-closed subset of  $\Gamma'_n$ , but it is possible to show that this set can be chosen so that it only involves finitely many triples  $(a, b, c)$ .  $\square$

The variety  $X_n$  admits a faithful action of  $\mathbb{G}_m^{2n-2}$ , 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 (30.1.1), so the action extends to  $\Gamma_n$ , and the construction of  $\Gamma'_n$  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  $\phi$ . It is this group action that allows one to control the proper Zariski-closed subset of  $\Gamma'_n$  arising out of Conjecture 14.6b. The group action also provides some additional structure, and in fact Conjecture 14.2 can be proved in this context [Vojta, 2000, Thm. 5.3].

### 31 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 curiosity, 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 [1991], who showed that ‘‘Mordell is as easy as abc,’’ i.e., the abc conjecture implies the Mordell

conjecture. This was extended by Bombieri [1994], who showed that abc implies Roth’s theorem, and then by van Frankenhuijsen [2002], who showed that the abc conjecture implies Conjecture 14.6b 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 14.6b. Here “the abc conjecture for  $k$ ” means Conjecture 22.5b over  $k$  with  $X = \mathbb{P}_k^1$  and  $D = [0] + [1] + [\infty]$ . It is further true that a “strong abc conjecture,” namely Conjecture 24.3b with  $X = \mathbb{P}_k^1$  and  $D = [0] + [1] + [\infty]$ , would imply Conjecture 24.1b 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 [1979]. He showed that a smooth complex projective curve  $X$  comes from a curve over  $\mathbb{Q}$  (i.e., there is a curve  $X_0$  over  $\mathbb{Q}$  such that  $X \cong X_0 \times_{\mathbb{Q}} \mathbb{C}$ ) if and only if there is a finite morphism from  $X$  to  $\mathbb{P}_{\mathbb{C}}^1$  ramified only over  $\{0, 1, \infty\}$ . For our purposes, this can be adapted as follows.

**Theorem 31.1.** (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  $f: X \rightarrow \mathbb{P}_k^1$  which is ramified only over  $\{0, 1, \infty\}$ , and such that  $S \subseteq f^{-1}(\{0, 1, \infty\})$ .*

*Proof.* See Belyĭ [1979] or Serre [1989]. □

The other ingredient is a complement (and actually, a converse) to Proposition 24.2, using truncated counting functions.

**Proposition 31.2.** *Let  $k$  be a number field or function field, let  $S \supseteq S_\infty$  be a finite set of places of  $k$ , let  $\pi: X' \rightarrow X$  be a surjective generically finite morphism of complete nonsingular varieties over  $k$ , and let  $D$  be a normal crossings divisor on  $X$ . Let  $D' = (\pi^*D)_{\text{red}}$ , and assume that it too has normal crossings. Assume also that the ramification divisor  $R$  of  $\pi$  satisfies*

$$(31.2.1) \quad R = \pi^*D - D'$$

*(and therefore that  $\pi$  is unramified outside of  $\text{Supp} D'$ ).<sup>4</sup> Let  $\mathcal{K}$  and  $\mathcal{K}'$  denote the canonical line sheaves on  $X$  and  $X'$ , respectively. Then, for all  $x \in X'(\bar{k})$  not lying on  $\text{Supp} D'$ ,*

$$(31.2.2) \quad \begin{aligned} N_S^{(1)}(D', x) + d_S(x) - h_{\mathcal{K}'(D'), k}(x) \\ \geq N_S^{(1)}(D, \pi(x)) + d_S(\pi(x)) - h_{\mathcal{K}(D), k}(\pi(x)) + O(1), \end{aligned}$$

*with equality if  $[\kappa(x) : k]$  is bounded. In particular, either part of Conjecture 24.3 for  $D'$  on  $X'$  is equivalent to that same part for  $D$  on  $X$ .*

*Proof.* First, by (31.2.1), we have  $\mathcal{K}'(D') \cong \pi^*(\mathcal{K}(D))$ . Thus

<sup>4</sup> This condition is equivalent to  $(X', D')$  being *log étale* over  $(X, D)$  by [Kato, 1989, (3.12)], using the fact that  $(X', D')$  is log smooth over  $\text{Spec} k$  by (3.7)(1) of *op. cit.*

$$h_{\mathcal{X}'(D'),k}(x) = h_{\mathcal{X}(D),k}(\pi(x)) + O(1),$$

so it suffices to show that

$$(31.2.3) \quad N_S^{(1)}(D', x) + d_S(x) \geq N_S^{(1)}(D, \pi(x)) + d_S(\pi(x)) + O(1),$$

with equality if  $[\kappa(x) : k]$  is bounded.

Now let  $Y_S = \text{Spec } \mathcal{O}_{k,S}$  (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  $\mathcal{X}$  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  $\mathcal{X}'$  be a proper model for  $X'$  over  $Y_S$ . We may assume that  $\pi$  extends to a morphism  $\pi: \mathcal{X}' \rightarrow \mathcal{X}$ , and that  $D'$  extends to an effective Cartier divisor on  $\mathcal{X}'$ . We assume further that  $\text{Supp } D' = \pi^{-1}(\text{Supp } D)$  (on  $\mathcal{X}'$ ). Indeed, one can obtain  $\mathcal{X}'$  and  $D'$  by blowing up the reduced sheaf of ideals corresponding to  $\pi^{-1}(\text{Supp } D)$ .

For  $x \in X'(\bar{k})$ , let  $L = \kappa(\pi(x))$ , let  $L' = \kappa(x)$ , and let  $Y_S^b$  and  $Y_S'$  be the normalizations of  $Y_S$  in  $L$  and  $L'$ , respectively. Then we have a commutative diagram

$$(31.2.4) \quad \begin{array}{ccc} Y_S' & \xrightarrow{i'} & \mathcal{X}' \\ \downarrow p & & \downarrow \pi \\ Y_S^b & \xrightarrow{i} & \mathcal{X} \end{array}$$

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  $Y_S^b$  lies in  $i^{-1}(\text{Supp } D)$ , and similarly for places  $w'$  of  $L'$ . By commutativity of (31.2.4) and the condition  $\text{Supp } D' = \pi^{-1}(\text{Supp } D)$ , 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  $(i')^*(\text{Supp } D')$ ,  $w$  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

$$d_S(x) - d_S(\pi(x)) = \frac{1}{[L' : k]} \sum_{Q'} \text{ord}_{Q'} \mathcal{D}_{L'/L} \cdot \log \#\mathbb{F}_{w'},$$

where the sum is over nonzero prime ideals  $Q'$  of  $\mathcal{O}_{L',S}$ , and  $w'$  is the corresponding place of  $L'$ . This sum can be restricted to primes corresponding to points in  $(i')^*(\text{Supp } D')$ , since the other primes are unramified over  $L$  by Lemma 23.14 ap-

plied to  $\pi: \mathcal{X}' \rightarrow \mathcal{X}$  and the relevant local rings. Therefore the inequality (31.2.3) follows from the elementary fact that

$$(31.2.5) \quad \text{ord}_Q \mathcal{D}_{L'/L} \geq e_{Q'/Q} - 1 ,$$

where  $Q = Q' \cap \mathcal{O}_{L,S}$ . (This inequality may be strict if  $Q'$  is wildly ramified over  $Q$ .)

Now if  $[L' : k]$  is bounded, then the differences in (31.2.5) add up to at most a bounded amount, so (31.2.3) holds up to  $O(1)$  in that case.

The last assertion of the proposition follows trivially from (31.2.2).  $\square$

The implications mentioned in the beginning of this section then follow immediately from Theorem 31.1 and Proposition 31.2, upon noting that (31.2.1) always holds for finite morphisms of nonsingular curves.

## References

### Ahlfors, Lars V.

- [1941] The theory of meromorphic curves. *Acta Soc. Sci. Fennicae. Nova Ser. A.*, 3(4):1–31, 1941.

### Artin, Emil

- [1967] *Algebraic numbers and algebraic functions*. Gordon and Breach Science Publishers, New York, 1967.

### Ashline, George Lawrence

- [1999] The defect relation of meromorphic maps on parabolic manifolds. *Mem. Amer. Math. Soc.*, 139(665):x+78, 1999. ISSN 0065-9266.

### Belyĭ, G. V.

- [1979] Galois extensions of a maximal cyclotomic field. *Izv. Akad. Nauk SSSR Ser. Mat.*, 43(2):267–276, 479, 1979. ISSN 0373-2436.

### Bloch, A.

- [1926] 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.

### Bombieri, Enrico

- [1994] Roth's theorem and the abc-conjecture. Preprint ETH Zürich, 1994.

### Borel, E.

- [1897] Sur les zéros des fonctions entières. *Acta Math.*, 20:357–396, 1897.

### Brownawell, W. D. and Masser, D. W.

- [1986] Vanishing sums in function fields. *Math. Proc. Cambridge Philos. Soc.*, 100(3):427–434, 1986. ISSN 0305-0041.

### Campana, Frédéric

- [2004] Orbifolds, special varieties and classification theory. *Ann. Inst. Fourier (Grenoble)*, 54(3):499–630, 2004. ISSN 0373-0956.

### Carlson, James and Griffiths, Phillip

- [1972] A defect relation for equidimensional holomorphic mappings between algebraic varieties. *Ann. of Math. (2)*, 95:557–584, 1972. ISSN 0003-486X.

### Cartan, Henri

- [1933] Sur les zéros des combinaisons linéaires de  $p$  fonctions holomorphes données. *Mathematica (Cluj)*, 7:5–29, 1933.

### Cherry, William and Ye, Zhuan

- [2001] *Nevanlinna's theory of value distribution*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2001. ISBN 3-540-66416-5.

### Corvaja, Pietro and Zannier, Umberto

- [2002] A subspace theorem approach to integral points on curves. *C. R. Math. Acad. Sci. Paris*, 334(4):267–271, 2002. ISSN 1631-073X.  
 [2004] On integral points on surfaces. *Ann. of Math. (2)*, 160(2):705–726, 2004. ISSN 0003-486X.

**Deligne, Pierre**

- [1970] *Équations différentielles à points singuliers réguliers*. Lecture Notes in Mathematics, Vol. 163. Springer-Verlag, Berlin, 1970.

**Dufresnoy, Jacques**

- [1944] Théorie nouvelle des familles complexes normales. Applications à l'étude des fonctions algébroides. *Ann. Sci. École Norm. Sup. (3)*, 61:1–44, 1944. ISSN 0012-9593.

**Eisenbud, David**

- [1995] *Commutative algebra (with a view toward algebraic geometry)*, volume 150 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995. ISBN 0-387-94268-8; 0-387-94269-6.

**Elkies, Noam D.**

- [1991] *ABC implies Mordell*. *Internat. Math. Res. Notices*, 1991(7):99–109, 1991. ISSN 1073-7928.

**Evertse, Jan-Hendrik**

- [1984] On sums of  $S$ -units and linear recurrences. *Compositio Math.*, 53(2):225–244, 1984. ISSN 0010-437X.

**Evertse, Jan-Hendrik and Ferretti, Roberto G.**

- [2008] A generalization of the Subspace Theorem with polynomials of higher degree. In *Diophantine approximation*, volume 16 of *Dev. Math.*, pages 175–198. SpringerWien-NewYork, Vienna, 2008. ArXiv:math.NT/0408381.

**Faltings, Gerd**

- [1983] Endlichkeitssätze für abelsche Varietäten über Zahlkörpern. *Invent. Math.*, 73(3):349–366, 1983. ISSN 0020-9910.
- [1986] Finiteness theorems for abelian varieties over number fields. In *Arithmetic geometry (Storrs, Conn., 1984)*, pages 9–27. Springer, New York, 1986. Translated from the German original [Invent. Math. **73** (1983), no. 3, 349–366; *ibid.* **75** (1984), no. 2, 381; MR 85g:11026ab] by Edward Shipz.
- [1991] Diophantine approximation on abelian varieties. *Ann. of Math. (2)*, 133(3):549–576, 1991. ISSN 0003-486X.
- [1994] The general case of S. Lang's conjecture. In *Barsotti Symposium in Algebraic Geometry (Abano Terme, 1991)*, volume 15 of *Perspect. Math.*, pages 175–182. Academic Press, San Diego, CA, 1994.

**Ferretti, Roberto G.**

- [2000] Mumford's degree of contact and Diophantine approximations. *Compositio Math.*, 121(3):247–262, 2000. ISSN 0010-437X.

**Fujimoto, Hirotaka**

- [1972] Extensions of the big Picard's theorem. *Tôhoku Math. J. (2)*, 24:415–422, 1972. ISSN 0040-8735.

**Gasbarri, C.**

- [2008] The strong abc conjecture over function fields [after McQuillan and Yamañoi]. To appear, Séminaire Bourbaki, 60e année (2007/08); arXiv:0811.3153, 2008.

**Gillet, H. and Soulé, C.**

- [1991] 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.

**Goldberg, Anatoly A. and Ostrovskii, Iossif V.**

- [2008] *Value distribution of meromorphic functions*, volume 236 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, 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.

**Green, Mark**

- [1972] Holomorphic maps into complex projective space omitting hyperplanes. *Trans. Amer. Math. Soc.*, 169:89–103, 1972. ISSN 0002-9947.
- [1975] Some Picard theorems for holomorphic maps to algebraic varieties. *Amer. J. Math.*, 97:43–75, 1975. ISSN 0002-9327.

**Green, Mark and Griffiths, Phillip**

- [1980] Two applications of algebraic geometry to entire holomorphic mappings. In *The Chern Symposium 1979 (Proc. Internat. Sympos., Berkeley, Calif., 1979)*, pages 41–74. Springer, New York, 1980.

**Grothendieck, A. et al.**

- [1971] *Revêtements étales et groupe fondamental*. Springer-Verlag, 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.

**Grothendieck, Alexander and Dieudonné, Jean-Alexandre-Eugène**

- [1960–1967] Éléments de géométrie algébrique. *Publ. Math. IHES*, 4, 8, 11, 17, 20, 24, 28, and 32, 1960–1967. ISSN 0073-8301.

**Gunning, Robert C.**

- [1990] *Introduction to holomorphic functions of several variables. Vol. II (Local theory)*. The Wadsworth & Brooks/Cole Mathematics Series. Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, CA, 1990. ISBN 0-534-13309-6.

**Hartshorne, Robin**

- [1977] *Algebraic geometry*. Springer-Verlag, New York, 1977. ISBN 0-387-90244-9. Graduate Texts in Mathematics, No. 52.

**Hayman, W. K.**

- [1964] *Meromorphic functions*. Oxford Mathematical Monographs. Clarendon Press, Oxford, 1964.

**Hodge, W. V. D. and Pedoe, D.**

- [1952] *Methods of algebraic geometry. Vol. II. Book III: General theory of algebraic varieties in projective space. Book IV: Quadrics and Grassmann varieties*. Cambridge, at the University Press, 1952.

**Iitaka, Shigeru**

- [1982] *Algebraic geometry, An introduction to birational geometry of algebraic varieties*, volume 76 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1982. ISBN 0-387-90546-4. North-Holland Mathematical Library, 24.

**Kato, Kazuya**

- [1989] Logarithmic structures of Fontaine-Illusie. In *Algebraic analysis, geometry, and number theory (Baltimore, MD, 1988)*, pages 191–224. Johns Hopkins Univ. Press, Baltimore, MD, 1989.



**Kawamata, Yujiro**

- [1980] On Bloch's conjecture. *Invent. Math.*, 57(1):97–100, 1980. ISSN 0020-9910.

**Koblitz, Neal**

- [1984] *p-adic numbers, p-adic analysis, and zeta-functions*, volume 58 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1984. ISBN 0-387-96017-1.

**Kunz, Ernst**

- [1986] *Kähler differentials*. Advanced Lectures in Mathematics. Friedr. Vieweg & Sohn, Braunschweig, 1986. ISBN 3-528-08973-3.

**Lang, Serge**

- [1960] Integral points on curves. *Inst. Hautes Études Sci. Publ. Math.*, 6:27–43, 1960. ISSN 0073-8301.
- [1970] *Algebraic number theory*. Addison-Wesley Publishing Co., Inc., Reading, Mass.-London-Don Mills, Ont., 1970.
- [1983] *Fundamentals of Diophantine geometry*. Springer-Verlag, New York, 1983. ISBN 0-387-90837-4.
- [1986] Hyperbolic and Diophantine analysis. *Bull. Amer. Math. Soc. (N.S.)*, 14(2):159–205, 1986. ISSN 0273-0979.
- [1991] *Number theory. III: Diophantine geometry*, volume 60 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 1991. ISBN 3-540-53004-5.

**Lang, Serge and Cherry, William**

- [1990] *Topics in Nevanlinna theory*, volume 1433 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1990. ISBN 3-540-52785-0. With an appendix by Zhuan Ye.

**Levin, Aaron**

- [2008] The dimensions of integral points and holomorphic curves on the complements of hyperplanes. *Acta Arith.*, 134(3):259–270, 2008. ISSN 0065-1036.
- [2009] Generalizations of Siegel's and Picard's theorems. *Ann. of Math. (2)*, 170(2):609–655, 2009. ISSN 0003-486X.

**Levin, Aaron, McKinnon, David, and Winkelmann, Jörg**

- [2008] 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. URL <http://dx.doi.org/10.1093/qmath/ham052>.

**Lu, Steven Shin-Yi**

- [1991] On meromorphic maps into varieties of log-general type. In *Several complex variables and complex geometry, Part 2 (Santa Cruz, CA, 1989)*, volume 52 of *Proc. Sympos. Pure Math.*, pages 305–333. Amer. Math. Soc., Providence, RI, 1991.

**Mahler, K.**

- [1949] On a theorem of Liouville in fields of positive characteristic. *Canadian J. Math.*, 1: 397–400, 1949. ISSN 0008-414X.

**McQuillan, Michael**

- [1998] Diophantine approximations and foliations. *Inst. Hautes Études Sci. Publ. Math.*, 87: 121–174, 1998. ISSN 0073-8301.
- [2000] Non-commutative Mori theory. IHES preprint IHES/M/00/15, 2000.
- [2008] Canonical models of foliations. *Pure Appl. Math. Q.*, 4(3, part 2):877–1012, 2008. ISSN 1558-8599.
- [2009] Old and new techniques in function field arithmetic. submitted, 2009.

**Neukirch, Jürgen**

- [1999] *Algebraic number theory*, volume 322 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, 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.

**Nevanlinna, Rolf**

- [1970] *Analytic functions*. Translated from the second German edition by Phillip Emig. Die Grundlehren der mathematischen Wissenschaften, Band 162. Springer-Verlag, New York, 1970.

**Noguchi, Junjiro**

- [1981] Lemma on logarithmic derivatives and holomorphic curves in algebraic varieties. *Nagoya Math. J.*, 83:213–233, 1981. ISSN 0027-7630.
- [1996] On Nevanlinna's second main theorem. In *Geometric complex analysis (Hayama, 1995)*, pages 489–503. World Sci. Publ., River Edge, NJ, 1996.
- [1998] On holomorphic curves in semi-abelian varieties. *Math. Z.*, 228(4):713–721, 1998. ISSN 0025-5874.

**Osgood, Charles F.**

- [1975] 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.
- [1981] A number theoretic-differential equations approach to generalizing Nevanlinna theory. *Indian J. Math.*, 23(1-3):1–15, 1981. ISSN 0019-5324.
- [1985] Sometimes effective Thue-Siegel-Roth-Schmidt-Nevanlinna bounds, or better. *J. Number Theory*, 21(3):347–389, 1985. ISSN 0022-314X.

**Roth, K. F.**

- [1955] Rational approximations to algebraic numbers. *Mathematika*, 2:1–20; corrigendum, 168, 1955. ISSN 0025-5793.

**Ru, Min**

- [1997] On a general form of the second main theorem. *Trans. Amer. Math. Soc.*, 349(12): 5093–5105, 1997. ISSN 0002-9947.
- [2001] *Nevanlinna theory and its relation to Diophantine approximation*. World Scientific Publishing Co. Inc., River Edge, NJ, 2001. ISBN 981-02-4402-9.
- [2009] Holomorphic curves into algebraic varieties. *Ann. of Math. (2)*, 169(1):255–267, 2009. ISSN 0003-486X. doi: 10.4007/annals.2009.169.255. URL <http://dx.doi.org/10.4007/annals.2009.169.255>.

**Schmidt, Wolfgang M.**

- [1991] *Diophantine approximations and Diophantine equations*, volume 1467 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1991. ISBN 3-540-54058-X.

**Serre, Jean-Pierre**

- [1955–1956] Géométrie algébrique et géométrie analytique. *Ann. Inst. Fourier, Grenoble*, 6:1–42, 1955–1956. ISSN 0373-0956.
- [1968] *Corps locaux*. Hermann, Paris, 1968. Deuxième édition, Publications de l'Université de Nancago, No. VIII.
- [1989] *Lectures on the Mordell-Weil theorem*. Aspects of Mathematics, E15. Friedr. Vieweg & Sohn, Braunschweig, 1989. ISBN 3-528-08968-7. Translated from the French and edited by Martin Brown from notes by Michel Waldschmidt.

**Shabat, B. V.**

- [1985] *Distribution of values of holomorphic mappings*, volume 61 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1985. ISBN 0-8218-4514-4. Translated from the Russian by J. R. King, Translation edited by Lev J. Leifman.

**Shiffman, Bernard**

- [1983] Introduction to the Carlson-Griffiths equidistribution theory. In *Value distribution theory (Joensuu, 1981)*, volume 981 of *Lecture Notes in Math.*, pages 44–89. Springer, Berlin, 1983.

**Silverman, Joseph H.**

- [1986] The theory of height functions. In *Arithmetic geometry (Storrs, Conn., 1984)*, pages 151–166. Springer, New York, 1986.

**Siu, Yum-Tong**

- [1995] Hyperbolicity problems in function theory. In *Five decades as a mathematician and educator*, pages 409–513. World Sci. Publ., River Edge, NJ, 1995.

**Siu, Yum-Tong and Yeung, Sai-Kee**

- [1996] 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.
- [1997] 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.

**Stewart, C. L. and Tijdeman, R.**

- [1986] On the Oesterlé-Masser conjecture. *Monatsh. Math.*, 102(3):251–257, 1986. ISSN 0026-9255.

**Stoll, Wilhelm**

- [1970] *Value distribution of holomorphic maps into compact complex manifolds*. Lecture Notes in Mathematics, Vol. 135. Springer-Verlag, Berlin, 1970.

**Szpiro, L., Ullmo, E., and Zhang, S.**

- [1997] Équirépartition des petits points. *Invent. Math.*, 127(2):337–347, 1997. ISSN 0020-9910.

**van der Poorten, A. J. and Schlickewei, H.P.**

- [1982] The growth conditions for recurrence sequences. *Macquarie Math. Reports*, 82-0041, 1982.

**van Frankenhuijsen, Machiel**

- [2002] The ABC conjecture implies Vojta's height inequality for curves. *J. Number Theory*, 95(2):289–302, 2002. ISSN 0022-314X.
- [2007] ABC implies the radicalized Vojta height inequality for curves. *J. Number Theory*, 127(2):292–300, 2007. ISSN 0022-314X.

**Vojta, Paul**

- [1987] *Diophantine approximations and value distribution theory*, volume 1239 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1987. ISBN 3-540-17551-2.
- [1989] A refinement of Schmidt's subspace theorem. *Amer. J. Math.*, 111(3):489–518, 1989. ISSN 0002-9327.
- [1996] Integral points on subvarieties of semiabelian varieties. I. *Invent. Math.*, 126(1):133–181, 1996. ISSN 0020-9910.
- [1997] On Cartan's theorem and Cartan's conjecture. *Amer. J. Math.*, 119(1):1–17, 1997. ISSN 0002-9327.

- [1998] A more general *abc* conjecture. *Internat. Math. Res. Notices*, 1998(21):1103–1116, 1998. ISSN 1073-7928.
- [1999] Integral points on subvarieties of semiabelian varieties. II. *Amer. J. Math.*, 121(2):283–313, 1999. ISSN 0002-9327.
- [2000] On the *ABC* conjecture and Diophantine approximation by rational points. *Amer. J. Math.*, 122(4):843–872, 2000. ISSN 0002-9327.
- [2001] Correction to: “On the *ABC* conjecture and Diophantine approximation by rational points” [*Amer. J. Math.* **122** (2000), no. 4, 843–872; MR1771576 (2001i:11094)]. *Amer. J. Math.*, 123(2):383–384, 2001. ISSN 0002-9327.
- [2007a] Jets via Hasse-Schmidt derivations. In *Diophantine geometry*, volume 4 of *CRM Series*, pages 335–361. Ed. Norm., Pisa, 2007. arXiv:math.AG/0407113.
- [2007b] Nagata’s embedding theorem. To appear; arXiv:0706.1907, 2007.
- [2008] On McQuillan’s “tautological inequality” and the Weyl-Ahlfors theory of associated curves. To appear; arXiv:0706.3044, 2008.

**Weil, André**

- [1928] L’arithmétique sur les courbes algébriques. *Acta Math.*, 52:281–315, 1928.
- [1951] Arithmetic on algebraic varieties. *Ann. of Math. (2)*, 53:412–444, 1951. ISSN 0003-486X.

**Weyl, Hermann and Weyl, Joachim**

- [1938] Meromorphic curves. *Ann. of Math. (2)*, 39(3):516–538, 1938. ISSN 0003-486X.

**Wong, Pit-Mann**

- [1989] On the second main theorem of Nevanlinna theory. *Amer. J. Math.*, 111(4):549–583, 1989. ISSN 0002-9327.

**Yamanoi, Katsutoshi**

- [2004a] Algebro-geometric version of Nevanlinna’s lemma on logarithmic derivative and applications. *Nagoya Math. J.*, 173:23–63, 2004. ISSN 0027-7630.
- [2004b] The second main theorem for small functions and related problems. *Acta Math.*, 192(2):225–294, 2004. ISSN 0001-5962.

# Index

## A

abc conjecture 115, 126, 132, 187–192, 202, 216–218  
for  $k$  219  
strong 219  
abelian variety 114, 159, 161  
Ahlfors, L. V. 128  
[1941] 215  
algebraic number 130, 132  
algebraic point 115, 153, 155, 168, 200, 216, 219  
analytic divisor *see* divisor, analytic  
analytic variety 199  
Arakelov degree 216  
Arakelov theory 140, 144, 165, 195, 215  
Arithmetic Chain Rule 216  
Artin, E.  
[1967] 163  
Ashline, G. L.  
[1999] 208  
associated curves 215  
auxiliary polynomial 128

## B

$B[r]$  203  
 $B(r)$  203  
 $B\langle r \rangle$  203  
ball quotient 206  
Belyĭ, G. V.  
[1979] 219  
Berkovich space 140  
Bloch, A. 114  
[1926] 159  
Bombieri, E.  
[1994] 219

Borel lemma 156  
Borel, A.  
[1897] 156  
Brownawell, W. D. and Masser, D. W.  
[1986] 191  
Buium, A. 114

## C

Campana, F.  
[2004] 167  
Carlson, J. and Griffiths, P.  
[1972] 158  
Cartan's theorem 115, 133–137, 159, 168, 213  
Cartan, H.  
[1933] 133  
Cartier divisor 138  
characteristic function 124  
Cherry, W. and Lang, S.  
[1990] 120, 128, 203, 205  
Cherry, W. and Ye, Z.  
[2001] 128  
Chevalley-Weil theorem 156, 172, 194, 196  
generalized 197  
Chow form 180  
Chow weight 180  
Chow's lemma 146, 147, 160  
closed immersion 216  
closure of the graph 155, 164, 167, 174, 197, 211  
compactness argument 210  
complex analytic space 150  
complexity 118, 124  
conic 169  
conjecture  
 $1 + \varepsilon$  202, 214

- abc *see* abc conjecture
  - for algebraic points 200
  - for algebraic points with truncated counting functions 201
  - for rational points 160
  - Griffiths' 115, 158, 205
  - tautological 115, 214–216
  - with truncated counting functions 189
  - Corvaja, P. 115, 180
  - Corvaja, P. and Zannier, U.
    - [2002] 171
    - [2004] 172
  - countable 162
  - counting function 123, 126, 130, 131, 134, 135, 148, 150, 155, 195, 204
    - of algebraic points 129
    - properties 148, 150, 204
    - ramification 132, 204, 210
    - relative to an ideal 195
    - truncated 187–192, 201, 202, 205, 212, 216, 219
  - covering construction 201, 202
  - covering spaces 128
  - curvature 128, 208, 214
  - curve 137
  - $\mathbb{C}_v$  139
- D**
- $d()$  *see* discriminant
  - de Franchis 206
  - Dedekind ring 163, 199
  - defect 129, 131
  - Deligne, P.
    - [1970] 209
  - derivative 114, 115, 128, 207–216
  - descent 214
  - dictionary 130, 131
  - Dieudonné, J. and Grothendieck, A.
    - [1960–1967] 214
  - different 192, 221
    - multiplicativity in towers 192, 198, 204
  - differential geometry 208
  - differentials 196
  - diophantine equation 120
  - discriminant 192–200
    - logarithmic 192, 206
  - divisor
    - analytic 133, 198, 200, 204
    - degree of 163
    - large 173, 174, 177–179
    - nef 179
    - normal crossings 158, 159, 167, 168, 182, 189, 208, 214, 215, 217
    - reduced 158, 201
    - strict normal crossings 159, 209, 210
    - very large 172–174, 176–179
  - domain 132
  - Dufresnoy, J.
    - [1944] 157
- E**
- EGA 119, 137, 210
  - Eisenbud, D.
    - [1995] 194
  - Elkies, N. D.
    - [1991] 218
  - Embeddings 168–171
  - error term 128, 136
  - étale 156, 216
  - étale cover 172
  - étale topology 159, 214
  - Euler characteristic 126
  - Evertse, J.-H. 115
    - [1984] 157
  - Evertse, J.-H. and Ferretti, R. G.
    - [2008] 180–184, 186
  - Exc() *see* exceptional set
  - $\leq_{\text{exc}}$  125
  - exceptional divisor 211
  - exceptional set 115, 137, 161, 162, 166–168, 188, 191–192, 210, 218
    - diophantine 166
    - holomorphic 166
  - exponential function 122, 124
- F**
- Faltings, G. 113, 167
    - [1983] 160
    - [1986] 160
    - [1991] 161
    - [1994] 161
  - Ferretti, R. G. 115
    - [2000] 180
  - Ferretti, R. G. and Evertse, J.-H.
    - [2008] 180–184, 186
  - fibration 167
  - finite place 116
  - finite ramified covering 115, 132, 168, 188, 198, 203–206, 212
  - First Main Theorem 122, 124, 127, 129, 134, 148, 151
  - Fitting ideal 194, 195, 200
  - flag variety 215
  - flat 174
  - fractional ideal 116

- Frenet formalism 215
- Fujimoto, H.  
[1972] 157
- function field 114, 115, 130, 155, 162–165,  
167, 168, 200, 215, 220  
split case 168, 206, 213
- G**
- GAGA 210
- Gasbarri, C.  
[2008] 202
- general position 115, 133, 159, 170
- general type 159, 161, 162, 166, 167
- generic subset 160, 162, 189, 200, 201, 215
- Z-generic subset 162
- generically finite ramified cover 201
- geometrically integral 137
- germ of holomorphic functions 199
- Gillet, H. and Soulé, C.  
[1991] 215
- global field 163
- Goldberg, Anatoly A. and Ostrovskii, Iossif V.  
[2008] 114, 152
- Grauert, H. 167
- Green, M. L.  
[1972] 157  
[1975] 169
- Green, M. L. and Griffiths, P.  
[1980] 159
- Griffiths, P. A. 158
- Griffiths, P. and Carlson, J.  
[1972] 158
- Griffiths, P. and Green, M. L.  
[1980] 159
- Grothendieck, A. and Dieudonné, J.  
[1960–1967] 214
- Grothendieck, A., et al.  
[1971] 214
- group action 218
- Gunning, R. C.  
[1990] 199
- H**
- $h()$  *see* height
- Hartshorne, R.  
[1977] 114, 137, 150, 174, 178, 215
- Hayman, W. K.  
[1964] 124
- height 114, 116–118, 123, 129, 131, 134,  
135, 143, 144, 151, 165, 176, 195, 205  
absolute 119  
Ahlfors-Shimizu 152
- big 147, 153, 160, 200  
defined using models 165  
logarithmic 118  
multiplicative 118  
properties of 128, 144, 152  
twisted 182
- Hermite-Minkowski theorem 156
- Hilbert function 171, 180
- Hilbert weight 181
- $H_k()$  118
- Hodge, W. V. D. and Pedoe, D.  
[1952] 180
- holomorphic curve 114, 133, 167, 177, 203
- hyperplane 115, 133–137, 176, 182
- I**
- Iitaka, S.  
[1982] 167
- integral point 153, 156, 165, 167, 169, 171,  
174
- integral set 155, 168
- $D$ -integral set 154
- intersection number 165
- J**
- Jacobian 172
- Jensen's formula 124, 131
- jet differential 213
- K**
- Kähler different 197
- Kato, K.  
[1989] 214, 219
- Kawamata locus 166
- Kawamata structure theorem 166
- Kawamata, Y.  
[1980] 159, 166
- Kim, Minhyong 114
- Kobayashi hyperbolic 114
- Kobayashi, R. 115, 208
- Koblitz, N.  
[1984] 139
- Kodaira dimension 167
- Kodaira's lemma 146
- Kunz, E.  
[1986] 197
- L**
- $\lambda_H$  133, 134
- Lang, S. 113, 114

- [1960] 160
  - [1970] 114
  - [1983] 114, 116, 138, 140, 141, 143, 206
  - [1986] 166
  - [1991] 120, 162, 166
  - Lang, S. and Cherry, W.
    - [1990] 120, 128, 203, 205
  - lattice 215
  - Laurent expansion 124
  - Lemma on the Logarithmic Derivative 115, 128, 208, 214
    - geometric 209, 210
  - Levin, A.
    - [2008] 158
    - [2009] 166, 172, 176, 177, 179
  - Levin, A., McKinnon, D., and Winkelmann, J.
    - [2008] 161
  - lies over 117
  - lies under 117
  - line sheaf 138
    - ample 145, 152
    - big 146, 152, 159, 201
    - canonical 158
    - tautological 210
  - Liouville's theorem 122
  - locally  $M_k$ -bounded function 140
  - $\log^+$  119, 123, 127
  - log étale 219
  - log cotangent sheaf 208, 214
  - log general type 159, 161, 166, 168
  - log tangent sheaf 209
  - Lu, S.
    - [1991] 213
- M**
- $m()$  *see* proximity function
  - Mahler, K.
    - [1949] 162
  - Manin, Yu. 167
  - Masser, D. *see* abc conjecture
  - Masser, D. W. and Brownawell, W. D.
    - [1986] 191
  - max-min 141
  - McKinnon, D., Winkelmann, J., and Levin, A.
    - [2008] 161
  - McQuillan, M. 115, 207, 208, 214
    - [1998] 210
    - [2000] 206
    - [2008] 202
    - [2009] 202
  - measure 204
  - meromorphic function 124, 130
  - metric 138, 215
  - Minkowski, H. 215
  - $M_k$  116
  - $M_k$ -constant 140, 154, 164
  - model 155, 163, 164, 195, 215, 220
  - Mordell conjecture 113, 160, 161, 167, 168
    - as easy as abc 218
  - Mordellic 114, 179
  - $M_{\mathbb{Q}}$  117
  - multisection 216
  - Mumford, D. 181
- N**
- $N()$  *see* counting function
  - $\mathbb{N}$  116
  - Néron, A. 138
  - Nagata's embedding theorem 140, 163, 195
  - Neukirch, J.
    - [1999] 117
  - Nevanlinna theory 113, 126–130
  - Nevanlinna, R. 215
    - [1970] 114, 124, 208
  - Noguchi, J.
    - [1981] 159, 166
    - [1996] 190
    - [1998] 159
  - norm 116, 131
  - normalization 220
  - Northcott's finiteness theorem 119, 145, 165
  - number field 116, 190
  - number theory 130
  - $N_v$  116
- O**
- Oesterlé, J. *see* abc conjecture
  - $\text{ord}^+$  123
  - $\text{ord}_p()$  116, 131
  - Osgood, C. F.
    - [1975] 162
    - [1981] 113, 130
    - [1985] 130
  - Ostrovskii, Iossif V. and Goldberg, Anatoly A.
    - [2008] 114, 152
- P**
- partitioning argument 183, 187
  - Pedoe, D. and Hodge, W. V. D.
    - [1952] 180
  - Picard's "little" theorem 125
  - Picard's theorem 113, 122
  - place 116, 117, 120, 130, 131
    - archimedean 116, 195, 215



- complex 116
  - infinite 116
  - non-archimedean 116, 130, 189
  - real 116
  - positive characteristic 162
  - Product Formula 117–119, 131, 135, 144, 163
  - projection formula 178
  - proximity function 123, 125, 126, 130, 131, 133, 135, 148, 150, 195, 204, 211
  - of algebraic points 129
  - properties 148, 150, 204
- Q**
- quasi-Mordellic 179
- R**
- ramification divisor 194, 195, 198, 200, 201, 204, 219
  - rational function 124
  - rational point 115
  - red *see* divisor, reduced
  - regular field extension 207
  - resolution of singularities 160
  - Riemann surface 132, 199, 212
  - Riemann-Hurwitz formula 207
  - Riemann-Roch 215
  - ring of integers 116
  - root of unity 119, 200
  - Roth's theorem 113, 114, 120–122, 125–132, 135, 160, 162, 171, 219
  - Roth, K. F. 215
    - [1955] 120
  - Ru, M. 115
    - [1997] 136
    - [2001] 114, 124
    - [2009] 180, 182, 185
- S**
- $S_\infty$  116
  - Schlickewei, H.P. and van der Poorten, A. J.
    - [1982] 157
  - Schmidt's Subspace Theorem 114, 115, 133–137, 160, 168, 169, 171, 175, 180, 184, 191, 210, 215
  - Schmidt, W. M.
    - [1991] 136
  - Schwarz lemma 213
  - second exact sequence for differentials 197
  - Second Main Theorem 125, 126, 128–130, 132–134, 148, 212
  - section 164
  - semi-generic subset 162
  - semiabelian variety 115, 159, 161, 166
  - Serre, J.-P.
    - [1955–1956] 210
    - [1968] 193
    - [1989] 153, 156, 219
  - Shabat, B. V.
    - [1985] 114
  - Shiffman, B. 180
    - [1983] 208
  - Siegel's theorem 171
  - Silverman, J. H.
    - [1986] 143
  - Siu, Y.-T.
    - [1995] 159
  - Siu, Y.-T. and Yeung, S.-K.
    - [1996] 159
    - [1997] 213
  - size 135
  - Soulé, C. and Gillet, H.
    - [1991] 215
  - Stewart, C. L. and Tijdeman, R.
    - [1986] 189
  - Stoll, W.
    - [1970] 158
  - strictly rational map 167
  - successive minima 114, 182, 215
  - Szpiro, L. 167
  - Szpiro, L., Ullmo, E., and Zhang, S.
    - [1997] 160
- T**
- $T()$  *see* height
  - tangent bundle
    - absolute 167, 207, 214
    - relative 214
  - tangent sheaf
    - relative 215
  - Tate, J. 167
  - tautological bundle 215
  - tautological conjecture *see* conjecture, tautological
  - tautological inequality 115, 210–214
  - Thue equation 120
  - Tijdeman, R. and Stewart, C. L.
    - [1986] 189
  - $v$ -topology 139
  - truncated counting function 115
- U**
- Ullmo, E., Zhang, S., and Szpiro, L.

- [1997] 160
  - unit 119, 156, 169
  - unit disk 206
  - unit theorem 157, 169
  - $m$ -uple embedding 169, 186
- V**
- valuative criterion of properness 195
  - value distribution theory *see* Nevanlinna theory
  - van der Poorten, A. J. and Schlickewei, H.P. [1982] 157
  - van Frankenhuijsen, M. [2002] 219 [2007] 189
  - variety 137
    - complete 146, 147
  - very generic 162
  - Vojta, P. [1987] 113, 114, 116, 130, 135, 146, 156, 169, 197, 201, 202, 215 [1989] 137 [1996] 161 [1997] 136, 137, 159 [1998] 202 [1999] 161 [2000] 213, 216, 218 [2001] 213 [2007a] 159 [2007b] 164 [2008] 213, 215
- W**
- Weierstrass polynomial 199
  - Weierstrass Preparation Theorem 199
  - Weil function 138, 140, 154, 157, 165, 174, 176, 177, 182, 183, 185, 187, 189, 195, 217
  - defined using model 164
  - of algebraic points 142
  - properties 140
  - Weil, A. 116, 118 [1928] 138 [1951] 138
  - Weyl, H. and J. [1938] 215
  - Winkelmann, J., Levin, A., and McKinnon, D. [2008] 161
  - Wong, P.-M. 115, 208, 209 [1989] 128
- X**
- $X(M_k)$  139
- Y**
- Yamanoi, K. [2004a] 196 [2004b] 202, 203
  - Ye, Z. and Cherry, W. [2001] 128
  - Yeung, S.-K. and Siu, Y.-T. [1996] 159 [1997] 213
- Z**
- Zannier, U. 115, 180
  - Zannier, U. and Corvaja, P. [2002] 171 [2004] 172
  - Zariski cotangent space 159, 210
  - Zhang, S., Szpiro, L., and Ullmo, E. [1997] 160