

SESHADRI CONSTANTS, DIOPHANTINE APPROXIMATION, AND ROTH'S THEOREM FOR ARBITRARY VARIETIES

DAVID MCKINNON AND MIKE ROTH

ABSTRACT. In this paper, we associate an invariant $\alpha_x(L)$ to an algebraic point x on an algebraic variety X with an ample line bundle L . The invariant α measures how well x can be approximated by rational points on X , with respect to the height function associated to L . We show that this invariant is closely related to the Seshadri constant $\epsilon_x(L)$ measuring local positivity of L at x , and in particular that Roth's theorem on \mathbb{P}^1 generalizes as an inequality between these two invariants valid for arbitrary projective varieties.

1. INTRODUCTION

Let k be a number field, and X an irreducible projective variety over $\text{Spec}(k)$. The Bombieri-Lang conjecture predicts that if X is of general type then the k -points of X are contained in a proper closed subset of X . We view this as a statement that a global fact about the canonical bundle of X (that it is “generically positive”, where positivity is used in a broad sense) implies a global fact about the accumulation of rational points. Following a well-established principle in geometry one should study the local influence of positivity on the local accumulation of rational points. To do this we need local measures of both these phenomena.

Let L be an ample line bundle on X , and x a point of $X(\bar{k})$. By slightly modifying the usual definition of approximation exponent on \mathbb{P}^1 (and inspired by a definition from [14] by the first author) we define a new invariant $\alpha_x(L) \in (0, \infty]$ which measures how quickly rational points accumulate around x , from the point of view of the line bundle L and a fixed place v of k .

The central theme of this paper is the interrelations between $\alpha_x(L)$ and the Seshadri constant $\epsilon_x(L)$, an invariant defined by Demailly [5] which measures local positivity of a line bundle L near a point x . The two share common formal properties, and this similarity is even more evident when α_x is interpreted through Arakelov theory. Moreover, the classic approximation results on \mathbb{P}^1 — the theorems of Liouville and Roth — generalize as inequalities between α_x and ϵ_x valid for arbitrary projective varieties. This general version of Roth's theorem admits further generalizations to simultaneous approximation and improvements via étale covers.

In order to motivate our results we first quickly review approximation on the line, and to simplify this part of the discussion we assume that $k = \mathbb{Q}$ and that the place is archimedean.

Date: Revised on May 8, 2014.

2010 *Mathematics Subject Classification.* Primary 14G05; Secondary 14G40.

David McKinnon was partially supported by an NSERC research grant.

Mike Roth was partially supported by an NSERC research grant.

Approximation on \mathbb{A}^1 . For a point $x \in \mathbb{R}$ the *approximation exponent* of x is defined as the smallest real number τ_x such that for any $\delta > 0$ the inequality

$$\left| x - \frac{a}{b} \right| \leq \frac{1}{b^{\tau_x + \delta}}$$

has only finitely many solutions $a/b \in \mathbb{Q}$ (see [8, Part D]). The approximation exponent measures a certain tension between our ability to closely approximate x by rational numbers (the $|x - a/b|$ term) and the complexity (the $1/b$ term) of the rational number needed to make this approximation.

If $x \in \mathbb{Q}$ then it is easy to see that $\tau_x = 1$. In 1842 Dirichlet proved his famous approximation theorem: if $x \in \mathbb{R} \setminus \mathbb{Q}$ then $\tau_x \geq 2$. One therefore seeks upper bounds on τ . In 1844 Liouville showed that if $x \in \mathbb{R}$ is algebraic of degree d over \mathbb{Q} then $\tau_x \leq d$, and used this to give concrete examples of transcendental numbers. Further improvements in the upper bound were obtained by Thue (1909), Siegel (1921), and Dyson and Gelfand (1947), culminating in the 1955 theorem of Roth: for $x \in \mathbb{R}$ algebraic over \mathbb{Q} , $\tau_x \leq 2$. Thus the theorems of Dirichlet and Roth give $\tau_x = 2$ for irrational algebraic $x \in \mathbb{R}$.

The invariant $\alpha_x(L)$. In §2 we generalize the approximation exponent to arbitrary projective varieties X defined over a number field k . To do this we replace the function $|x - a/b|$ by a distance function $d_v(x, \cdot)$ depending on a place v of k , and measure the complexity of a rational point via a height function $H_L(\cdot)$ depending on an ample line bundle L . The one essential change in our definition is to move the exponent from the height to the distance. As a result, as Proposition 2.12 shows, for $x \in \mathbb{R} = \mathbb{A}^1(\mathbb{R}) \subset \mathbb{P}^1(\mathbb{R})$ we have $\alpha_x(\mathcal{O}_{\mathbb{P}^1}(1)) = \frac{1}{\tau_x}$. The choice of moving the exponent is justified by Proposition 2.15(a,b) which shows that this form is more natural when we vary L , and by the resulting similarities with the Seshadri constant.

In particular, for $x \in \mathbb{R} \setminus \mathbb{Q}$, algebraic of degree d over \mathbb{Q} , and $L = \mathcal{O}_{\mathbb{P}^1}(1)$ the theorems of Liouville and Roth become the lower bounds $\alpha_x(L) \geq \frac{1}{d}$ and $\alpha_x(L) \geq \frac{1}{2}$ respectively. One of the main goals of this paper is to generalize these statements to lower bounds for $\alpha_x(L)$ on an arbitrary variety X .

Examples. Here are three examples of lower bounds on α given by previously known results on Diophantine approximation. We work over an arbitrary number field k .

- (a) If $X = \mathbb{P}_k^1$, $x \in X(\bar{k})$, $L = \mathcal{O}_{\mathbb{P}^1}(1)$, then $\alpha_x(L) \geq \frac{1}{2}$.
- (b) If $X = \mathbb{P}_k^n$, $x \in X(\bar{k})$, $L = \mathcal{O}_{\mathbb{P}^n}(1)$, then either $\alpha_x(L) \geq \frac{n}{n+1}$ or there is a smaller linear space $Z \cong \mathbb{P}_k^m \subset \mathbb{P}_k^n$, with $m < n$, and $x \in Z(\bar{k})$.
- (c) If X is an abelian variety, $x \in X(\bar{k})$, and L any ample line bundle then $\alpha_x(L) = \infty$.

Example (a) is Roth's theorem for a general number field (and place v), and example (b) follows from the Schmidt subspace theorem. In both of these cases by using a Dirichlet-type argument [3] one obtains exact values for α_x . In the case of \mathbb{P}^1 , if $x \in \mathbb{P}^1(\bar{k} \cap k_v) \setminus \mathbb{P}^1(k)$ then $\alpha_x(\mathcal{O}_{\mathbb{P}^1}(1)) = \frac{1}{2}$. In the case of \mathbb{P}^n , if $x \in \mathbb{P}^n(\bar{k} \cap k_v) \setminus \mathbb{P}^n(k)$, and m is the smallest value so that there exists a linear subspace $\mathbb{P}_k^m \subset \mathbb{P}_k^n$ with $x \in \mathbb{P}^m(\bar{k})$, then $\alpha_x(\mathcal{O}_{\mathbb{P}^n}(1)) = \frac{m}{m+1}$. Finally example (c) is [19, p. 98; second theorem].

The basic interpretation of $\alpha_x(L)$ is as the cost in complexity required to get closer to x . When α_x is finite this indicates that the complexity has polynomial growth in the reciprocal of the distance, with α_x as the exponent. In example (c) the complexity grows roughly exponentially in the reciprocal of the distance (see [19, p. 98 again]) and thus $\alpha_x = \infty$.

The invariant $\epsilon_x(L)$. The definition and elementary properties of the Seshadri constants are given in §3. We list two of these properties, and the corresponding properties for α , here in order to emphasise the similarity between α_x and ϵ_x , and to use one of the properties in the discussion below. Both α and ϵ make sense for \mathbb{Q} -bundles. Fix $x \in X(\bar{k})$, then

(a) for any ample \mathbb{Q} -bundle L , and any $m \in \mathbb{Q}_{>0}$,

$$\alpha_x(mL) = m\alpha_x(L) \text{ and } \epsilon_x(mL) = m\epsilon_x(L);$$

(b) α_x and ϵ_x are concave functions of the line bundle. For any ample \mathbb{Q} -bundles L_1 and L_2 , and any $a, b \in \mathbb{Q}_{>0}$,

$$\alpha_x(aL_1 + bL_2) \geq a\alpha_x(L_1) + b\alpha_x(L_2) \text{ and } \epsilon_x(aL_1 + bL_2) \geq a\epsilon_x(L_1) + b\epsilon_x(L_2).$$

These and other parallel properties appear in Propositions 2.15 and 3.4.

Examples.

(a) If $X = \mathbb{P}^n$, $x \in X(\mathbb{C})$, and $L = \mathcal{O}_{\mathbb{P}^n}(1)$ then $\epsilon_x(L) = 1$, and so $\epsilon_x(\mathcal{O}_{\mathbb{P}^n}(e)) = e$ for all $e > 0$.

(b) If X is a smooth cubic surface, and $L = \mathcal{O}_{\mathbb{P}^3}(1)|_X$ then

$$\epsilon_x(L) = \begin{cases} 1 & \text{if } x \text{ is on a line} \\ \frac{3}{2} & \text{otherwise.} \end{cases}$$

If X is a variety with a transitive group action, such as \mathbb{P}^n or an abelian variety, then the value of $\epsilon_x(L)$ is independent of $x \in X(\mathbb{C})$. One thesis of this paper is that ϵ_x affects approximation results. On varieties where ϵ_x does not depend on the point this effect is essentially invisible since it becomes a global property of the line bundle. On arbitrary varieties however one can expect more precise approximation theorems by taking the differing values of ϵ into account. This will be a feature of the results below.

Roth theorems.¹ If X is a variety over $\text{Spec}(k)$, and $x \in X(\bar{k})$ with field of definition K , then for any ample line bundle L on X we have $\alpha_x(L) \geq \frac{1}{d}\epsilon_x(L)$, where $d = [K:k]$. On \mathbb{P}^1 , this is the inequality $\alpha_x(\mathcal{O}_{\mathbb{P}^1}(1)) \geq \frac{1}{d}$, and hence we regard this as the general version of Liouville's theorem. This result follows from elementary properties of the height of the exceptional divisor (see the end of §3 or [15, §3] for a proof).

Our main concern is proving general ‘‘Roth’’ theorems. By this we mean lower bounds on $\alpha_x(L)$ that are: (1) independent of the field of definition of x , and (2) (following the philosophy of this paper) expressed in terms of $\epsilon_x(L)$. The examples of \mathbb{P}^n and \mathbb{P}^1 suggest two possible interpretations of this goal.

First, based on the example of \mathbb{P}^n one might hope for a theorem of the form: for every $n \geq 1$ there is a constant c_n so that for every irreducible n -dimensional variety X , ample line bundle L and $x \in X(\bar{k})$, either $\alpha_x(L) \geq c_n\epsilon_x(L)$ or there is a proper subvariety Z , with $x \in Z(\bar{k})$, such that $\alpha_x(L) = \alpha_{x,Z}(L|_Z)$.

Second, one might seek to generalize the \mathbb{P}^1 example: there is a constant c so that for every variety X , every ample line bundle L , and every $x \in X(\bar{k})$ the inequality $\alpha_x(L) \geq c\epsilon_x(L)$ holds. Considering varieties of the form $X = \mathbb{P}^1 \times Y$ shows that $c = \frac{1}{2}$ is the best possible constant (i.e., it does not help to have the constant vary with the dimension of X).

We establish versions of both of these statements; here is our version of the first type.

¹All uses of ‘‘Roth’’ as an adjective in this paper are in homage to the theorem proved by Klaus F. Roth and its later extensions by Ridout and Lang, and do not refer to the second named author of the paper.

Theorem (6.2, “Schmidt type”): *Let X be an irreducible n -dimensional variety over $\text{Spec}(k)$. For any ample \mathbb{Q} -bundle L and any $x \in X(\bar{k})$ either*

$$(a) \alpha_x(L) \geq \frac{n}{n+1} \epsilon_x(L)$$

or

(b) *there exists a proper subvariety $Z \subset X$, irreducible over \bar{k} , with $x \in Z(\bar{k})$ so that $\alpha_{x,X}(L) = \alpha_{x,Z}(L|_Z)$, i.e., “ $\alpha_x(L)$ is computed on a proper subvariety of X ”.*

This theorem has an equivalent version expressed in more familiar terms.

Theorem (6.2, alternate statement): *Let L be any ample \mathbb{Q} -bundle on X , and choose any $x \in X(\bar{k})$. Then there is a proper subvariety $Z \subset X$ so that for each $\delta > 0$ there are only finitely many solutions $y \in X(k) \setminus Z(k)$ to*

$$d_v(x, y) < H_L(y)^{-\left(\frac{n+1}{n\epsilon_x(L)} + \delta\right)}.$$

Theorem 6.2 generalizes the Schmidt subspace theorem, insofar as the Schmidt theorem concerns approximating a point. It is an important part of the Schmidt theorem that Z be a union of linear spaces so that the theorem may be applied inductively. Since Theorem 6.2 applies to arbitrary varieties, the ability to apply induction of this type is automatic. In particular, since the Seshadri constant is weakly increasing when restricting to a subvariety (Proposition 3.4(c)), Theorem 6.3 and induction on dimension yield a theorem of the second type.

Theorem (6.3, “Roth type”): *For all varieties X over $\text{Spec}(k)$ (possibly reducible), all $x \in X(\bar{k})$ and all ample line bundles L , $\alpha_x(L) \geq \frac{1}{2} \epsilon_x(L)$.*

In order for equality to hold in Theorem 6.3 the induction must have gone down to a one-dimensional variety, and from this we deduce that if equality holds then there is a k -rational curve C passing through x , and unibranch at x , which also computes the Seshadri constant, i.e., $\epsilon_x(L) = \epsilon_{x,C}(L|_C)$. The exact statement and its converse appear as part of Theorem 6.3, as fully stated in §6. This is one of the few examples we know of where an arithmetic condition about approximation implies a geometric condition about X (namely that there must be a rational curve passing through x .) If there is no rational curve passing through x then the lower bound in Theorem 6.3 may be improved; see Corollary 6.6.

It is useful to state Theorem 6.3 in an equivalent form closer to that of the usual statement of Roth’s theorem on \mathbb{P}^1 .

Corollary (6.4): *For any $\delta > 0$ there are only finitely many $y \in X(k)$ such that*

$$d_v(x, y) < H_L(y)^{-\left(\frac{2}{\epsilon_x(L)} + \delta\right)}.$$

Heuristic explanation. Given an ample line bundle L and $x \in X(\bar{k})$ consider the problem of finding an exponent e so that for all $\delta > 0$ there are only finitely many solutions $y \in X(k)$ to $d_v(x, y) < H_L(y)^{-(e+\delta)}$. If m is such that mL is very ample, then embedding X via mL , projecting on coordinates, and using Roth’s theorem for \mathbb{P}^1 shows that the exponent $e = 2m$ will do. The smaller the value of e , the stronger such a statement is, so we now ask the question: what is the smallest value of m so that mL is very ample?

If A is a very ample line bundle, then $\epsilon_{x'}(A) \geq 1$ for all $x' \in X(\bar{k})$ (see Proposition 3.4(d)). In particular, if mL is very ample then we must have $m\epsilon_x(L) = \epsilon_x(mL) \geq 1$, and thus $m \geq \frac{1}{\epsilon_x(L)}$. In general $m = \frac{1}{\epsilon_x(L)}$ does not guarantee that mL is very ample. There are basically three problems. (1) We need $\alpha_{x'}(mL) \geq 1$ for all $x' \in X$, and not just x . (2) Even if the previous condition holds, this does not guarantee that mL is very ample. (3) With this value of m , mL may not be an integral (or, conjecturally, even a rational) line bundle.

As an example of two of these issues, let X be a smooth cubic surface, $L = \mathcal{O}_{\mathbb{P}^3}(1)|_X$, and $x \in X(\bar{k})$ a point not on a line. As stated above $\epsilon_x(L) = \frac{3}{2}$. However $\frac{2}{3}L$ is not an integral line bundle (it has degree $\frac{2}{3}$ on every line), nor is $\epsilon_{x'}(\frac{2}{3}L) \geq 1$ for points x' on a line.

The essential point of Corollary 6.4 is that these concerns don't matter: as long as we only care about approximating x the local estimate of amplitude $m = \frac{1}{\epsilon_x(L)}$ works. This is a good illustration of the effects of local positivity on approximation.

Simultaneous approximation. As with Roth's theorem on \mathbb{P}^1 , our theorems admit generalizations to simultaneous approximation. In order to indicate the nature of the results let us consider the two equivalent statements for a single place given by Theorem 6.3 and Corollary 6.4 above and see how they generalize. In §2, as part of defining $\alpha_x(L)$ we also define $\alpha_x(\{x_i\}, L)$ for any sequence $\{x_i\}$ of k -points of X , and we will need this notation to state our results below. In particular, Theorem 6.3 can be equivalently stated as $\alpha_x(\{x_i\}, L) \geq \frac{1}{2}\epsilon_x(L)$ for all sequences $\{x_i\}$ of k -points of X .

To set up the simultaneous approximation problem let S be a finite set of places of k , each extended to \bar{k} . For each $v \in S$ let $d_v(\cdot, \cdot)$ be the distance function computed with respect to $v \in S$ and choose a point $x_v \in X(\bar{k})$. To simplify notation, we set α_v to be α_{x_v} computed with respect to d_v .

We are interested in understanding how well sequences of k -points can simultaneously approximate each x_v . The generalizations of Theorem 6.3 and Corollary 6.4 to simultaneous approximation (see Corollary 7.6) are respectively:

- (1) for any sequence $\{x_i\}$ of k -points, $\sum_{v \in S} \frac{\epsilon_{x_v}(L)}{\alpha_v(\{x_i\}, L)} \leq 2$, and
- (2) for any $\delta > 0$ there are only finitely many $y \in X(k)$ such that

$$\prod_{v \in S} d_v(x_v, y)^{\epsilon_{x_v}(L)} < H_L(y)^{-(2+\delta)}.$$

The other results (e.g., Theorems 6.1 and 6.2) also have their simultaneous versions. Full statements and further discussion appear in §7.

Improvements via étale covers. Given X (which we assume normal to simplify the discussion), an ample line bundle L on X , and $x \in X(\bar{k})$ we define $\hat{\epsilon}_x^{\text{ét}}(L)$ by

$$\hat{\epsilon}_x^{\text{ét}}(L) = \sup_{y \in \varphi^{-1}(x)} \epsilon_y(\varphi^*L)$$

where the supremum is over all irreducible étale covers $\varphi: Y \rightarrow X$. In §8 we show that all the previous theorems, for one place or simultaneous places, hold with $\epsilon_x(L)$ replaced by $\hat{\epsilon}_x^{\text{ét}}(L)$ (see Corollary 8.9). Since $\hat{\epsilon}^{\text{ét}}$ is in general larger, this can be a significant strengthening of the results. For instance, if X is an abelian variety and L an ample line bundle, then $\epsilon_x(L)$ is always finite, while $\hat{\epsilon}_x^{\text{ét}}(L) = \infty$ (see the example on page 48). Thus Theorem 6.3 applied with $\hat{\epsilon}^{\text{ét}}$ in place of ϵ shows that $\alpha_x(L) = \infty$ on an abelian variety.

The results proved in §8 are slightly more general (for instance, one can take the supremum over irreducible unramified covers) and the reader is referred there for more detailed statements.

In his 1962 book *Diophantine Geometry*, Lang ([10, p. 119]) suggests three directions for future progress on Roth's theorem. The first is to make the result quantitative, and we seem to know as much about this now as was known in 1962; the second is to deal with approximation in \mathbb{A}^n or \mathbb{P}^n , which has been fully answered by the Schmidt subspace theorem; and the third (in paraphrase) is to generalize Roth's theorem to projective varieties in a way which is compatible with unramified covers. We feel that the results of this paper are a partial fulfillment of the third suggestion. (We say partial since Lang wanted a generalization of his "geometric formulation" of Roth's theorem, which applied to maps, and since it is not completely clear to us what Lang intended by this suggestion. Unfortunately we can no longer ask him.)

Other results. The proofs of the theorems (in particular Theorem 6.2) hinge on a third invariant of a point and ample line bundle L . This invariant, $\beta_x(L)$, is defined in §4 and further explored in §9. This invariant is purely geometric in the sense that, like $\epsilon_x(L)$, it only depends on the base change of X to the algebraic closure.

This invariant is obtained by integrating a function $f(\gamma)$ which measures the "relative asymptotic volume" of the subspace of sections of L vanishing to order $\geq \gamma$ at x . One of the reasons for using $\beta_x(L)$ is that the asymptotic behaviour of a line bundle is often better than any particular multiple.

In order to prove Theorem 6.2 we first prove an approximation result using $\beta_x(L)$.

Theorem (6.1): *Let X be an irreducible variety over $\text{Spec}(k)$. Then for any ample \mathbb{Q} -bundle L and any $x \in X(\bar{k})$ either*

$$(a) \alpha_x(L) \geq \beta_x(L)$$

or

$$(b) \alpha_x(L) \text{ is computed on a proper subvariety of } X.$$

If X is n -dimensional then there is an easy estimate $\beta_x(L) \geq \frac{n}{n+1}\epsilon_x(L)$ (see Corollary 4.4) and so Theorem 6.1 immediately implies Theorem 6.2. It is interesting to study when $\beta_x(L) = \frac{n}{n+1}\epsilon_x(L)$, i.e., when replacing $\beta_x(L)$ by $\frac{n}{n+1}\epsilon_x(L)$ does not diminish the strength of the result. Equivalent conditions for this equality are given in Theorem 9.1. The reader will also find a heuristic interpretation of $\beta_x(L)$ in §9.

Finally, we note that §8 also proves that all theorems involving $\beta_x(L)$ hold with $\beta_x(L)$ replaced by its limit $\hat{\beta}_x(L)$ over unramified covers.

Remarks on the proof. The central motor of this paper, which largely implies the other approximation results, is Theorem 5.2 to which §5 is devoted. This theorem is a simultaneous approximation theorem written in terms of $\{\beta_{x_v}(L)\}_{v \in S}$ where S is a finite set of places of k , and $x_v \in X(\bar{k})$ for $v \in S$. Theorem 5.2 is proved using the Faltings-Wüstholz theorem and the definition of $\beta_x(L)$ has been chosen in order to optimize an estimate used in applying that theorem. The basic idea is explained at the beginning of the proof of Theorem 5.2, which appears at the end of §5.

The Faltings-Wüstholz theorem implies Roth's theorem for \mathbb{P}^1 and the Schmidt subspace theorem, and thus the values of $\frac{n}{n+1}$ and $\frac{1}{2}$ when approximating on \mathbb{P}^n and \mathbb{P}^1 respectively.

Our theorems (e.g., Theorem 6.2 applied to \mathbb{P}^n) also produce these values, but we deduce them from the Faltings-Wüstholz theorem by a different method than their paper, and it is worth commenting on this difference.

In the argument of [7, §9] the value $\frac{n}{n+1}$ arises as the ratio of the dimension of the subspace of $\Gamma(\mathbb{P}^n, \mathcal{O}_{\mathbb{P}^n}(1))$ vanishing at a point x , and the dimension of the entire space. In our result the value $\frac{n}{n+1}$ arises as the integral $\beta_x(\mathcal{O}_{\mathbb{P}^n}(1)) = \int_0^1 f(\gamma) d\gamma$ of the relative asymptotic volume function $f(\gamma) = 1 - \gamma^n$ for the line bundle $\mathcal{O}_{\mathbb{P}^n}(1)$. Thus — as mentioned above as a motivation for β_x — we deduce the constant $\frac{n}{n+1}$ from asymptotic properties of $\mathcal{O}_{\mathbb{P}^n}(1)$ and not from its global sections.

Organization of the paper. Sections 2, 3, and 4 are devoted to the definitions and basic properties of $\alpha_x(L)$, $\epsilon_x(L)$, and $\beta_x(L)$ respectively. In §5 we prove Theorem 5.2, which will be used to prove all the other approximation results in the paper. In §6 we prove approximation results for a single place, and in §7 we prove simultaneous approximation results for several places. In §8 we show that all of the previous theorems hold with β_x and ϵ_x replaced by their suprema $\hat{\beta}_x$ and $\hat{\epsilon}_x$ over unramified covers. In §9 we provide some complementary material about $\beta_x(L)$, and finally in §10 we give an elementary application of our theorems to establish some previously unknown special cases of Vojta's main conjecture.

Notation and Conventions. Unless otherwise specified we work over a fixed number field k . By “variety over $\text{Spec}(k)$ ” we mean a (possibly reducible, possibly singular) projective variety over $\text{Spec}(k)$, i.e., a reduced projective scheme over $\text{Spec}(k)$. We use additive notation for line bundles since this is in line with the behaviour of α_x , ϵ_x , and β_x . On a product $X \times Y$ we therefore use $L_1 \boxplus L_2$ instead of $L_1 \boxtimes L_2$ for a line bundle of the form $pr_X^* L_1 + pr_Y^* L_2$, with pr_X and pr_Y being the projections.

If X is a variety over $\text{Spec}(k)$, a point $x \in X(\bar{k})$ is a map $\text{Spec}(\bar{k}) \rightarrow X$ of k -schemes. Such a point gives rise to a point of $X \times_k \bar{k}$, and a closed point (the image of this map) of X . The symbol $\kappa(x)$ denotes the residue field of this closed point of X , called the field of definition of x . We say that “ x is defined over K ” if $\kappa(x)$ is a subfield of K (this inclusion may be implicit). A sequence of k -points of X (or a sequence in $X(k)$) means an infinite sequence of distinct points of $X(k)$. We denote such a sequence by $\{x_i\}$ rather than $\{x_i\}_{i \geq 0}$.

The absolute values are normalized with respect to k : if v is a finite place of k , π a uniformizer of the corresponding maximal ideal, and κ the residue field then $\|\pi\|_v = 1/\#\kappa$; if v is an infinite place corresponding to an embedding $i: k \hookrightarrow \mathbb{C}$ then $\|x\|_v = |i(x)|^{m_v}$ for all $x \in k$, where $m_v = 1$ or 2 depending on whether v is real or complex.

Two real-valued functions g and g' with the same domain are called *equivalent* if there are positive real constants $c \leq C$ so that $cg \leq g' \leq Cg$ for all values of the domain. We will apply this terminology in three situations: to distance functions $d_v(\cdot, \cdot)$, to height functions $H_L(\cdot)$, and to partially evaluated distance functions $d_v(x, \cdot)$. Typical domains are $X(\bar{k}) \times X(\bar{k})$, $X(\bar{k})$, and Zariski open subsets or v -adically compact subsets of these.

Acknowledgements. We thank Chris Dionne, Laurence Ein, Robert Lazarsfeld, Victor Lozovanu, and Damien Roy for helpful discussions. We are also extremely grateful to the referees of this paper for pointing out several mathematical and expositional errors in the initial versions, and for their suggestions on how to correct them. Finally, we wish to acknowledge an intellectual debt to Michael Nakamaye who has long advocated the point of view that Seshadri constants are diophantine.

2. APPROXIMATION BY RATIONAL POINTS

Let k be a number field, and X a projective variety over $\text{Spec}(k)$. We begin by discussing the distance functions in the archimedean and non-archimedean cases.

Distance Functions: Archimedean case. Fix an archimedean place v_0 of k , and an extension of v_0 to \bar{k} , which we denote by v . We choose a distance function on $X(\bar{k})$ by choosing an embedding $X \hookrightarrow \mathbb{P}_k^r$ and pulling back (via v) the function on $\mathbb{P}^r(\mathbb{C}) \times \mathbb{P}^r(\mathbb{C})$ given by the formula

$$(2.1.1) \quad d_v(x, y) = \left(1 - \frac{|\sum_{i=0}^r x_i \bar{y}_i|^2}{(\sum_{i=0}^r |x_i|^2)(\sum_{j=0}^r |y_j|^2)} \right)^{[k_v:\mathbb{R}]/2}$$

where $x = [x_0:\cdots:x_r]$, and $y = [y_0:\cdots:y_r]$ are points of $\mathbb{P}^r(\mathbb{C})$, and $|\cdot|$ is the absolute value on \mathbb{C} extending the usual absolute value on \mathbb{R} , i.e., such that $|3 + 4\sqrt{-1}| = 5$.

Note that if $k_v = \mathbb{C}$ then this function does not satisfy the triangle inequality, nonetheless we continue to call it a distance function. (To see that this function does satisfy the triangle inequality if $k_v = \mathbb{R}$ see [1, Proposition 2.8.18].)

Distance Functions: Non-archimedean case. Fix a non-archimedean place v_0 of k , and an extension of v_0 to \bar{k} , which we denote by v . The place v defines an absolute value $\|\cdot\|_v$ on \bar{k} , normalized according to our conventions in the introduction. (This normalization agrees with the use of the symbol $\|\cdot\|_v$ in the books of Bombieri-Gubler and Hindry-Silverman; see [1, 1.3.6 and 1.4.3] and [8, p. 171–172] respectively.) We choose a distance function on $X(\bar{k})$ by choosing an embedding $X \hookrightarrow \mathbb{P}_k^r$ and pulling back the distance function on $\mathbb{P}^r(\bar{k})$ given by the formula

$$(2.1.2) \quad d_v(x, y) = \frac{\max_{0 \leq i < j \leq r} (\|x_i y_j - x_j y_i\|_v)}{\max_{0 \leq i \leq r} (\|x_i\|_v) \max_{0 \leq j \leq r} (\|y_j\|_v)}$$

where $x = [x_0:\cdots:x_r]$, and $y = [y_0:\cdots:y_r]$ are points of $\mathbb{P}^r(\bar{k})$.

Basic properties of distance functions. These definitions are somewhat opaque on first reading, but they are standard distance functions in Arakelov theory, albeit normalized with respect to k , rather than \mathbb{Q} . (See for instance [1, §2.8] where a distance function $\delta_v(\cdot, \cdot)$ is defined for each place v ; the distance functions are related by the formula $d_v(\cdot, \cdot) = \delta_v(\cdot, \cdot)^{[k:\mathbb{Q}]}$.) We will also briefly discuss the geometric meaning of $d_v(\cdot, \cdot)$ for non-archimedean v below.

We note two elementary properties of the distance function, whose proofs follow easily from the definitions.

Proposition 2.2. *Let v be a place of k extended to \bar{k} , and $d_v(\cdot, \cdot)$ the distance function constructed by choosing an embedding $X \hookrightarrow \mathbb{P}_k^r$. Then*

- (a) *For all $x, y \in X(\bar{k})$ we have $d_v(x, y) \in [0, 1]$, with $d_v(x, y) = 0$ if and only if $x = y$.*
- (b) *If K is a finite extension of k , then $d_v(\cdot, \cdot)_K = d_v(\cdot, \cdot)_k^{m_v}$, where $m_v = [K_v : k_v]$ is the local degree. (Here $d_v(\cdot, \cdot)_K$ refers to the distance function defined by using the same embedding and normalizing with respect to K and $d_v(\cdot, \cdot)_k$ the distance function normalized with respect to k , as above.)*

We will use the next result several times in proving equivalence of different types of distance functions. Let \mathbb{C}_v be the completion of \bar{k} with respect to the place v .

Lemma 2.3. *Let Y be a variety over $\text{Spec}(k)$, U an affine open subset of $Y_K = Y \times_k K$ for some finite extension K/k , and u_1, \dots, u_r and u'_1, \dots, u'_s two collections of elements of $\Gamma(U, \mathcal{O}_Y)$ which generate the same ideal. Then the functions $\max(\|u_1(\cdot)\|_v, \dots, \|u_r(\cdot)\|_v)$ and $\max(\|u'_1(\cdot)\|_v, \dots, \|u'_s(\cdot)\|_v)$ are equivalent on any compact subset of $U(\mathbb{C}_v)$.*

Proof: Since u_1, \dots, u_r and u'_1, \dots, u'_s generate the same ideal on U there are functions $f_{j,\ell} \in \Gamma(U, \mathcal{O}_{Y_K})$ such that $u_j = \sum_{\ell=1}^s f_{j,\ell} u'_\ell$ for each $j = 1, \dots, r$. Similarly there are functions $g_{\ell,j} \in \Gamma(U, \mathcal{O}_{Y_K})$ such that $u'_\ell = \sum_{j=1}^r g_{\ell,j} u_j$ for all $\ell = 1, \dots, s$. On any compact subset T of $U(\mathbb{C}_v)$ the functions $\|f_{j,\ell}(\cdot)\|_v$ and $\|g_{\ell,j}(\cdot)\|_v$ are bounded on T . It follows that the functions $\max(\|u_1(\cdot)\|_v, \dots, \|u_r(\cdot)\|_v)$ and $\max(\|u'_1(\cdot)\|_v, \dots, \|u'_s(\cdot)\|_v)$ are equivalent on T . \square

Remark. Let x be a point of $X(\bar{k})$ and let K be the field of definition of x . Throughout the paper we will be interested in approximating x by points of $X(k)$. If $K \not\subseteq k_v$, or equivalently, $K_v \neq k_v$ then it will be impossible to find a sequence of points of $X(k)$ converging (in terms of d_v) to x (e.g., when v is archimedean this happens when $k_v = \mathbb{R}$ and $K_v = \mathbb{C}$). Thus, in all cases we can approximate x by points of $X(k)$ we may assume that $K_v = k_v$.

Lemma 2.4. *Let V and W be vector spaces over k , $j: X \hookrightarrow \mathbb{P}(V^*)$ and $j': X \hookrightarrow \mathbb{P}(W^*)$ embeddings, $\mathbb{P}(V^*) \cong \mathbb{P}^r$ and $\mathbb{P}(W^*) \cong \mathbb{P}^s$ choices of coordinates, and d_v and d'_v the induced distance functions on X . Let K/k be any finite extension. Then for any point $x \in X(K_v)$ and any inclusion $V \hookrightarrow W$ of k -vector spaces so that the resulting rational map $f: \mathbb{P}^s \rightarrow \mathbb{P}^r$ is defined at $j(x)$ and such that $f \circ j = j'$ near x , there is a compact v -adic neighbourhood T of (x, x) in $X(K_v) \times X(K_v)$ such that d_v and d'_v are equivalent on T .*

Proof: Change of basis by k -linear transformation only changes the distance function by bounded amount (see [3, Theorem 3] for this statement for δ_v). We may therefore change coordinates and assume that the map f is given by dropping the last $s - r$ coordinates on \mathbb{P}^s . In the non-archimedean case we are therefore reduced to comparing the behaviour of

$$\frac{\max_{0 \leq i < j \leq s} (\|x_i y_j - x_j y_i\|_v)}{\max_{0 \leq i \leq s} (\|x_i\|_v) \max_{0 \leq j \leq s} (\|y_j\|_v)} \quad \text{and} \quad \frac{\max_{0 \leq i < j \leq r} (\|x_i y_j - x_j y_i\|_v)}{\max_{0 \leq i \leq r} (\|x_i\|_v) \max_{0 \leq j \leq r} (\|y_j\|_v)}$$

near (x, x) . Let X_0, \dots, X_s and Y_0, \dots, Y_s be the coordinates on $\mathbb{P}^s \times \mathbb{P}^s$. Choose an affine open U containing x such that the embedding line bundle is trivial on U and so we may identify sections with functions. Since f is defined at x the functions X_0, \dots, X_r have no common zero at x , and so by shrinking U we may assume that they generate the unit ideal on U . The same is therefore true for the larger collection of functions X_0, \dots, X_s . By Lemma 2.3 we thus have that $\max_{0 \leq i \leq s} (\|x_i\|_v)$ and $\max_{0 \leq i \leq r} (\|x_i\|_v)$ are equivalent on any compact neighbourhood of $x \in X(K_v) \subset X(\mathbb{C}_v)$. (Note that in the non-archimedean case these compact neighbourhoods in $X(K_v)$ will have empty interior as a subset of $X(\mathbb{C}_v)$. Indeed, \mathbb{C}_v is not locally compact, so there are no compact neighbourhoods of x in $X(\mathbb{C}_v)$ at all.) A similar statement applies to the coordinates Y_i . The functions $\{X_i Y_j - X_j Y_i\}_{0 \leq i < j \leq s}$ and $\{X_i Y_j - X_j Y_i\}_{0 \leq i < j \leq r}$ generate the ideal of the diagonal on $\mathbb{P}^s \times \mathbb{P}^s$ and $\mathbb{P}^r \times \mathbb{P}^r$ respectively, and so restricted to $U \times U$ generate the ideal of the diagonal there. Applying Lemma 2.3 again, the functions $\max_{0 \leq i < j \leq s} (\|x_i y_j - x_j y_i\|_v)$ and $\max_{0 \leq i < j \leq r} (\|x_i y_j - x_j y_i\|_v)$ are equivalent on a compact neighbourhood of (x, x) in $U(K_v) \times U(K_v)$. This proves the lemma in the non-archimedean case. In the archimedean case one uses the same strategy, the identity

$$1 - \frac{|\sum_{i=0}^r x_i \bar{y}_i|^2}{(\sum_{i=0}^r |x_i|^2)(\sum_{j=0}^r |y_j|^2)} = \frac{\sum_{0 \leq i < j \leq r} |x_i y_j - x_j y_i|^2}{(\sum_{0 \leq i \leq r} |x_i|^2)(\sum_{0 \leq j \leq r} |y_j|^2)},$$

and the fact that $\max(\|f_1\|_v, \dots, \|f_r\|_v)$ and $(\|f_1\|^2 + \dots + \|f_r\|^2)^{[k_v:\mathbb{R}]/2}$ are equivalent for functions f_1, \dots, f_r . \square

We now check that distance functions coming from two different embeddings are equivalent. We are indebted to one of the referees of this paper for the following argument, which is substantially simpler and shorter than our original one.

Proposition 2.5. *Let d_v and d'_v be two distance functions coming from different embeddings of X . Then for all finite extensions K/k , d_v is equivalent to d'_v on $X(K_v) \times X(K_v)$.*

Proof: It suffices to show that for each $x \in X(K_v)$ there is a compact neighbourhood of (x, x) in $X(K_v) \times X(K_v)$ where d_v and d'_v are equivalent. Since X is projective, $X(K_v) \times X(K_v)$ can be covered by finitely many such neighbourhoods which then proves the proposition.

We first observe that we may assume that each of the embeddings is by a complete linear series. Suppose that L is a very ample line bundle and $V \subset W = H^0(X, L)$ a basepoint-free subseries such that the associated map $j: X \rightarrow \mathbb{P}(V^*)$ is an embedding. Then $j': X \hookrightarrow \mathbb{P}(W^*)$ is also an embedding and the rational map $\mathbb{P}(W^*) \rightarrow \mathbb{P}(V^*)$ arising from the inclusion $V \hookrightarrow W$ is defined everywhere along $j(X)$. Thus the result we want follows from Lemma 2.4.

Now let d_v and d'_v be two distance functions coming from embeddings $j: X \hookrightarrow \mathbb{P}(V^*)$ and $j': X \hookrightarrow \mathbb{P}(W^*)$, with $V = H^0(X, L)$ and $W = H^0(X, L')$ where L and L' are very ample line bundles. Assume that $L - L'$ is basepoint free. Then for any point $x \in X(K_v)$ there is a section s of $L - L'$ that does not vanish at x . Multiplication by s induces a rational map $f_s: \mathbb{P}(V^*) \rightarrow \mathbb{P}(W^*)$ that is defined at $j(x)$ and such that $f_s \circ j = j'$ near x . Thus the neighbourhood we want is again guaranteed by Lemma 2.4.

Finally, for general L and L' , we may replace L with a multiple mL large enough so that $mL - L$ and $mL - L'$ are both basepoint free, and the proposition immediately follows. \square

We next turn to local descriptions of the distance function useful in computations.

Lemma 2.6. *Let x be a point of $X(\bar{k})$ and K any finite extension of k over which x is defined. Then there exists an open affine subset U of $X_K := X \times_k K$ containing x , and elements u_1, \dots, u_r of $\Gamma(U, \mathcal{O}_{X_K})$ which generate the maximal ideal of x and positive real constants $c \leq C$ such that*

$$(2.6.1) \quad cd_v(x, y) \leq \min(1, \max(\|u_1(y)\|_v, \dots, \|u_m(y)\|_v)) \leq C d_v(x, y)$$

for all $y \in U(K_v)$. That is, on $U(K_v)$ the function $\min(1, \max(\|u_1(\cdot)\|_v, \dots, \|u_r(\cdot)\|_v))$ is equivalent to the function $d_v(x, \cdot)$.

Proof: We start with two reductions. First, since the absolute value $\|\cdot\|_v$ and the distance function $d_v(\cdot, \cdot)$ transform the same way under field extensions, we may assume that x is defined over k . Second, by Proposition 2.5 we may choose whichever embedding of X we wish when performing the calculation.

Given these reductions, choose an embedding $X \hookrightarrow \mathbb{P}^r$ so that x is sent to $[1:0:\dots:0]$. Let Z_0, \dots, Z_r be homogenous coordinates on \mathbb{P}^r , and choose the open set U of X to be the set $Z_0 \neq 0$, and $u_i = Z_i/Z_0$ for $i = 1, \dots, r$ as the generators of the maximal ideal of x . If v is non-archimedean, then (2.1.2) and the fact that x is sent to $[1:0:\dots:0]$ give

$$d_v(x, y) = \frac{\max(\|Z_1(y)\|_v, \|Z_2(y)\|_v, \dots, \|Z_r(y)\|_v)}{\max(\|Z_0(y)\|_v, \|Z_1(y)\|_v, \dots, \|Z_r(y)\|_v)} \text{ for all } y \in X(\bar{k}).$$

For $y \in U(K_v)$, this is equal to $\min(1, \max(\|u_1(y)\|_v, \dots, \|u_r(y)\|_v))$.

In the case that v is archimedean, we may further assume that $k_v = \mathbb{C}$, again using the fact that the functions to be compared transform the same way under field extensions. From (2.1.1) and the fact that x is sent to $[1:0:\dots:0]$ we obtain

$$d_v(x, y) = 1 - \frac{|y_0|^2}{|y_0|^2 + \dots + |y_r|^2} = \frac{|y_1|^2 + \dots + |y_r|^2}{|y_0|^2 + |y_1|^2 + \dots + |y_r|^2} = \frac{\|y_1\|_v + \dots + \|y_r\|_v}{\|y_0\|_v + \|y_1\|_v + \dots + \|y_r\|_v}.$$

For $y \in U(K_v)$, $y_0 \neq 0$, and $u_j(y) = y_j/y_0$ for $j = 1, \dots, r$. Thus $d_v(x, y) = \frac{\|u_1(y)\|_v + \dots + \|u_r(y)\|_v}{1 + \|u_1(y)\|_v + \dots + \|u_r(y)\|_v}$; it is then elementary to check that (2.6.1) holds with $c = \frac{1}{r}$ and $C = 2$. \square

We will need an extension of this lemma which applies to any affine open U containing x , and any choice u_1, \dots, u_r of generators the maximal ideal. To do this we need to drop the requirement that the result hold for all $y \in U(K_v)$, and restrict to points v -adically close to x ; for our purposes it will be sufficient to restrict to sequences converging to x .

Lemma 2.7. *Let x be a point of $X(\bar{k})$ and K any finite extension of k over which x is defined. Let U be any open affine subset of $X_K := X \times_k K$ containing x . Let u_1, \dots, u_r be any elements of $\Gamma(U, \mathcal{O}_{X_K})$ which generate the maximal ideal of x . Then for any sequence $\{x_i\}$ of points of $U(K_v)$ such that $d_v(x, x_i) \rightarrow 0$ as $i \rightarrow \infty$ the functions $d_v(x, \cdot)$ and $\max(\|u_1(\cdot)\|_v, \dots, \|u_r(\cdot)\|_v)$ are equivalent on $\{x_i\}$. In other words, there are positive constants $c < C$ such that for all $i \geq 0$ we have*

$$c d_v(x, x_i) \leq \max(\|u_1(x_i)\|_v, \dots, \|u_m(x_i)\|_v) \leq C d_v(x, x_i).$$

Proof: By Lemma 2.6 there is an affine open set U' containing x and functions u'_1, \dots, u'_s generating the maximal ideal of x (on U') such that $\min(1, \max(\|u'_1(\cdot)\|_v, \dots, \|u'_s(\cdot)\|_v))$ is equivalent to $d_v(x, \cdot)$ on U' . By replacing U and U' with their intersection and possibly omitting initial members of the sequence we may assume $U = U'$. Since $d_v(x, x_i) \rightarrow 0$ as $i \rightarrow \infty$, each of the $\|u'_j(x_i)\|_v \rightarrow 0$ as well, and so $d_v(x, \cdot)$ is equivalent to $\max(\|u'_1(\cdot)\|_v, \dots, \|u'_s(\cdot)\|_v)$ on $\{x_i\}$. It therefore suffices to show the equivalence of $\max(\|u_1(\cdot)\|_v, \dots, \|u_r(\cdot)\|_v)$ and $\max(\|u'_1(\cdot)\|_v, \dots, \|u'_s(\cdot)\|_v)$. Since $d_v(x, x_i) \rightarrow 0$ as $i \rightarrow \infty$, for large enough i the points x_i are contained in a compact neighbourhood of x in $U(K_v)$. Thus the equivalence follows by Lemma 2.3. \square

One warning: Lemma 2.7, with the freedom to choose U and u_1, \dots, u_r does not hold in the generality of Lemma 2.6, i.e., for all $y \in U(K_v)$. As an elementary example, let $k = \mathbb{Q}$, $K = \mathbb{Q}(\sqrt{2})$, $X = \mathbb{P}^1$, $x = [-\sqrt{2}: 1]$, and let v be an extension of the archimedean absolute value on \mathbb{Q} . Let $\mathbb{A}_K^1 = \text{Spec}(K[t])$ be the open affine subset of \mathbb{P}^1 obtained by removing $[1:0]$, and U the open subset of \mathbb{A}_K^1 obtained by removing the point $[\sqrt{2}: 1]$. Then on U the function $u_1 = t^2 - 2$ generates the maximal ideal of x . Pick a sequence of points $\{x_i\}$ in $\mathbb{A}^1(\mathbb{Q})$ converging to $[\sqrt{2}: 1]$. Then $d_v(x, x_i)$ does not go to zero as $i \rightarrow \infty$, while $\|u_1(x_i)\|_v$ does, so the two functions are not equivalent.

The issue is fairly clear. The function $u'_1 = t + \sqrt{2}$ also generates the maximal ideal of x on U , and $\min(1, \|u'_1(\cdot)\|_v)$ is equivalent to the distance function $d_v(x, \cdot)$ on $U(\mathbb{R})$. The problem is that although u_1 and u'_1 satisfy the relation $u'_1 = \frac{1}{t-\sqrt{2}}u_1$ we cannot deduce that $u'_1(x_i) \rightarrow 0$ as $i \rightarrow \infty$ from the fact that $u_1(x_i) \rightarrow 0$ because the function $\frac{1}{t-\sqrt{2}}$ is unbounded on $\{x_i\}$.

Lemma 2.6 gives one justification that the distance functions chosen are reasonable: they are globally defined functions which locally, around any point x , behave like the standard v -adic distance functions induced from an embedding into an affine space. The following

discussion connecting the non-archimedean distance functions to order of contact provides another justification. This discussion is not necessary for any of the arguments in the paper, but is included to provide further geometric intuition behind the definition.

Geometric meaning of non-archimedean distance. The formula in (2.1.2) is a compact way of stating a very concrete notion of v -adic distance: points x and y are close if the corresponding curves in an integral model of X have high order of contact at the place v .

To see this, we will define a distance function, $d'_v(\cdot, \cdot)$ via order of contact, suitably normalized, and show that it equals $d_v(\cdot, \cdot)$. Let \mathfrak{X} be the projective integral model of X over $\text{Spec}(\mathcal{O}_k)$ obtained by taking the closure of X in $\mathbb{P}^m_{\mathcal{O}_k}$, under the inclusions $X \hookrightarrow \mathbb{P}^m \hookrightarrow \mathbb{P}^m_{\mathcal{O}_k}$. Let $\widehat{\mathcal{O}}_{k,v}$ be the completion of \mathcal{O}_k at the maximal ideal corresponding to v , and set \mathfrak{X}_v to be the base-change of \mathfrak{X} to $\widehat{\mathcal{O}}_{k,v}$.

Suppose that $x, y \in X(k)$, then x and y give rise to sections σ_x and σ_y of \mathfrak{X}_v over $\text{Spec}(\widehat{\mathcal{O}}_{k,v})$. If $x = y$ we set $d'_v(x, y) = 0$. If $x \neq y$, then let Z be the scheme of intersection of σ_x and σ_y in \mathfrak{X}_v . The ring $\Gamma(Z, \mathcal{O}_Z)$ of global sections of the structure sheaf of Z has finitely many elements, and we set $d'_v(x, y) = 1/(\#\Gamma(Z, \mathcal{O}_Z))$ where $\#$ denotes the number of elements in the ring. Note that if Z is empty then $\Gamma(Z, \mathcal{O}_Z)$ is the zero ring with a single element (namely 0). I.e., if $Z = \emptyset$ then $d'_v(x, y) = 1$.

In the general case that $x, y \in X(\bar{k})$, let F/k be any finite extension so that x and y are defined over F and set $\mathfrak{X}_{F,v}$ to be the base change of \mathfrak{X} to $\text{Spec}(\widehat{\mathcal{O}}_{F,v})$, where $\widehat{\mathcal{O}}_{F,v}$ is the completion of \mathcal{O}_F at v . As before, x and y give rise to sections σ_x and σ_y of $\mathfrak{X}_{F,v}$ over $\text{Spec}(\widehat{\mathcal{O}}_{F,v})$. If $x = y$ then set $d'_v(x, y) = 0$. Otherwise let Z be the scheme of intersection, and set $d'_v(x, y) = 1/(\#\Gamma(Z, \mathcal{O}_Z))^{\frac{1}{[F_v:k_v]}}$. If F'/F is any finite extension, and Z' the scheme of intersection of the corresponding sections σ'_x and σ'_y of $\mathfrak{X}_{F',v}$ then $\#\Gamma(Z', \mathcal{O}_{Z'}) = \#\Gamma(Z, \mathcal{O}_Z)^{[F'_v:F_v]}$. It follows that d'_v is well defined.

To see that $d'_v(\cdot, \cdot)$ is equal to $d_v(\cdot, \cdot)$ we make the following observations: (1) Since both functions transform in the same way when extending the field, we may assume that x and y are defined over k . (2) The section σ_x is obtained by multiplying the coordinates of x by an element in k so that all coordinates are in $\mathcal{O}_{k,v}$ and such that at least one coordinate is not in the maximal ideal corresponding to v . After multiplying, we have $\max_{0 \leq i \leq m} (\|x_i\|_v) = 1$. Similarly, we may assume that the section σ_y is given by $[y_0 : \dots : y_m]$ and that $\max_{0 \leq j \leq m} (\|y_j\|_v) = 1$. (3) The diagonal of $\mathbb{P}^m \times \mathbb{P}^m$ is cut out by the equations $\{X_i Y_j - X_j Y_i\}$ for $0 \leq i < j \leq m$, where X_0, \dots, X_m and Y_0, \dots, Y_m are the coordinates on the product. Thus the ideal of $\widehat{\mathcal{O}}_{F,v}$ generated by $\{x_i y_j - x_j y_i\}_{0 \leq i < j \leq m}$ is the ideal of the scheme of intersection Z . (4) Our normalization for $\|\cdot\|_v$ now shows that $\max_{0 \leq i < j \leq m} (\|x_i y_j - x_j y_i\|_v) = 1/\#\Gamma(Z, \mathcal{O}_Z)$, i.e., that $d'_v(x, y) = d_v(x, y)$.

Height Functions. A height function is a function $H: X(\bar{k}) \rightarrow \mathbb{R}_{>0}$. Two height functions H and H' are equivalent if there are positive real constants c and C with $0 < c \leq C$ such that

$$cH(x) \leq H'(x) \leq CH(x)$$

for all $x \in X(\bar{k})$ (see also “Notations and Conventions” in the introduction). The set of height functions forms a group under multiplication and the group operation descends to equivalence classes of height functions.

For any line bundle L on X we may associate a height function H_L , well defined up to equivalence, in such a way that the map from $\text{Pic}(X)$ to the equivalence classes of height functions is a group homomorphism and the height function is functorial with respect to

pullbacks. For details on how to do this, see for example any one of [1, Chap. 2], [8, Part B], [10, Chap. III], or [19, Chap. 2]. One caveat: the normalizations used in these references are not all the same. In this paper we normalize our height functions so that for a point $x = [x_0 : \cdots : x_r] \in \mathbb{P}^r(k)$, the height with respect to $\mathcal{O}_{\mathbb{P}^r}(1)$ is

$$H(x) = \prod_v \max(\|x_0\|_v, \dots, \|x_r\|_v)$$

where the product ranges over all the places v of k . Unless otherwise specified all height functions in this paper are multiplicative and relative to k .

Approximation Constants. We now define the main objects of study in this paper, inspired by similar definitions from [14]. We fix a single place v , archimedean or non-archimedean, and a corresponding distance function d_v as described above.

Definition 2.8. *Let X be a projective variety, $x \in X(\bar{k})$, L a line bundle on X . For any sequence $\{x_i\} \subseteq X(k)$ of distinct points with $d_v(x, x_i) \rightarrow 0$ (which we denote by $\{x_i\} \rightarrow x$), we set*

$$A(\{x_i\}, L) = \left\{ \gamma \in \mathbb{R} \mid d_v(x, x_i)^\gamma H_L(x_i) \text{ is bounded from above} \right\}.$$

If $\{x_i\}$ does not converge to x then we set $A(\{x_i\}, L) = \emptyset$.

Remarks. (a) It follows easily from the definition that if $A(\{x_i\}, L)$ is nonempty then it is an interval unbounded to the right, i.e., if $\gamma \in A(\{x_i\}, L)$ then $\gamma + \delta \in A(\{x_i\}, L)$ for any $\delta > 0$.

(b) If $\{x'_i\}$ is a subsequence of $\{x_i\}$ then $A(\{x_i\}, L) \subseteq A(\{x'_i\}, L)$.

Definition 2.9. *For any sequence $\{x_i\}$ we set $\alpha_x(\{x_i\}, L)$ to be the infimum of $A(\{x_i\}, L)$ (in particular if $A(\{x_i\}, L) = \emptyset$ then $\alpha_x(\{x_i\}, L) = \infty$). We call $\alpha_x(\{x_i\}, L)$ the approximation constant of $\{x_i\}$ with respect to L .*

It follows immediately from the definition that for any $\delta > 0$, $d_v(x, x_i)^{\alpha_x(\{x_i\}, L) + \delta} H_L(x_i) \rightarrow 0$ as $i \rightarrow \infty$ whenever α_x is finite. We will frequently use this fact. By remark (b) above, if $\{x'_i\}$ is a subsequence of $\{x_i\}$ then $\alpha_x(\{x'_i\}, L) \leq \alpha_x(\{x_i\}, L)$.

Definition 2.10. *The approximation constant $\alpha_{x,X}(L)$ of x with respect to L is defined to be the infimum of all approximation constants of sequences of points in $X(k)$ converging to x . If no such sequence exists, we set $\alpha_{x,X}(L) = \infty$.*

Remarks. (a) The asymptotics of the approximation are unchanged if we replace the distance and height functions by equivalent ones. Since the approximation constant α_x is local to x , we are also free to replace the distance function by one which is only equivalent to d_v in some open set (in the analytic, v -adic, or Zariski topology) around x without changing α_x . In particular, by Proposition 2.5 the definition of α_x does not depend on the choice of projective embedding used to define d_v .

(b) Slightly more generally, two height functions H and H' are called quasi-equivalent if for every $\delta > 0$ there exist $0 < c < C$ (depending on δ) so that

$$cH^{1-\delta} \leq H' \leq CH^{1+\delta}.$$

The definitions of $\alpha_x(\{x_i\}, L)$ and $\alpha_x(L)$ only depend on the quasi-equivalence class of the height function. For ample L and any $M \in \text{Pic}^0(X)$, the heights H_L and $H_{L \otimes M}$ are quasi-equivalent (see [19, p. 26]; the proof also applies to singular varieties). For ample L , the

functions $\alpha_x(L)$ and $\alpha_x(\{x_i\}, L)$ therefore only depend on the class of L in $\text{Pic}(X)/\text{Pic}^0(X)$, i.e., on the class of L in the Néron-Severi group.

(c) If L is ample, then there exists $c > 0$ so that $H_L(x_i) \geq c$ for all $x_i \in X(k)$. Thus if the sequence $d_v(x, x_i)^\gamma H_L(x_i)$ is bounded we must have $\gamma \geq 0$. We therefore conclude that $\alpha_x(L) \geq 0$. In Proposition 2.15(d) we will show the slightly stronger statement $\alpha_x(L) > 0$ for ample L . Similarly, if some multiple of L is an effective divisor and x a point outside the asymptotic base locus Z of L , we can again conclude that $\alpha_x(L) \geq 0$, since again $H_L(x_i) \geq c > 0$ for all $x_i \in X(k) \setminus Z(k)$.

(d) When L is ample, $H_L(x_i)$ is a proxy for how complicated the point x_i is. The number $\alpha_x(\{x_i\}, L)$ therefore measures the cost (in terms of the growth of complexity of the approximating points) required to get closer and closer to x . Thus under this definition (for ample L) smaller approximation constants correspond to better approximating sequences.

(e) It is possible that $\alpha_x(L) = \infty$. This occurs if either there is no sequence of points in $X(k)$ converging to x , or, if for every such sequence $\{x_i\}$ the set $A(\{x_i\}, L)$ is empty. It is also possible that $\alpha_x(L) = -\infty$. This can occur in either of the ways suggested by the definition. For instance there may be one sequence $\{x_i\}$ so that $A(\{x_i\}, L) = (-\infty, \infty)$. Alternatively given any $C > 0$, there may be a sequence $\{x_i\}$ such that $\alpha_x(\{x_i\}, L) < -C$. This happens, for instance on \mathbb{P}^n with $L = \mathcal{O}_{\mathbb{P}^n}(-1)$ and $x \in \mathbb{P}^n(k)$. See later comments and examples for more on these extreme situations.

(f) The definition given above is different from the definition of the ‘‘approximation constant’’ given in [14], since it is the infimum of the set described above rather than the minimum, as in [14]. In [14] this difference is not important to the results, since in all examples that appear in that paper, the minimum exists and is equal to the infimum.

More significantly, the distance function used in [14] is computed with respect to all of the archimedean places of k , rather than a single archimedean or non-archimedean place, and is not normalized by local degree. Thus, when $k = \mathbb{Q}$ and we choose the archimedean place, this is no difference at all, but in general the distance functions will be different. Where necessary, we will reprove results from [14] using the new definitions.

We next give an alternate characterization of $\alpha_x(L)$, valid for those line bundles whose heights satisfy the Northcott property, similar to the usual definition of the approximation constant on the affine line. Recall that a line bundle L has the Northcott property if for any constant $c \in \mathbb{R}$, the set of points $y \in X(k)$ such that $H_L(y) \leq c$ is finite. Note in particular that every ample line bundle has the Northcott property.

Definition 2.11. *For any point $x \in X(\bar{k})$ and any line bundle L we set*

$$B_x(L) = \left\{ \gamma \in \mathbb{R}_{\geq 0} \mid \begin{array}{l} \text{for all } C > 0 \text{ the number of } x_i \in X(k) \text{ such} \\ \text{that } d_v(x, x_i)^\gamma H_L(x_i) < C \text{ is finite.} \end{array} \right\}$$

Remarks.

- (a) $0 \in B_x(L)$ if and only if L has the Northcott property.
- (b) $B_x(L) \neq \emptyset$ if and only if L has the Northcott property.
- (c) $B_x(L)$ (if nonempty) is an interval: if $\gamma \in B_x(L)$ then $\gamma - \delta \in B_x(L)$ for all $0 \leq \delta \leq \gamma$.

Part (a) is obvious from the definition. For part (b), if L has the Northcott property then $B_x(L)$ is nonempty by (a). If L does not have the Northcott property then there is a constant C so that the number of $x_i \in X(k)$ with $H_L(x_i) < C$ is infinite. Since $d_v(x, x_i) \leq 1$,

for any $\gamma > 0$ these infinitely many x_i also satisfy

$$d_v(x, x_i)^\gamma H_L(x_i) \leq H_L(x_i) < C$$

and therefore $\gamma \notin B_x(L)$. Thus $B_x(L)$ is empty. Part (c) follows by again using the fact that $d_v(x, x_i)$ is bounded.

We remark that there are line bundles which have the Northcott property but which are not ample. For instance, let X be the blowup of \mathbb{P}^2 at the base locus of a k -rational pencil of plane curves of genus at least three. There is a morphism $\pi: X \rightarrow \mathbb{P}^1$ whose fibres are exactly the curves in the pencil. If the pencil is chosen so that the singular fibres all have a single nodal singularity and the curves in the pencil intersect transversely at smooth points (as is the case for a generic pencil), then every fibre of π contains finitely many k -rational points, by Faltings' Theorem. Thus, the height associated to the nef line bundle $\pi^* \mathcal{O}_{\mathbb{P}^1}(1)$ satisfies the Northcott property, but is not ample.

Proposition 2.12. *Suppose that L has the Northcott property. Then $\alpha_x(L) = \sup(B_x(L))$.*

Proof: Set $\alpha_x = \alpha_x(L)$ and $b_x = \sup(B_x(L))$. By definition of α_x , for any $\delta > 0$ there exists a sequence $\{x_i\}$ such that $\alpha_x(\{x_i\}, L) < \alpha_x + \delta$ and hence (by the definition of $\alpha_x(\{x_i\}, L)$) we conclude that $d_v(x, x_i)^{\alpha_x + \delta} H_L(x_i)$ is bounded. Therefore $\alpha_x + \delta \notin B_x(L)$ and so $\alpha_x + \delta \geq b_x$. Letting δ go to zero we conclude $\alpha_x \geq b_x$.

On the other hand, by the definition of b_x , for any $\delta > 0$ there is a C such that there are infinitely many solutions $x_i \in X(k)$ to $d_v(x, x_i)^{b_x + \delta} H_L(x_i) < C$. Since L has the Northcott property, the set of heights $H_L(x_i)$ must be unbounded, and we can therefore choose a subsequence $\{x_i\}$ of these points so that $H_L(x_i) \rightarrow \infty$ as $i \rightarrow \infty$. By the boundedness of the product, we conclude that $d_v(x, x_i) \rightarrow 0$, and so $\{x_i\}$ converges to x . But then

$$d_v(x, x_i)^{b_x + 2\delta} H_L(x_i) < C \cdot d_v(x, x_i)^\delta \rightarrow 0$$

and so $b_x + 2\delta \in A(\{x_i\}, L)$. Thus $b_x + 2\delta \geq \alpha_x$, and letting δ go to zero we conclude that $b_x \geq \alpha_x$ and so $\alpha_x = b_x$. \square

Remark. If L has the Northcott property then $0 \in B_x(L)$ and hence $\alpha_x(L) = \sup(B_x(L)) \geq 0$ by Proposition 2.12. In particular this shows again that for ample bundles $\alpha_x(L) \geq 0$.

It will be useful to know how the approximation constant changes when we change the field k . We use the notation that for an extension field K/k , $\alpha_x(\{x_i\}, L)_K$ (respectively $\alpha_x(L)_K$) denotes the approximation constant of a sequence (resp. point x) computed with respect to K . This means that when computing α , we use the height H_L relative to K and normalize d_v relative to K . If $d = [K:k]$ and $m_v = [K_v:k_v]$ then this means simply that $H_L(x_i)_K = H_L(x_i)_k^d$ and $d_v(x, x_i)_K = d_v(x, x_i)_k^{m_v}$.

Proposition 2.13. *Suppose $x \in X(\bar{k})$, L a line bundle on X , and $\{x_i\} \rightarrow x$ a sequence of points in $X(k)$ approximating x . Let K be any finite extension of k . Then $\{x_i\} \rightarrow x$ can also be considered to be a set of points of $X(K)$ approximating x . Set $m_v = [K_v:k_v]$, and let $d = [K:k]$. Then*

$$\alpha_x(\{x_i\}, L)_K = \frac{d}{m_v} \alpha_x(\{x_i\}, L)_k.$$

In particular, we have the bound $\alpha_x(L)_K \leq \frac{d}{m_v} \alpha_x(L)_k$.

Proof: The claim that $\alpha_x(\{x_i\}, L)_K = \frac{d}{m_v} \alpha_x(\{x_i\}, L)_k$ follows immediately from the equalities $H_L(\cdot)_K = H_L(\cdot)_k^d$ and $d_v(\cdot, \cdot)_K = d_v(\cdot, \cdot)_k^{m_v}$. The inequality $\alpha_x(L)_K \leq \frac{d}{m_v} \alpha_x(L)_k$ then

follows since the sequences of k -points approximating x are a subset of the sequences of K -points approximating x . \square

Basic properties of α . We start by computing α when $x \in \mathbb{P}^n(k)$.

Lemma 2.14. *Let x be any k -point of \mathbb{P}^n . Then $\alpha_{x, \mathbb{P}^n}(\mathcal{O}_{\mathbb{P}^n}(1)) = 1$.*

Proof: Without loss of generality, we may assume that $x = [1 : 0 : \dots : 0]$. We first show that $\alpha_x(\{x_i\}, \mathcal{O}_{\mathbb{P}^n}(1)) \geq 1$ for all sequences $\{x_i\}$ of k -points. Let Z_0, \dots, Z_n be the coordinates on \mathbb{P}^n and $\{x_i\}$ a sequence of k -points converging to x . Since $d_v(x, x_i) \rightarrow 0$ as $i \rightarrow \infty$ we conclude that $\|Z_j(x_i)/Z_0(x_i)\|_v \rightarrow 0$ for each $j = 1, \dots, n$. By passing to a subsequence of the x_i , which can only possibly lower the value of α , we may assume that for all i we have that $\|Z_0(x_i)\|_v$ is the largest of the $\|Z_j(x_i)\|_v$ and that there is a fixed $j \in \{1, \dots, n\}$ so that $\max(\|Z_1(x_i)\|_v, \dots, \|Z_n(x_i)\|_v) = \|Z_j(x_i)\|_v$. By Lemma 2.7 we then have $d_v(x, x_i) = \|Z_j(x_i)/Z_0(x_i)\|_v$ for all i (at least up to equivalence). Thus, for any $\gamma \geq 0$

$$\begin{aligned} d_v(x, x_i)^\gamma H(x_i) &= \left(\frac{\|Z_j(x_i)\|_v}{\|Z_0(x_i)\|_v} \right)^\gamma \cdot \|Z_0(x_i)\|_v \cdot \prod_{w \neq v} \max(\|Z_0(x_i)\|_w, \dots, \|Z_n(x_i)\|_w) \\ &\geq \left(\frac{\|Z_j(x_i)\|_v}{\|Z_0(x_i)\|_v} \right)^\gamma \cdot \|Z_0(x_i)\|_v \cdot \prod_{w \neq v} \|Z_j(x_i)\|_w = \left(\frac{\|Z_0(x_i)\|_v}{\|Z_j(x_i)\|_v} \right)^{1-\gamma}, \end{aligned}$$

where in the last step we have used the product formula. If $\gamma < 1$ then the lower bound above goes to infinity as $i \rightarrow \infty$, and hence $\alpha_x(\{x_i\}, \mathcal{O}_{\mathbb{P}^n}(1)) \geq 1$.

We next show that we can achieve $\alpha = 1$. Since we can always choose to approximate along a rational line containing x it suffices to treat the case $n = 1$ and approximate the point $[1:0]$. We will handle the archimedean and non-archimedean cases separately.

In the archimedean case embed \mathcal{O}_k as a lattice in the Minkowski space $\prod_{w \text{ arch}} k_w = \mathbb{R}^r \times \mathbb{C}^s$. For any $D > 0$, there are infinitely many elements of \mathcal{O}_k that lie in the cylinder $\{b \in \mathcal{O}_k \mid \|b\|_w \leq D \text{ for } w \neq v\}$. These elements b_i satisfy $H([b_i : 1])d_v([1 : 0], [b_i : 1]) \leq D^{r+2s}$, and so for the sequence $x_i = [b_i : 1]$ we conclude that $\alpha_x(\{x_i\}, \mathcal{O}_{\mathbb{P}^1}(1)) \leq 1$, and therefore that $\alpha_x(\{x_i\}, \mathcal{O}_{\mathbb{P}^1}(1)) = 1$.

In the non-archimedean case, since the ideal class group is finite some power of the maximal ideal corresponding to v is principal, generated by $b \in \mathcal{O}_k$. Thus we have $\|b\|_v < 1$ and $\|b\|_w = 1$ for all other finite places w of k . After taking a further power of b , and multiplying by a suitably chosen unit, we may suppose in addition that $\|b\|_w > 1$ for all infinite places w . Set $x_i = [1:b^i]$ for $i \geq 0$. Then $H(x_i) = \prod_{w \text{ arch}} \|b^i\|_w = 1/\|b\|_v^i$, where the last equality follows from the product formula. Since $d_v(x, x_i) = \|b\|_v^i$, it is clear that $\alpha_x(\{x_i\}, \mathcal{O}_{\mathbb{P}^1}(1)) = 1$ for this sequence. \square

The next proposition collects some elementary properties of α .

Proposition 2.15. *Let X and Y be projective varieties over $\text{Spec}(k)$, $x \in X(\bar{k})$, and L a line bundle on X .*

- (a) *For any positive integer m , $\alpha_{x, X}(m \cdot L) = m \cdot \alpha_{x, X}(L)$. This allows an extension of the definition of $\alpha_{x, X}(L)$ to \mathbb{Q} -divisors L .*
- (b) *α_x is a concave function of L : for any positive rational numbers a and b , and any \mathbb{Q} -divisors L_1 and L_2 (with the exception of the case that $\{\alpha_x(L_1), \alpha_x(L_2)\} = \{-\infty, \infty\}$)*

we have

$$\alpha_x(aL_1 + bL_2) \geq a\alpha_x(L_1) + b\alpha_x(L_2).$$

- (c) If Z is a subvariety of X then for any point $z \in Z(\bar{k})$ we have $\alpha_{z,Z}(L|_Z) \geq \alpha_{z,X}(L)$.
 (d) If $x \in X(k)$ and L is very ample then $\alpha_{x,X}(L) \geq 1$; if $x \in X(\bar{k})$ and L is ample then $\alpha_x(L) > 0$.
 (e) Let L_X and L_Y be line bundles on X and Y which are asymptotically base point free, and $x \in X(\bar{k})$, $y \in Y(\bar{k})$. If neither x nor y are defined over k , then

$$\alpha_{x \times y, X \times Y}(L_X \boxplus L_Y) \geq \alpha_x(L_X) + \alpha_y(L_Y)$$

If x is defined over k but y is not, then

$$\alpha_{x \times y, X \times Y}(L_X \boxplus L_Y) = \alpha_{y,Y}(L_Y).$$

If x and y are both defined over k , then

$$\alpha_{x \times y, X \times Y}(L_X \boxplus L_Y) = \min\{\alpha_{x,X}(L_X), \alpha_{y,Y}(L_Y)\}.$$

- (f) Suppose that X is reducible over k and let X_1, \dots, X_r be the irreducible components (over k) containing x . Then $\alpha_{x,X}(L) = \min(\alpha_{x,X_1}(L|_{X_1}), \dots, \alpha_{x,X_r}(L|_{X_r}))$.

Proof: Since (up to equivalence) $H_{mL} = H_L^m$, part (a) follows immediately.

To simplify notation in part (b) set $\alpha_1 = \alpha_x(L_1)$ and $\alpha_2 = \alpha_x(L_2)$. We will first prove (b) under the assumption that both α_1 and α_2 are finite. We further note that in light of part (a), we may assume that $a + b = 1$.

Suppose that there is a sequence $\{x_i\}$ with $\alpha_x(\{x_i\}, aL_1 + bL_2) < a\alpha_1 + b\alpha_2$. Fix $\delta > 0$ small enough so that $a\alpha_1 + b\alpha_2 - \delta > \alpha_x(\{x_i\}, aL_1 + bL_2)$.

$$(2.15.1) \quad d_v(x, x_i)^{a\alpha_1 + b\alpha_2 - \delta} H_{aL_1 + bL_2}(x_i) = \left(d_v(x, x_i)^{\alpha_1 - \delta} H_{L_1}(x_i)\right)^a \left(d_v(x, x_i)^{\alpha_2 - \delta} H_{L_2}(x_i)\right)^b.$$

By definition of α_1 the term $d_v(x, x_i)^{\alpha_1 - \delta} H_{L_1}(x_i)$ is unbounded. Hence, by passing to a subsequence of the x_i (which can only lower the value of $\alpha_x(\{x_i\}, L)$, we can assume that $d_v(x, x_i)^{\alpha_1 - \delta} H_{L_1}(x_i) \rightarrow \infty$ as $i \rightarrow \infty$. By definition of α_2 the term $d_v(x, x_i)^{\alpha_2 - \delta} H_{L_2}(x_i)$ is also unbounded, and hence the left side of (2.15.1) is unbounded as well. This implies that $\alpha_x(\{x_i\}, aL_1 + bL_2) \geq \alpha_1 + \alpha_2 - \delta$, in contradiction to the way δ was chosen. Hence, for all sequences $\{x_i\}$ of k -points we have $\alpha_x(\{x_i\}, aL_1 + bL_2) \geq a\alpha_1 + b\alpha_2$. Taking the infimum over all sequences we conclude that $\alpha_x(aL_1 + bL_2) \geq a\alpha_1 + b\alpha_2$, which is the inequality in (b).

When one or both of α_1 and α_2 are infinite, with the exception of the case $\{\alpha_1, \alpha_2\} = \{\infty, -\infty\}$ either the resulting statement is obvious (for instance if both $\alpha_1 = \alpha_2 = -\infty$ then the bound is $\alpha_x(\{x_i\}, aL_1 + bL_2) \geq -\infty$ which is automatically true) or a minor variation of the argument above works. In the case that $\{\alpha_1, \alpha_2\} = \{\infty, -\infty\}$ then it is not possible to deduce an upper bound for $\alpha_x(\{x_i\}, aL_1 + bL_2)$ from the data given (and also not clear what the purported upper bound, of the form “ $\infty - \infty$ ” is supposed to mean).

Part (c) is simple: We may assume that the distance function on Z is the restriction of the distance function on X and that the height function on Z is the restriction of H_L to $Z(\bar{k})$. Then for any sequence $\{z_i\}$ of points of $Z(k)$ converging to z we have $\alpha_{z,Z}(\{z_i\}, L|_Z) = \alpha_{z,X}(\{z_i\}, L)$. The statement in (c) then follows from the observation that the set of k -points of Z is a subset of the set of k -points of X , and so the infimum used to define $\alpha_{z,Z}(L|_Z)$ is over a subset of the sequences used to define $\alpha_{z,X}(L)$.

For (d), if L is very ample then L induces an embedding $X \hookrightarrow \mathbb{P}^n$ in some projective space. If $x \in X(k)$ then by part (c) and Lemma 2.14 we conclude that $\alpha_{x,X}(L) \geq \alpha_{x,\mathbb{P}^n}(\mathcal{O}_{\mathbb{P}^n}(1)) = 1$. If L is ample then some multiple mL is very ample, and so if $x \in X(k)$ then $\alpha_x(L) \geq \frac{1}{m}$ by the first part of this statement and homogeneity. Finally, if $x \in X(\bar{k})$ let K be the field of definition of x . We have just established that $\alpha_x(L)_K > 0$, hence by Proposition 2.13 we have $\alpha_x(L) = \alpha_x(L)_k \geq \frac{m_v}{d} \alpha_x(L)_K > 0$.

To prove claim (e), notice that the height function with respect to $L_X \boxplus L_Y$ is the product of the height functions of L_X and L_Y . Since $d_v((x_1, y_1), (x_2, y_2)) = d_v(x_1, x_2) + d_v(y_1, y_2)$ is a distance function on $X \times Y$, we may take that as our distance function for $X \times Y$.

Let $\{(x_i, y_i)\}$ be a sequence of k -points approximating (x, y) . If $\{x_i\}$ and $\{y_i\}$ are both eventually contained in $X - \{x\}$ and $Y - \{y\}$, respectively, then by the definition of α_x and α_y , we must have

$$\alpha_{x \times y, X \times Y}(\{(x_i, y_i)\}, L_X \boxplus L_Y) \geq \alpha_x(L_X) + \alpha_y(L_Y)$$

as desired.

If $\{x_i\}$ is eventually equal to x , we get

$$\alpha_{x \times y, X \times Y}(L_X \boxplus L_Y) = \alpha_{y,Y}(L_Y).$$

Similarly, if $\{y_i\}$ is eventually equal to y , we get

$$\alpha_{x \times y, X \times Y}(L_X \boxplus L_Y) = \alpha_{x,X}(L_X).$$

To finish the proof, it remains only to note that $\{x_i\}$ and $\{y_i\}$ are sequences of k -rational points, so that $\{x_i\}$ can only be eventually the constant sequence $\{x\}$ if x is k -rational, and similarly for y .

Finally, statement (f) follows by the pigeonhole principle: if $\{x_i\}$ is a sequence approximating x , then infinitely many x_i must lie on some component X_j , and by passing to a subsequence we may assume that all x_i lie on X_j . Thus $\alpha_{x,X}(L)$ is no more than the minimum in part (f). The opposite inequality follows from part (c). \square

Remarks on extreme cases. (a) If $\alpha_x(L) = \infty$ for one line bundle then $\alpha_x(A) = \infty$ for all ample line bundles A . Indeed, for any sequence $\{x_i\}$, if $\alpha_x(\{x_i\}, L) = \infty$ then $\alpha_x(\{x_i\}, A) = \infty$ for all ample line bundles A . This follows immediately from the fact that there is some positive integer n such that $nA - L$ is effective, giving $H_{nA}(x_i) \geq H_L(x_i) + O(1)$ for all i and so $\alpha_x(\{x_i\}, nA) \geq \alpha_x(\{x_i\}, L) = \infty$. Thus, by Proposition 2.15, part (a), $\alpha_x(\{x_i\}, A) = \infty$.

(b) Assume that there is no nef line bundle L so that $\alpha_x(L) = \infty$. The concavity condition shows that α_x is a continuous function on the ample cone.

(c) If X is smooth and L is ample, then any sequence $\{x_i\}$ such that $\alpha_x(\{x_i\}, L)$ is finite must eventually lie in a fibre of the Albanese map $\pi: X \rightarrow A$. This follows from the fact that α is infinite on Abelian varieties (see Example (c) in the introduction). More precisely, let D be an ample divisor on the Albanese variety A . Then there is some positive integer n such that $L_n = nL + \pi^*D$ is ample. If $\alpha_x(\{x_i\}, \pi^*D)$ is finite, then clearly $\{x_i\}$ is eventually contained in a fibre of the Albanese map, since π does not increase distances by more than a bounded multiple. Since $L_n - \pi^*D$ is effective, this means that $\alpha_x(\{x_i\}, L_n)$ is also infinite unless $\{x_i\}$ is eventually contained in a fibre of the Albanese map. By Remark (a), this means that for any ample divisor L , $\alpha_x(\{x_i\}, L) = \infty$, unless $\{x_i\}$ is eventually contained in a fibre of the Albanese map.

Lemma 2.16. *Let d be a positive integer, $L = \mathcal{O}_{\mathbb{P}^1}(d)$, and $x \in \mathbb{P}^1(\bar{k})$. Then*

$$\alpha_x(L) = \begin{cases} \infty & \text{if } \kappa(x) \not\subseteq k_v \\ d & \text{if } \kappa(x) = k \\ \frac{d}{2} & \text{otherwise.} \end{cases}$$

Proof: If $\kappa(x) \not\subseteq k_v$ then there is no sequence of k -points converging (with respect to $d_v(\cdot, \cdot)$) to x (see the Remark on page 9), and hence $\alpha_x(L) = \infty$. If $x \in \mathbb{P}^1(k)$ then this is Lemma 2.14 and Proposition 2.15(a). If $\kappa(x) \subseteq k_v$ but $\kappa(x) \neq k$ then $\alpha_x(\mathcal{O}_{\mathbb{P}^1}(1)) \geq \frac{1}{2}$ by Roth's theorem for \mathbb{P}^1 , while $\alpha_x(\mathcal{O}_{\mathbb{P}^1}(1)) \leq \frac{1}{2}$ by a Dirichlet-type argument. (This follows, for example, from Theorem 7.8.) Thus $\alpha_x(\mathcal{O}_{\mathbb{P}^1}(1)) = \frac{1}{2}$, and so $\alpha_x(L) = \frac{d}{2}$ by Proposition 2.15(a) again. \square

Remark: Note that the cases in Lemma 2.16 depend sensitively upon the choice of extension of v_0 to \bar{k} . For example, if $\kappa(x)$ is not a Galois extension of k , then it is possible that for some choices of v on \bar{k} , k_v contains $\kappa(x)$, and for others it does not. This leads to radically different values of $\alpha_x(L)$.

Theorem 2.17. *Let C be any singular k -rational curve and $\varphi: \mathbb{P}^1 \rightarrow C$ the normalization map. Then for any ample line bundle L on C , and any $x \in C(\bar{k})$ we have the equality:*

$$\alpha_{x,C}(L) = \min_{q \in \varphi^{-1}(x)} d/r_q m_q$$

where $d = \deg(L)$, m_q is the multiplicity of the branch of C through x corresponding to q , and

$$r_q = \begin{cases} 0 & \text{if } \kappa(q) \not\subseteq k_v \\ 1 & \text{if } \kappa(q) = k \\ 2 & \text{otherwise.} \end{cases}$$

Here we use $r_q = 0$ as a shorthand for $d/r_q m_q = \infty$.

Proof: Given any sequence $\{x_i\} \rightarrow x$ then by passing to a subsequence we can assume that all x_i are on a single branch through x . More precisely, we can assume that none of the x_i are the finitely many points where φ is not an isomorphism, and that $\{\varphi^{-1}(x_i)\}$ converges (with respect to $d_v(\cdot, \cdot)$) to one of the points $q \in \varphi^{-1}(x)$. Conversely, given a sequence $\{q_i\}$ of points of $\mathbb{P}^1(k)$ converging to some q , then $\{\varphi(q_i)\}$ converges to x . Thus it suffices to study only sequences of this kind to compute $\alpha_x(L)$.

Given a sequence $\{q_i\} \rightarrow q$ we have $H_{\varphi^*L}(q_i) = H_L(\varphi(q_i))$ for all i . Furthermore since the branch corresponding to q has multiplicity m_q , locally φ is described by functions in the m_q -th power of the maximal ideal of q , and thus $d_v(x, \varphi(q_i))$ is equivalent to $d_v(q, q_i)^{m_q}$ as $i \rightarrow \infty$. Therefore, as in Proposition 2.13 we have $\alpha_x(\{\varphi(q_i)\}, L) = \frac{1}{m_q} \alpha_q(\{q_i\}, \varphi^*L)$, and the theorem then follows from Lemma 2.16. \square

Remark: This is similar to Theorem 2.8 of [14], except that it is actually correct. (The conclusion of Theorem 2.8 of [14] neglects the possibility that the r_q defined in Theorem 2.17 might not be one.) Theorem 2.17 also uses the definition of α from this paper, rather than that of [14], and generalises the results to points defined over \bar{k} .

Examples

- (a) If $X = \mathbb{P}^N$, $L = \mathcal{O}_{\mathbb{P}^N}(d)$ for some $d > 0$, then $\alpha_x(L) = d$ for all points x in $\mathbb{P}^N(k)$. This follows from Lemma 2.14 and Proposition 2.15(a).
- (b) If $X = \mathbb{P}^1 \times \mathbb{P}^1$, $L = \mathcal{O}_{\mathbb{P}^1 \times \mathbb{P}^1}(a, b)$, with $a, b \geq 0$ then $\alpha_x = \min(a, b)$ for all $x \in X(k)$. This follows immediately from Proposition 2.15(e).
- (c) Similarly if $X = \mathbb{P}^{N_1} \times \dots \times \mathbb{P}^{N_r}$, $L = \mathcal{O}_X(d_1, \dots, d_r)$ with $d_i \geq 0$ then $\alpha_x(L) = \min(d_1, \dots, d_r)$.
- (d) Taking $X = \mathbb{P}^1 \times \mathbb{P}^1$, $L_1 = L_2 = \mathcal{O}_X(2, 1)$, $L_3 = \mathcal{O}_X(1, 2)$ then example (b) gives $\alpha_x(L_i) = 1$ for $i = 1, 2, 3$, but $\alpha_x(L_1 + L_2) = 2$ and $\alpha_x(L_1 + L_3) = 3$.

Part (d) shows that there can be no formula for determining $\alpha_x(L_i + L_j)$ in terms of $\alpha_x(L_i)$ and $\alpha_x(L_j)$ alone, and that Proposition 2.15(b) is the best possible general relation of this type.

The following lemma, which we will use several times in the paper, allows us to reduce to the case of geometrically irreducible varieties when studying α .

Lemma 2.18. *Let Z be a variety over $\text{Spec}(k)$, and set Y to be the Zariski closure of the points of $Z(k)$. Then each irreducible component of Y is geometrically irreducible and for any line bundle L on Z and any $x \in Y(\bar{k})$ we have $\alpha_{x,Z}(L) = \alpha_{x,Y}(L|_Y)$.*

Proof: Let Y_1, \dots, Y_r be the irreducible components of $\bar{Y} := Y \times_k \bar{k}$; we will show that each Y_i is actually defined over k . Let Y_i be one such component. Since Y is a variety over $\text{Spec}(k)$, all $\text{Gal}(\bar{k}/k)$ conjugates of Y_i are also components of \bar{Y} . Let $I \subseteq \{1, \dots, r\}$ be the subset of indices such that each Y_j , $j \in I$, is a Galois conjugate of Y_i , and set $I' = \{1, \dots, r\} \setminus I$. Any point $y \in Z(k)$ contained in Y_i is also contained in Y_j for $j \in I$. Therefore all points of $Z(k)$ are contained in $Y' := (\bigcap_{j \in I} Y_j) \cup (\bigcup_{j' \in I'} Y_{j'})$. By construction Y' is closed and defined over k . If $I \neq \{i\}$ then Y' is a proper subset of \bar{Y} . This contradicts the construction of Y as the Zariski closure of $Z(k)$. Thus $I = \{i\}$ and so Y_i is defined over k . Finally since $Y(k) = Z(k)$, it is clear that $\alpha_{x,Z}(L) = \alpha_{x,Y}(L|_Y)$ for all line bundles L and $x \in Y(\bar{k})$. \square

3. SESHADRI CONSTANTS

In this section, we review some basic properties of Seshadri constants, first introduced and studied in [5]. Many foundational results on Seshadri constants are given in [11, chap. 5]. The Seshadri constant is purely geometric in the sense that it only depends on the base change of the variety to the algebraic closure.

Definition 3.1. *Let X be a projective variety over $\text{Spec}(k)$, x a point of $X(\bar{k})$, and L a nef line bundle on X . The Seshadri constant, $\epsilon_{x,X}(L)$, is defined to be*

$$\epsilon_{x,X}(L) := \sup \{ \gamma \geq 0 \mid \pi^* L - \gamma E \text{ is nef} \}$$

where $\pi : \tilde{X} \rightarrow X_{\bar{k}}$ is the blowup of $X_{\bar{k}} := X \times_k \bar{k}$ at x with exceptional divisor E . Here, by abuse of notation, we also use L for the base change of L to $X_{\bar{k}}$.

The Seshadri constant is defined on the level of \mathbb{Q} - or \mathbb{R} -divisors, and in the above definition $\gamma \geq 0$ is an element of \mathbb{Q} . If γ is allowed to be a real number, then the sup in the definition can be replaced by a max.

The idea behind the Seshadri constant is that it measures the local positivity of L at x . From the definition, the Seshadri constant only depends on the numerical equivalence class of L . We will often just use $\epsilon_x(L)$ or ϵ_x for $\epsilon_{x,X}(L)$ if X or L are clear from the context.

Since the Seshadri constant only depends on $X_{\bar{k}}$, for the rest of this section we assume our varieties are defined over a fixed algebraically closed field. From Definition 3.1, all of the properties of the Seshadri constant established below will hold for varieties over $\text{Spec}(k)$.

Another characterization of the Seshadri constant is given by the following.

Proposition 3.2. *Let X be a projective variety, $x \in X$, and L a nef line bundle on X , then*

$$\epsilon_{x,X}(L) = \inf_{x \in C \subseteq X} \left\{ \frac{(L \cdot C)}{\text{mult}_x(C)} \right\}$$

where the infimum is taken over all reduced irreducible curves C passing through x .

This alternate description of the Seshadri constant follows immediately from the definition that a bundle L' on a variety \tilde{X} is nef if and only if $L' \cdot C' \geq 0$ for all reduced irreducible curves C' in \tilde{X} , and the straightforward observation that if C' is the proper transform of C in the blowup, then $E \cdot C' = \text{mult}_x(C)$, and $(\pi^*L) \cdot C' = L \cdot C$.

Basic properties of ϵ . We start by computing ϵ when $X = \mathbb{P}^n$.

Lemma 3.3. *For any point $x \in \mathbb{P}^n$, $\epsilon_x(\mathcal{O}_{\mathbb{P}^n}(1)) = 1$.*

Proof: Let $\pi: \tilde{\mathbb{P}}^n \rightarrow \mathbb{P}^n$ be the blowup of \mathbb{P}^n at x . For any $\gamma > 0$ set $L_\gamma := \pi^*(\mathcal{O}_{\mathbb{P}^n}(1)) - \gamma E$. Then L_1 is base point free and defines the projection morphism $\tilde{\mathbb{P}}^n \rightarrow \mathbb{P}^{n-1}$ with fibres the proper transforms of lines in \mathbb{P}^n passing through x . Thus L_1 is nef on $\tilde{\mathbb{P}}^n$. For any such fibre the degree of L_γ on the fibre is $1 - \gamma$, hence L_1 is the boundary of the nef cone, and $\epsilon_x(\mathcal{O}_{\mathbb{P}^n}(1)) = 1$. \square

Note that Lemma 3.3 shows that if $x \in \mathbb{P}^n(k)$, then $\epsilon_x = \alpha_x$. The following proposition extends the list of similarities between ϵ and α much further.

Proposition 3.4. *Let X be a projective variety, $x \in X(\bar{k})$, and L a nef line bundle on X .*

- (a) *For any positive integer m , $\epsilon_{x,X}(m \cdot L) = m \cdot \epsilon_{x,X}(L)$. This allows an extension of the definition of $\epsilon_{x,X}(L)$ to \mathbb{Q} -divisors L .*
- (b) *ϵ_x is a concave function of L : for any positive rational numbers a and b , and any nef \mathbb{Q} -divisors L_1 and L_2*

$$\epsilon_x(aL_1 + bL_2) \geq a\epsilon_x(L_1) + b\epsilon_x(L_2).$$

- (c) *If Z is a subvariety of X then for any point $z \in Z$ we have $\epsilon_{z,Z}(L|_Z) \geq \epsilon_{z,X}(L)$.*
- (d) *If L is very ample then $\epsilon_x(L) \geq 1$, if L is ample then $\epsilon_{x,X}(L) > 0$.*
- (e) *If x and y are points of varieties X and Y , with nef line bundles L_X and L_Y then*

$$\epsilon_{x \times y, X \times Y}(L_X \boxplus L_Y) = \min(\epsilon_{x,X}(L_X), \epsilon_{y,Y}(L_Y)).$$

- (f) *Suppose that X is reducible and let X_1, \dots, X_r be the irreducible components containing x . Then $\epsilon_{x,X}(L) = \min(\epsilon_{x,X_1}(L|_{X_1}), \dots, \epsilon_{x,X_r}(L|_{X_r}))$.*

Proof: The definition implies (a) immediately. Part (b) is also clear from the definition: if $\pi^*L_1 - \epsilon_1 \cdot E$ and $\pi^*L_2 - \epsilon_2 \cdot E$ are nef on \tilde{X} , then so is $\pi^*(aL_1 + bL_2) - (a\epsilon_1 + b\epsilon_2) \cdot E = a(\pi^*(L_1) - \epsilon_1 \cdot E) + b(\pi^*(L_2) - \epsilon_2 \cdot E)$.

To prove (c), it is enough to remark that the proper transform of Z in the blow up \tilde{X} of X at z is the blow up \tilde{Z} of Z at z , and that the restriction of a nef bundle on \tilde{X} will be a nef bundle on \tilde{Z} .

For (d), if L is very ample then L induces an embedding $X \hookrightarrow \mathbb{P}^n$ in some projective space. By part (c) and Lemma 3.3 we conclude that $\epsilon_{x,X}(L) \geq \epsilon_{x,\mathbb{P}^n}(\mathcal{O}_{\mathbb{P}^n}(1)) = 1$. If L is ample then some positive multiple mL is very ample and so $\alpha_x(L) \geq \frac{1}{m}$ by the first part of this statement and homogeneity.

The proper transforms of $X \times y$ and $x \times Y$ in the blow-up of $X \times Y$ at $x \times y$ are the blowups \tilde{X} and \tilde{Y} of X at x and Y at y . This and the observation that the restriction of a nef bundle must be nef give

$$\epsilon_{x \times y, X \times Y}(L_X \boxplus L_Y) \leq \min(\epsilon_{x,X}(L_X), \epsilon_{y,Y}(L_Y)).$$

To prove the other direction, we will use the description of $\epsilon_{x \times y}$ from Proposition 3.2. Let π_X and π_Y be the projections from $X \times Y$ to X and Y and let C be any irreducible curve in $X \times Y$ passing through $x \times y$.

Let $\pi_X(C)$ be the reduced image of C . Suppose that C is not contained in a fibre of π_X . Then $\pi_X(C)$ is not equal to a point, and if d is the generic degree of the map $C \rightarrow \pi_X(C)$ we have $\pi^* L_X \cdot C = d(L_X \cdot \pi_X(C))$, and $\text{mult}_{x \times y}(C) \leq d \cdot \text{mult}_x(\pi_X(C))$.

Since ϵ_x is the Seshadri constant for L_X at x , we have

$$\epsilon_x \leq \frac{L_X \cdot \pi_X(C)}{\text{mult}_x(\pi_X(C))} \leq \frac{d(L_X \cdot \pi_X(C))}{\text{mult}_{x \times y}(C)} = \frac{\pi_X^* L_X \cdot C}{\text{mult}_{x \times y}(C)} \leq \frac{(\pi_X^* L_X + \pi_Y^* L_Y) \cdot C}{\text{mult}_{x \times y}(C)}$$

where the first inequality follows from Proposition 3.2 applied to ϵ_x , the second from the inequality on the multiplicities, and the third from the fact that $\pi^* L_Y$ is nef.

Similarly, if C is not contained in a fibre of π_Y we have the corresponding inequality with ϵ_y in place of ϵ_x . Since for any given curve C one of these must be true we have

$$\min(\epsilon_x, \epsilon_y) \leq \inf_{x \times y \in C \subseteq X \times Y} \left\{ \frac{(\pi_X^* L_X + \pi_Y^* L_Y) \cdot C}{\text{mult}_{x \times y}(C)} \right\} \stackrel{3.2}{=} \epsilon_{x \times y}$$

finishing the proof of (e).

For part (f) we use the fact that a line bundle is ample if and only if it is ample restricted to each component, and that the blow up of each X_i at x is a component of \tilde{X} . \square

Examples

- (a) If $X = \mathbb{P}^n$, $L = \mathcal{O}_{\mathbb{P}^n}(d)$ then $\epsilon_x(L) = d$ for all points x in \mathbb{P}^n . This follows from the computation for \mathbb{P}^n and $\mathcal{O}_{\mathbb{P}^n}(1)$ in Lemma 3.3 along with Proposition 3.4(a).
- (b) If $X = \mathbb{P}^1 \times \mathbb{P}^1$, $L = \mathcal{O}_{\mathbb{P}^1 \times \mathbb{P}^1}(a, b)$, with $a, b \geq 0$ then $\epsilon_x = \min(a, b)$ for all $x \in X$. This follows immediately from Proposition 3.4(e) and part (a) of the examples, but we can also prove this as follows. Let \tilde{X} be the blow up of $X = \mathbb{P}^1 \times \mathbb{P}^1$ at a point x , E the exceptional divisor and F_1 and F_2 the pullback of the class of fibres from X . The effective cone of \tilde{X} is generated by $F_1 - E$, $F_2 - E$, and E . Dually, the nef cone of \tilde{X} is generated by F_1 , F_2 and $F_1 + F_2 - E$.

Therefore for $aF_1 + bF_2 - \gamma E$ to be in the nef cone, the condition is exactly that $\gamma \leq \min(a, b)$, i.e., $\epsilon_x(aF_1 + bF_2) = \min(a, b)$.

- (c) Similarly if $X = \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_r}$, $L = \mathcal{O}(d_1, \dots, d_r)$ with $d_i \geq 0$, for $i = 1, \dots, r$ then $\epsilon_x(L) = \min(d_1, \dots, d_r)$.

As evidenced by our parallel statements in Proposition 2.15 and Proposition 3.4 (and Lemmas 2.14 and 3.3, and the examples) there is a great deal of formal similarity between α_x and ϵ_x . See the discussion below on the Arakelov point of view for some heuristic reasons for this similarity.

For future reference we record the exact conditions on a curve C and point $x \in C(\bar{k})$ so that $\alpha_x(L) = \frac{1}{2}\epsilon_x(L)$.

Lemma 3.5. *Let C be an irreducible curve over $\text{Spec}(k)$, $x \in C(\bar{k})$ and L any ample line bundle on C . Then $\alpha_x(L) = \frac{1}{2}\epsilon_x(L)$ if and only if C is a k -rational curve, C is unibranch at x , $\kappa(x) \neq k$, and $\kappa(x) \subseteq k_v$.*

Proof: Since $\epsilon_x(L)$ is always finite, the equality implies that $\alpha_x(L)$ is finite, and hence that C is a k -rational curve. Let $\varphi: \mathbb{P}^1 \rightarrow C$ be the normalization map, and for any $q \in \varphi^{-1}(x)$ let m_q be the multiplicity at x of the branch corresponding to q , and define r_q as in Theorem 2.17. By that theorem we have $\alpha_x(L) = \min_{q \in \varphi^{-1}(x)} \left\{ \frac{d}{r_q m_q} \right\}$ where $d = \deg(L)$. By the definition of the Seshadri constant we have $\epsilon_x(L) = \frac{d}{\text{mult}_x C} = \frac{d}{\sum_{q \in \varphi^{-1}(x)} m_q}$. Thus the equality $\alpha_x(L) = \frac{1}{2}\epsilon_x(L)$ amounts to the equality

$$\max_{q \in \varphi^{-1}(x)} \{r_q m_q\} = 2 \sum_{q \in \varphi^{-1}(x)} m_q.$$

Since $r_q \in \{0, 1, 2\}$ for each q , the only possible way to have equality above is if $\varphi^{-1}(x)$ consists of a single point q with $r_q = 2$. Given the definition of r_q in Theorem 2.17 this proves the lemma. \square

Arakelov point of view. For the rest of this section we discuss some parallels between α and ϵ from the point of view of Arakelov theory. Although it does not explain those parallels, we think that this heuristic discussion is useful.

Let X be a projective variety over $\text{Spec}(k)$ and x a point of $X(k)$. Let \tilde{X} be the blow up of X at x with exceptional divisor E . By Kleiman's characterization of the ample cone, the definition of the Seshadri constant $\epsilon_x = \epsilon_x(L)$ is that for any $0 < \gamma < \epsilon_x$ the set

$$\left\{ B \subseteq \tilde{X}_{\bar{k}} \mid B \text{ an irreducible curve, } (L - \gamma E) \cdot B < 0 \right\}$$

is empty, and ϵ_x is the largest number with this property.

Let $\tilde{\mathfrak{X}}$ be a proper integral model for \tilde{X} over $\text{Spec}(\mathcal{O}_k)$. We consider each point $y \in X(k)$, $y \neq x$, to also be a point of $\tilde{X}(k)$, and hence each y gives rise to a section σ_y of $\tilde{\mathfrak{X}}$ over $\text{Spec}(\mathcal{O}_k)$. Choose suitable metrizations of L and E on the archimedean places of k . By the Arakelov construction of the intersection product on $\tilde{\mathfrak{X}}$, for any $\gamma > 0$ we have

$$h_{L-\gamma E}(y) = (L - \gamma E) \cdot \sigma_y.$$

(For details on Arakelov intersection theory, see for example [20, §III.2].)

Choose an embedding $\varphi: X \hookrightarrow \mathbb{P}^r$ so that $x \mapsto [1:0:\cdots:0]$. Let Z_0, \dots, Z_r be the coordinates on \mathbb{P}^r and define functions u_i , $i = 1, \dots, r$ on the open subset U where $Z_0 \neq 0$ by $u_i = Z_i/Z_0$. For each place w of k , define a function $e_w: X(k) \rightarrow \mathbb{R}_{\geq 0}$ by

$$e_w(y) = \begin{cases} 1 & \text{if } y \notin U(k), \\ \min(1, \max(\|u_1(y)\|_w, \dots, \|u_r(y)\|_w)) & \text{if } y \in U(k). \end{cases}$$

A short local calculation (see [15, Lemma 3.1]) shows that $-h_E(y) = \sum_w \log(e_w(y))$. By Proposition 2.5 and Lemma 2.6, $e_w(\cdot)$ is equivalent to $d_w(x, \cdot)$ on $U(k)$ for each place w . Thus, up to a bounded constant which we ignore, we have

$$(3.5.1) \quad (L - \gamma E) \cdot \sigma_y = h_{L-\gamma E}(y) = h_L(y) + \gamma \log(d_v(x, y)) + \gamma \left(\sum_{w \neq v} \log(d_w(x, y)) \right).$$

By Proposition 2.12 for any ample line bundle L an equivalent description of $\alpha_x = \alpha_x(L)$ is that for any $\gamma < \alpha_x$ the set

$$(3.5.2) \quad \left\{ y \in X(k) \mid d_v(x, y)^\gamma H_L(y) < 1 \right\}$$

is finite, and α_x is the largest number with this property.² Taking log, the finiteness of (3.5.2) is equivalent to the finiteness of

$$(3.5.3) \quad \left\{ y \in X(k) \mid h_L(y) + \gamma \log(d_v(x, y)) < 0 \right\}$$

where h_L is the logarithmic height. Since the logarithmic height is only defined up to a bounded constant, “finitely many” is the best substitute for “none”, and this makes the definition of α_x look very much like the definition of ϵ_x . Equation (3.5.1) suggests an even closer parallel: that we interpret $h_L(y) + \gamma \log(d_v(x, y))$ as the intersection “ $(L - \gamma E_v) \cdot \sigma_y$ ”, where E_v is meant to indicate that we only count the local contribution of E at the place v when computing the intersection with σ_y .

From this point of view the statements in Propositions 2.15 and 3.4 have almost identical proofs. For instance, here are the arguments for the superadditivity of α_x and ϵ_x (part (b) of the respective propositions). For ϵ_x the argument is: if there are no curves B such that $(L_1 - \gamma_1 E) \cdot B < 0$ and none such that $(L_2 - \gamma_2 E) \cdot B < 0$ then there are no curves B such that $((L_1 - \gamma_1 E) + (L_2 - \gamma_2 E)) \cdot B < 0$. For α_x the argument is: if there are only finitely many $y \in X(k)$ such that $(L_1 - \gamma_1 E_v) \cdot \sigma_y < 0$, and only finitely many such that $(L_2 - \gamma_2 E_v) \cdot \sigma_y < 0$, then there are only finitely many $y \in X(k)$ such that $((L_1 - \gamma_1 E_v) + (L_2 - \gamma_2 E_v)) \cdot \sigma_y < 0$.

We finish the discussion with another connection between the two invariants. Since $d_w(x, y) \in (0, 1]$, we have $\log(d_w(x, y)) \leq 0$ and thus by (3.5.1) if $\gamma > 0$ the set (3.5.3) is contained in the set

$$(3.5.4) \quad \left\{ y \in X(k) \mid (L - \gamma E) \cdot \sigma_y < 0 \right\}.$$

If $0 < \gamma < \epsilon_x(L)$ then the line bundle $L - \gamma E$ is ample, and therefore satisfies the Northcott property. For such γ , the set (3.5.4) and hence the subset (3.5.3) is finite. This proves the inequality $\alpha_x(L) \geq \epsilon_x(L)$, a stronger equality than $\alpha_x(L) \geq \frac{1}{2}\epsilon_x(L)$, one of the main goals of this paper. However, in the discussion above we have assumed that $x \in X(k)$. For an arbitrary point $x \in X(\bar{k})$ with field of definition K this argument, along with the change of field formula of Proposition 2.13, yields the inequality $\alpha_x(L) \geq \frac{1}{[K:k]} \epsilon_x(L)$. This is the Liouville theorem mentioned in the introduction. (See also [15, Theorem 3.3] for an extension involving γ in the big cone and the asymptotic base locus.)

²The extra quantifier “ C ” in Definition 2.11 can be absorbed by the condition that the finiteness is supposed to hold for all $\gamma < \alpha_x$. The purpose of this quantifier in Definition 2.11 is to simplify arguments.

4. THE CONSTANT $\beta_x(L)$

The proof of the general version of Roth's theorem will involve a third invariant of a point and an ample line bundle. In this section we define this invariant and establish some basic facts to be used in the proof. As with the Seshadri constant this invariant only depends on the base change of the variety to an algebraically closed field. We start by describing the invariant in this case, and then give the general definition below.

First suppose that X is an irreducible n -dimensional variety defined over an algebraically closed field. For any ample line bundle L on X and $x \in X$, let $\pi: \tilde{X} \rightarrow X$ be the blow up at x with exceptional divisor E , and for any $\gamma \in \mathbb{R}_{\geq 0}$ set $L_\gamma := \pi^*L - \gamma E$.

Let $\text{NS}(\tilde{X})_{\mathbb{R}}$ be the real Néron-Severi group of \tilde{X} and let $\text{Vol}(\cdot)$ be the volume function on $\text{NS}(\tilde{X})_{\mathbb{R}}$. Recall that the *volume*, $\text{Vol}(M)$, of a line bundle M on an n -dimensional variety measures the asymptotic growth of the global sections of M . Specifically $\text{Vol}(M)$ is the unique real number so that $h^0(mM) = \frac{\text{Vol}(M)}{n!} m^n + O(m^{n-1})$ for $m \gg 0$. From the definition it follows that $\text{Vol}(mM) = m^n \text{Vol}(M)$ for $m \geq 0$, so that $\text{Vol}(\cdot)$ may be extended to \mathbb{Q} -bundles. By [11, Corollary 2.2.45] $\text{Vol}(\cdot)$ depends only on the numerical class of M and extends uniquely to a continuous function on the real Néron-Severi group. A line bundle M is called *big* if $\text{Vol}(M) \neq 0$.

Let $\gamma_{\text{eff}} = \gamma_{\text{eff},x}(L) = \sup\{\gamma \in \mathbb{R}_{\geq 0} \mid L_\gamma \text{ is effective}\}$. We note that γ_{eff} is always finite: if A is an ample bundle on \tilde{X} and L_γ effective then $L_\gamma \cdot A^{n-1} = (\varphi^*L) \cdot A^{n-1} - \gamma(E \cdot A^{n-1}) > 0$ and hence $\gamma_{\text{eff}} \leq \frac{(\varphi^*L) \cdot A^{n-1}}{E \cdot A^{n-1}}$. The big cone is the interior of the effective cone, and therefore by definition of γ_{eff} we have $\text{Vol}(L_\gamma) > 0$ for all $\gamma \in [0, \gamma_{\text{eff}})$, $\text{Vol}(L_\gamma) = 0$ for all $\gamma > \gamma_{\text{eff}}$, and so also $\text{Vol}(L_{\gamma_{\text{eff}}}) = 0$ by continuity of the volume function. We define a decreasing function (the ‘‘asymptotic relative volume function’’) $f: [0, \infty) \rightarrow [0, 1]$ by

$$f(\gamma) = \frac{\text{Vol}(L_\gamma)}{\text{Vol}(L)},$$

and note that f is supported on $[0, \gamma_{\text{eff}}]$. Finally, define³

$$\beta_x(L) = \int_0^\infty f(\gamma) d\gamma = \int_0^{\gamma_{\text{eff}}} f(\gamma) d\gamma$$

to be the area under f .

Example. Let $X = \mathbb{P}^n$, and $L = \mathcal{O}_{\mathbb{P}^n}(1)$. We will check below that for any $x \in \mathbb{P}^n$ we have $\gamma_{\text{eff},x}(L) = 1$, $f(\gamma) = 1 - \gamma^n$ on $[0, 1]$, and hence $\beta_x(L) = \frac{n}{n+1}$. This will turn out (via Theorem 6.3 or 6.1) to explain the approximation constants of $\frac{1}{2}$ for \mathbb{P}^1 (from the classical Roth's theorem) or $\frac{n}{n+1}$ for \mathbb{P}^n (from the Schmidt subspace theorem).

We now verify the claims above. As in the proof of Lemma 3.3, let $\pi: \tilde{\mathbb{P}}^n \rightarrow \mathbb{P}^n$ be the blowup of \mathbb{P}^n at x and recall that L_1 is base point free and defines a projection morphism $\varphi: \tilde{\mathbb{P}}^n \rightarrow \mathbb{P}^{n-1}$. The degree of L_γ on the fibres of φ is $1 - \gamma$. Hence for rational $\gamma > 1$, and $m > 0$ such that $m\gamma$ is an integer, any global section of mL_γ vanishes on all fibres of φ , and is therefore zero. Thus $\gamma_{\text{eff}} \leq 1$. When $0 < \gamma < 1$, L_γ is ample, and thus effective, and we conclude that $\gamma_{\text{eff}} = 1$.

³To the best of our knowledge, the number $\beta_x(L)$ was first defined by Per Salberger in unpublished work dating from 2006, where it was used to improve results of R. Heath-Brown on uniform upper bounds for the number of rational points of bounded height. Salberger also proved Corollary 4.2 as a key step in this work.

If M is a big and nef line bundle then $\text{Vol}(M) = c_1(M)^n$ (see [11, p. 148]), and therefore $\text{Vol}(L_\gamma) = c_1(L_\gamma)^n = c_1(L)^n + (-\gamma)^n E^n = 1 - \gamma^n$ on $[0, 1]$. By this formula, $\text{Vol}(L) = \text{Vol}(L_0) = 1$, and so $f(\gamma) = 1 - \gamma^n$ as claimed.

We may also compute the volume directly. Choosing an affine chart \mathbb{A}^n centered at x we can identify the global sections of $\mathcal{O}_{\mathbb{P}^n}(m)$ with polynomials in z_1, \dots, z_n of degree $\leq m$. For γ rational and m such that $m\gamma$ is an integer, the global sections of mL_γ can be identified with the subspace of those polynomials whose lowest degree term has degree at least $m\gamma$. This subspace therefore has dimension $\binom{m+n}{n} - \binom{m\gamma-1+n}{n}$. From the definition of volume we then compute that $\text{Vol}(L_\gamma) = 1 - \gamma^n$ as before.

Example. Let $X = \mathbb{P}^1 \times \mathbb{P}^1$, $L = \mathcal{O}_X(d_1, d_2)$ with $d_1, d_2 \geq 1$, and let $x \in X$ be any point. Choosing an affine chart \mathbb{A}^2 centered at x , global sections of mL may be identified with the polynomials in two variables z_1, z_2 on \mathbb{A}^2 spanned by the monomials $z_1^a z_2^b$ with $0 \leq a \leq md_1$ and $0 \leq b \leq md_2$. For rational γ and m such that $m\gamma$ is integral, the global sections of mL_γ may be identified with the subspace of these polynomials vanishing to order $\geq m\gamma$ at x , or equivalently, with the subspace spanned by the monomials $z_1^a z_2^b$ with $a+b \geq m\gamma$. For $\gamma > d_1 + d_2$ we therefore have $H^0(mL_\gamma) = 0$ for all $m > 0$, and for $\gamma < d_1 + d_2$ we have $H^0(mL_\gamma) \neq 0$ for all $m \gg 0$ (and sufficiently divisible so that $m\gamma$ is an integer). Thus $\gamma_{\text{eff},x}(L) = d_1 + d_2$.

The exponent vectors (a, b) of the monomials in $H^0(mL)$ are the lattice points in the rectangle $[0, md_1] \times [0, md_2]$, while those of the monomials in the subspace $H^0(mL_\gamma)$ are the subset of those lattice points satisfying $a + b \geq m\gamma$. Scaling the rectangle by $1/m$ and letting $m \rightarrow \infty$, we conclude that for $\gamma \in [0, d_1 + d_2]$ the ratio $f(\gamma) = \text{Vol}(L_\gamma)/\text{Vol}(L)$ is the fraction of the area of the rectangle $[0, d_1] \times [0, d_2]$ satisfying $a + b \geq \gamma$ (the shaded region shown below):

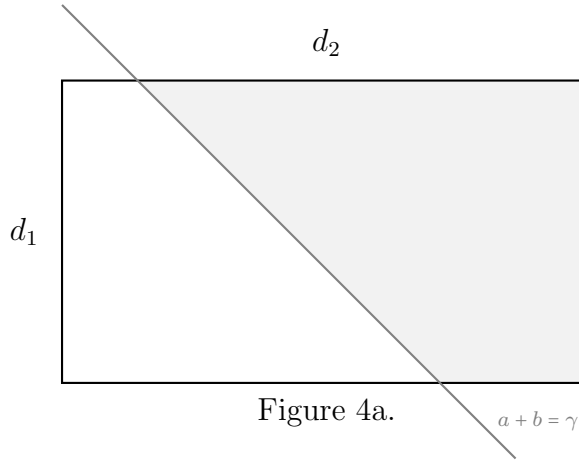


Figure 4a.

So that for any $x \in X$ (and assuming that $d_1 \leq d_2$ for the purposes of this formula)

$$f(\gamma) = \begin{cases} 1 - \frac{\gamma^2}{2d_1d_2} & \text{if } 0 \leq \gamma \leq d_1 \\ 1 + \frac{d_1}{2d_2} - \frac{\gamma}{d_2} & \text{if } d_1 \leq \gamma \leq d_2 \\ \frac{(d_1+d_2-\gamma)^2}{2d_1d_2} & \text{if } d_2 \leq \gamma \leq d_1 + d_2 \end{cases}$$

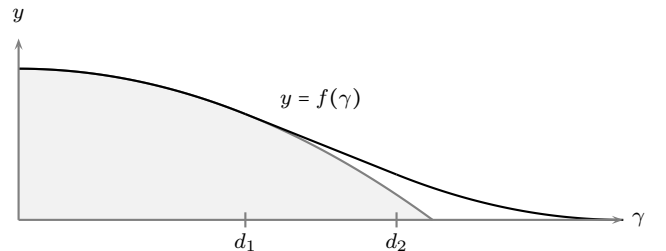


Figure 4b.

with area $\beta_x(L) = \int_0^{d_1+d_2} f(\gamma) d\gamma = \frac{d_1+d_2}{2}$. (The shaded region in Figure 4b is not connected with the shaded region in Figure 4a and will be explained below.)

Lemma 4.1. *For any ample L , $x \in X$, and $\gamma \geq 0$ we have $\text{Vol}(L_\gamma) \geq \text{Vol}(L) - (\text{mult}_x X) \cdot \gamma^n$.*

Proof: Since $\text{Vol}(\cdot)$ is a continuous function, it suffices to prove the formula for rational γ . For m large and sufficiently divisible (i.e., so that $m\gamma$ is an integer) we have the exact sequence of sheaves

$$(4.1.1) \quad 0 \longrightarrow mL_\gamma \xrightarrow{m\gamma E} mL_0 \longrightarrow mL_0|_{m\gamma E} \longrightarrow 0$$

on \tilde{X} where $m\gamma E$ is the subscheme defined by the $(m\gamma)^{\text{th}}$ power of the ideal sheaf of the Cartier divisor E , and where $L_0 = \pi^*L$. This yields an exact sequence on global sections:

$$0 \longrightarrow \Gamma(\tilde{X}, mL_\gamma) \longrightarrow \Gamma(\tilde{X}, mL_0) \longrightarrow \Gamma(m\gamma E, mL_0|_{m\gamma E}).$$

Since $h^0(mL_0) = h^0(mL) = \frac{\text{Vol}(L)}{n!}m^n + O(m^{n-1})$ the lemma will follow if we show that $h^0(mL_0|_{m\gamma E}) \leq \frac{\text{mult}_x X}{n!}(\gamma m)^n + O(m^{n-1})$.

Because L can be trivialized in a neighbourhood of x , $L_0 = \pi^*L$ is trivial in a neighbourhood of E , and hence $L_0|_{m\gamma E} = \mathcal{O}_{m\gamma E}$. Let \mathcal{J}_E be the ideal sheaf of E on \tilde{X} . For any $\ell \geq 1$ we have $\mathcal{J}_E^\ell / \mathcal{J}_E^{\ell+1} = \mathcal{O}_E(-\ell E)$, and thus the exact sequence of sheaves

$$(4.1.2) \quad 0 \longrightarrow \mathcal{O}_E(-\ell E) \longrightarrow \mathcal{O}_{(\ell+1)E} \longrightarrow \mathcal{O}_{\ell E} \longrightarrow 0.$$

This gives the inductive estimate

$$(4.1.3) \quad h^0(L_0|_{m\gamma E}) = h^0(\mathcal{O}_{m\gamma E}) \leq \sum_{\ell=0}^{m\gamma-1} h^0(\mathcal{O}_E(-\ell E)).$$

Choose an embedding $X \hookrightarrow \mathbb{P}^m$ and let $\tilde{\mathbb{P}}^m$ be the blow up of \mathbb{P}^m at the image of x , with exceptional divisor $E' \cong \mathbb{P}^{m-1}$. Then \tilde{X} is the proper transform of X in $\tilde{\mathbb{P}}^m$, and $E = \tilde{X} \cap E'$. Furthermore, E has degree $\text{mult}_x X$ as an $(n-1)$ -dimensional subvariety of $\mathbb{P}^{m-1} \cong E'$. We thus have

$$(4.1.4) \quad h^0(\mathcal{O}_E(-\ell E)) = \frac{\text{mult}_x X}{(n-1)!} \ell^{n-1} + O(\ell^{n-2}) \text{ for } \ell \gg 0$$

since $\mathcal{O}_{E'}(-E') = \mathcal{O}_{\mathbb{P}^{m-1}}(1)$, and so $h^0(\mathcal{O}_E(-\ell E)) = h^0(\mathcal{O}_{\mathbb{P}^{m-1}}(\ell)|_E)$ is simply given by the Hilbert polynomial of E for large ℓ . Summing (4.1.4) and using (4.1.3) we obtain the estimate $h^0(mL_0|_{m\gamma E}) \leq \frac{\text{mult}_x X}{n!}(\gamma m)^n + O(m^{n-1})$, proving the lemma. \square

Remark. If M is a big and nef line bundle, then $\text{Vol}(M) = c_1(M)^n$. In particular, for $\gamma \in [0, \epsilon_x(L)]$, $\text{Vol}(L_\gamma) = c_1(L_\gamma)^n = c_1(L)^n + (-\gamma)^n E^n = \text{Vol}(L) - (\text{mult}_x X) \cdot \gamma^n$, i.e., the lower bound from Lemma 4.1 is an equality on $[0, \epsilon_x(L)] \subseteq [0, \gamma_{\text{eff}}]$. In general the inequality in Lemma 4.1 is strict on $(\epsilon_x(L), \gamma_{\text{eff}}]$ (i.e., $H^0(mL_0|_{m\gamma E})$ fails to impose independent conditions on $H^0(mL_0)$ for γ in that range). As an example, the shaded region in Figure 4b shows the (normalized) lower bound $\frac{1}{\text{Vol}(L)}(\text{Vol}(L) - \gamma^2)$ in the case $X = \mathbb{P}^1 \times \mathbb{P}^1$. The lower bound is equal to $f(\gamma)$ up until $d_1 = \epsilon_x(L)$, but drops away from $f(\gamma)$ immediately after.

Corollary 4.2. *For any ample L and $x \in X$ we have $\beta_x(L) \geq \frac{n}{n+1} \sqrt[n]{\frac{\text{Vol}(L)}{\text{mult}_x X}} \geq \frac{n}{n+1} \epsilon_x(L)$. In general, both these inequalities are strict.*

Proof: Let $g(\gamma) = 1 - \frac{\text{mult}_x X}{\text{Vol}(L)} \gamma^n$ and set $\omega = \sqrt[n]{\frac{\text{Vol}(L)}{\text{mult}_x X}}$ (i.e., the solution to $g(\omega) = 0$). By Lemma 4.1 we have $f(\gamma) \geq g(\gamma)$ for all $\gamma \geq 0$. For any $\gamma \in [0, \omega)$, $g(\gamma) > 0$ and hence $f(\gamma) > 0$, so we conclude that $\omega \leq \gamma_{\text{eff}}$. Therefore

$$\beta_x(L) = \int_0^{\gamma_{\text{eff}}} f(\gamma) d\gamma \geq \int_0^{\omega} g(\gamma) d\gamma = \frac{n}{n+1} \sqrt[n]{\frac{\text{Vol}(L)}{\text{mult}_x X}}.$$

The inequality $\sqrt[n]{\frac{\text{Vol}(L)}{\text{mult}_x X}} = \sqrt[n]{\frac{c_1(L)^n}{\text{mult}_x X}} \geq \epsilon_x(L)$ is [11, Proposition 5.1.9]. In the example of $X = \mathbb{P}^1 \times \mathbb{P}^1$, $L = \mathcal{O}_X(d_1, d_2)$ (with $d_1 \leq d_2$) the inequalities are $\frac{d_1+d_2}{2} > \frac{2}{3}\sqrt{2d_1d_2} > \frac{2}{3}d_1$, i.e., all are strict. \square

We now give the definition of β in general.

Definition 4.3. *Let X be a variety over $\text{Spec}(k)$, $x \in X(\bar{k})$, and L an ample line bundle on X . Then we define*

$$\beta_x(L) = \min(\beta_{x, \bar{X}_1}(L|_{\bar{X}_1}), \dots, \beta_{x, \bar{X}_\ell}(L|_{\bar{X}_\ell})),$$

where $\bar{X}_1, \dots, \bar{X}_\ell$ are the irreducible components of $\bar{X} = X \times_k \bar{k}$ containing x .

It will be important for us that part of Corollary 4.2 holds in the general case.

Corollary 4.4. *Let X be an irreducible n -dimensional variety over $\text{Spec}(k)$. Then for any $x \in X(\bar{k})$ and any ample L we have $\beta_x(L) \geq \frac{n}{n+1} \epsilon_x(L)$. In general, this inequality is strict.*

Proof: Let $\bar{X} = X \times_k \bar{k}$ with irreducible components $\bar{X}_1, \dots, \bar{X}_\ell$. Then each component is n -dimensional, hence applying Corollary 4.2 we have $\beta_{x, \bar{X}_i}(L|_{\bar{X}_i}) \geq \frac{n}{n+1} \epsilon_{x, \bar{X}_i}(L|_{\bar{X}_i})$ for each $i = 1, \dots, \ell$. By Definition 4.3 and Proposition 3.4(f) we then conclude that $\beta_x(L) \geq \frac{n}{n+1} \epsilon_x(L)$. \square

Remark. Let X be absolutely irreducible, $x \in X(\bar{k})$ be any point and K its field of definition. Set $X_K = X \times_k K$, $\pi_K: \tilde{X}_K \rightarrow X_K$ to be the blow up of X_K at the closed point corresponding to x , and E_K to be the exceptional divisor. For any $\gamma \geq 0$ set $L_{\gamma, K} = \pi_K^* L_K - \gamma E_K$, where L_K is the base change of L to X_K . We similarly define $\tilde{X}_{\bar{k}}$, $E_{\bar{k}}$, and $L_{\gamma, \bar{k}}$. Since x is defined over K it follows that $\tilde{X}_K \times_K \bar{k} = \tilde{X}_{\bar{k}}$ and hence that $\dim_{\bar{k}} H^0(\tilde{X}_{\bar{k}}, mL_{\gamma, \bar{k}}) = \dim_K H^0(\tilde{X}_K, mL_{\gamma, K})$ for all $m > 0$ and $\gamma \geq 0$ with $m\gamma$ an integer. Thus the dimension of $mL_{\gamma, \bar{k}}$, and hence the asymptotic growth (i.e., the volume) of $L_{\gamma, \bar{k}}$ may be computed “over K ”. In particular, $\text{Vol}(L_{\gamma, K})/\text{Vol}(L) = \text{Vol}(L_{\gamma, \bar{k}})/\text{Vol}(L)$ for all $\gamma \geq 0$.

We will investigate $\beta_x(L)$ further in §9. The facts above are all we need for our application to the general versions of Roth’s theorem.

5. AN APPROXIMATION THEOREM

This section is devoted to proving Theorem 5.2 below. This theorem is the central theorem of the paper in the sense that, together with lines of reasoning common in diophantine approximation⁴ this theorem implies most of the results in §6–§8.

We fix the following notation for the rest of the section. Let S be a finite set of places of k , each extended in some way to \bar{k} . Let X be an irreducible n -dimensional variety over $\text{Spec}(k)$. For each $v \in S$ choose a point $x_v \in X(\bar{k})$, and let $d_v(\cdot, \cdot)$ be a distance function (as

⁴... as well as Propositions 2.15(f) and 3.4(c), and Corollary 4.4 ...

in §2) computed with respect to $v \in S$. We are interested in simultaneously approximating each x_v , where the distance to x_v is computed with d_v . To simplify notation, let $\alpha_v = \alpha_{x_v}$ computed with respect to d_v .

A large part of this article is concerned with the approximation constant α , and we will state the results of this section in terms of α and in terms of the usual finiteness conditions; both versions are equivalent.

Fix an ample \mathbb{Q} -bundle L . For a sequence of positive real numbers $\{R_v\}_{v \in S}$ we consider the following two equivalent conditions:

$$(5.1.1) \left\{ \begin{array}{l} \text{There is a proper subvariety } Z \text{ of } X \text{ such that for all infinite sequences } \{x_i\} \text{ of} \\ \text{distinct points of } X(k) \setminus Z(k), \text{ there is at least one } v \in S \text{ so that} \\ \alpha_v(\{x_i\}, L) \geq \frac{1}{R_v}. \end{array} \right\}$$

and

$$(5.1.2) \left\{ \begin{array}{l} \text{There is a proper subvariety } Z \text{ of } X \text{ such that for any collection } \{\delta_v\}_{v \in S} \text{ with} \\ \text{each } \delta_v > 0, \text{ there are only finitely many solutions } y \in X(k) \setminus Z(k) \text{ to} \\ d_v(x_v, y) \leq H_L(y)^{-(R_v + \delta_v)} \text{ for all } v \in S. \end{array} \right\}.$$

We think of the constants R_v as ‘‘Roth constants’’ for this approximation problem, generalizing $R = 2$ in the case $X = \mathbb{P}^1$. Although indexed by the place $v \in S$, it is the local geometry around x_v , also indexed by v , which influences the constants R_v for which (5.1.1) and (5.1.2) hold.

Theorem 5.2. *Given a collection $\{R_v\}_{v \in S}$ of positive real numbers, if*

$$(5.2.1) \quad \sum_{v \in S} \beta_{x_v}(L) R_v > 1$$

then (5.1.1) and (5.1.2) hold with respect to the collection $\{R_v\}_{v \in S}$.

It is sometimes common (e.g., as in the Schmidt subspace theorem) to switch the order of quantifiers in condition (5.1.2) and specify $\{\delta_v\}_{v \in S}$ before having to specify Z . In this case one may relax the strict inequality in (5.2.1) to allow equality. We record this variation for future use.

Corollary 5.3. *If $\{R_v\}_{v \in S}$ is a sequence of positive real numbers such that $\sum_{v \in S} \beta_{x_v}(L) R_v \geq 1$, then given any sequence $\{\delta_v\}_{v \in S}$ of positive real numbers there is a proper subvariety Z so that there are only finitely many solutions $y \in X(k) \setminus Z(k)$ to*

$$d_v(x_v, y) \leq H_L(y)^{-(R_v + \delta_v)} \text{ for all } v \in S.$$

Proof of Corollary 5.3. Given such collections $\{R_v\}_{v \in S}$ and $\{\delta_v\}_{v \in S}$ set $\delta'_v = \frac{\delta_v}{2}$ and $R'_v = R_v + \delta'_v$ for each $v \in S$. Since each $\beta_{x_v}(L) > 0$, and since each $R'_v > R_v$, we conclude that $\sum_v \beta_{x_v}(L) R'_v > 1$, and thus we may apply Theorem 5.2 to the collection $\{R'_v\}_{v \in S}$. By the theorem, there exists a $Z \subset X$ such that there are only finitely many $y \in X(k) \setminus Z(k)$ satisfying

$$d_v(x_v, y) \leq H_L(y)^{-(R'_v + \delta'_v)} \text{ for all } v \in S.$$

Since $R'_v + \delta'_v = R_v + \delta_v$ for all $v \in S$, this establishes the corollary. \square

The following slight improvement in Theorem 5.2 is useful as a first step in induction. The content is that if $\dim X = 1$ one may again relax the strict inequality in (5.2.1) to allow equality and also take $Z = \emptyset$.

Corollary 5.4. *Assume that $\dim X = 1$. If $\{R_v\}_{v \in S}$ is a sequence of positive real numbers such that $\sum_{v \in S} \beta_{x_v}(L)R_v \geq 1$, then given any sequence $\{\delta_v\}_{v \in S}$ of positive real numbers, there are only finitely many solutions $y \in X(k)$ to*

$$(5.4.1) \quad d_v(x_v, y) \leq H_L(y)^{-(R_v + \delta_v)} \text{ for all } v \in S.$$

In particular, there is no need for an exceptional subset Z .

Proof: By Corollary 5.3 there is a Z (depending on our choice of $\{\delta_v\}_{v \in S}$) so that there are only finitely many solutions $y \in X(k) \setminus Z(k)$ to (5.4.1). Since Z is of dimension zero, $Z(k)$ is finite, and so there are only finitely many $y \in X(k)$ satisfying (5.4.1). \square

We will prove Theorem 5.2 at the end of this section, after dealing with some preliminary material. The key input in the proof of the theorem is the powerful and flexible approximation theorem of Faltings-Wüstholz, which we now outline in the form we will use.

For each $v \in S$ let $K^{(v)}$ be a finite extension of k (we use this notation so that there is no confusion with K_v , the completion of a field K at v). Let L be a very ample line bundle on X and set $V = \Gamma(X, L)$. For each $v \in S$, set $V_{K^{(v)}} = V \otimes_k K^{(v)}$. We suppose that for each v we're given a decreasing filtration

$$V_{K^{(v)}} = V_{K^{(v)}}^0 \supseteq V_{K^{(v)}}^1 \supseteq V_{K^{(v)}}^2 \supseteq \cdots \supseteq V_{K^{(v)}}^{r_v} \supseteq V_{K^{(v)}}^{r_v+1} = \{0\}$$

of $K^{(v)}$ -vector spaces, and an increasing sequence $0 < c_{v,1} < c_{v,2} < \cdots < c_{v,r_v}$ of positive real numbers. For any k -subspace $W \subseteq V$ we set $W_{K^{(v)}} = W \otimes_k K^{(v)}$ and $W_{K^{(v)}}^j = V_{K^{(v)}}^j \cap W_{K^{(v)}}$ for $j = 1, \dots, r_v + 1$. We define the v -th piece of the slope, $\mu_v(W)$, by

$$\mu_v(W) = \frac{1}{\dim W} \sum_{j=1}^{r_v} c_{v,j} \dim(W_{K^{(v)}}^j / W_{K^{(v)}}^{j+1}) = \frac{1}{\dim W} \sum_{j=1}^{r_v} c_{v,j} (\dim W_{K^{(v)}}^j - \dim W_{K^{(v)}}^{j+1}).$$

Finally, we define the slope $\mu(W)$ of W to be $\mu(W) = \sum_{v \in S} \mu_v(W)$.

Although there are an infinite number of possible subspaces W , once the data of the filtration is fixed, there are only finitely many possible values for the slope. Let μ_\circ be the largest slope appearing, and among the subspaces of slope μ_\circ , let W_\circ be one of the largest dimension. A short calculation shows that if W' is a subspace with slope μ_\circ , then $W' \subseteq W_\circ$, so W_\circ is the largest subspace of slope μ_\circ both in dimension and in the partial ordering induced by inclusion. The bundle W_\circ is often called the ‘‘maximal destabilizing bundle’’, or the ‘‘first step in the Harder-Narasimhan filtration’’. We now fix W_\circ to be this subspace (rather than an arbitrary variable subspace). Note that $W_\circ \neq \{0\}$.

Given the destabilizing bundle W_\circ , set $Z = \{z \in X \mid s(z) = 0 \text{ for all } s \in W_\circ\}$. Since W_\circ is a nonzero subspace of $H^0(X, L)$, Z is a proper subvariety of X .

Next, for each $v \in S$ we fix a v -adic norm on L extending our chosen valuation v . Given a global section s of L and a point $y \in X(\bar{k})$ we denote the v -adic norm of s in the fibre at y by $|s(y)|_v$.

Choosing an affine open set U where L is trivial, each global section s may be identified with a function g_s via the trivialization. The only fact about the norm which we will need is that for any $x \in U(\bar{k})$, locally (with respect to d_v) near x the functions $|s(\cdot)|_v$ and $\|g_s(\cdot)\|_v$ are equivalent. In particular, if $d_v(x_v, y_i) \rightarrow 0$ as $i \rightarrow \infty$ then the asymptotics of $|s(y_i)|_v$ and $\|g_s(y_i)\|_v$ are the same.

Finally, for each $v \in S$ and $j \in \{1, \dots, r_v\}$ we choose a $K^{(v)}$ -basis $\{s_{v,j,\ell}\}_{\ell \in I_{v,j}}$ for $W_{K^{(v)}}^j$. With this notation, the theorem [7, Theorem 9.1] of Faltings-Wüstholz is:

Theorem 5.5. (*Faltings-Wüstholz*) *If $\mu(W_\circ) > 1$ then there are only finitely many solutions $y \in X(k) \setminus Z(k)$ such that*

$$|s_{v,j,\ell}(y)|_v < H_L(y)^{-c_{v,j}} \text{ for all } v \in S, j \in \{1, \dots, r_v\}, \ell \in I_{v,j}.$$

By definition of W_\circ we have the elementary estimate $\mu(W_\circ) \geq \mu(V)$ and we will ensure the hypothesis $\mu(W_\circ) > 1$ by simply checking that $\mu(V) > 1$. The next lemma allows us deduce $\mu(V) > 1$ from condition (5.2.1).

Lemma 5.6. *Suppose that f is a continuous function defined on an interval $[0, \gamma_{\text{eff}}]$ with $f(\gamma_{\text{eff}}) = 0$, and set $\beta = \int_0^{\gamma_{\text{eff}}} f(\gamma) d\gamma$. Given any positive real number R and any $\delta' > 0$ it is possible to choose a non-negative integer r and rational numbers $0 = \gamma_0 < \gamma_1 < \gamma_2 < \dots < \gamma_r < \gamma_{\text{eff}}$ so that, if we define c_j by $c_j = \gamma_j R$ and set $\gamma_{r+1} = \gamma_{\text{eff}}$, we have*

$$\sum_{j=1}^r c_j (f(\gamma_j) - f(\gamma_{j+1})) > \beta R - \delta'.$$

Proof: Substituting $c_j = R\gamma_j$ we have

$$\sum_{j=1}^r c_j (f(\gamma_j) - f(\gamma_{j+1})) = \sum_{j=1}^r R\gamma_j (f(\gamma_j) - f(\gamma_{j+1})) = R \left(\sum_{j=1}^r (\gamma_j - \gamma_{j-1}) f(\gamma_j) \right),$$

and we recognize the final term as R times the right-hand-sum approximation to the integral of f . By choosing r and rational $\gamma_1, \dots, \gamma_r \in (0, \gamma_{\text{eff}})$ we can clearly arrange for this approximation to be as close as we want to β . \square

Proof of Theorem 5.2: The idea of the proof is simple. For each $v \in S$ we filter the space of global sections of mL (with $m \gg 0$) by the order of vanishing at x_v . (Using sections of mL instead of L allows us to get the better estimate on the resulting slope.) Writing out what the Faltings-Wüstholz theorem gives us with respect to the resulting filtration yields Theorem 5.2. We now explicitly carry out these steps.

If $X(k)$ is not Zariski-dense, then (5.1.1) and (5.1.2) hold with $Z = \overline{X(k)}$. We may therefore assume that $X(k)$ is Zariski dense and hence by Lemma 2.18 that X is geometrically irreducible.

For each $v \in S$ we let $K^{(v)}$ be the field of definition of x_v , $\pi_v: \widetilde{X}^{(v)} \rightarrow X^{(v)}$ the blow up of $X^{(v)} = X \times_k K^{(v)}$ at the closed point corresponding to x_v . Let $E^{(v)}$ denote the exceptional divisor and for $\gamma \geq 0$ we put $L_\gamma^{(v)} = \pi_v^* L - \gamma E^{(v)}$. Then π_v , $E^{(v)}$ and $L_\gamma^{(v)}$ are all varieties over $\text{Spec}(K^{(v)})$. For $\gamma \geq 0$ set $f_v(\gamma) = \frac{\text{Vol}(L_\gamma^{(v)})}{\text{Vol}(L)}$. Since X is geometrically irreducible (and by the remark on page 28) the integral of f_v is $\beta_{x_v}(L)$. By Lemma 5.6 and the hypothesis (5.2.1), for each $v \in S$ we may choose r_v and rational $0 < \gamma_{v,1} < \dots < \gamma_{v,r_v} < \gamma_{\text{eff},x_v}(L)$ so that

$$\sum_{v \in S} \left(\sum_{j=1}^{r_v} c_{v,j} (f_v(\gamma_{v,j}) - f_v(\gamma_{v,j+1})) \right) > 1,$$

with $c_{v,j} = R_v \gamma_{v,j}$ for $v \in S$, $j = 1, \dots, r_v$. By taking m sufficiently divisible we may ensure that mL is an integral line bundle and that each $m\gamma_{v,j}$ is an integer.

For any $\gamma \geq 0$, $\dim_{K^{(v)}} \Gamma(mL_\gamma^{(v)}) / \dim_{K^{(v)}} \Gamma(mL) \rightarrow f_v(\gamma)$ as $m \rightarrow \infty$, and so by taking m sufficiently large we may also ensure that each $\dim_{K^{(v)}} \Gamma(mL_{\gamma_{v,j}}^{(v)}) / \dim_{K^{(v)}} \Gamma(mL)$ is sufficiently close to $f_v(\gamma_{j,v})$ so that

$$(5.6.1) \quad \sum_{v \in S} \frac{1}{\dim_{K^{(v)}} \Gamma(mL)} \left(\sum_{j=1}^{r_v} c_{v,j} \left(\dim_{K^{(v)}} \Gamma(mL_{\gamma_{v,j}}^{(v)}) - \dim_{K^{(v)}} \Gamma(mL_{\gamma_{v,j+1}}^{(v)}) \right) \right) > 1.$$

Set $V = \Gamma(X, mL)$ and we identify $V_{K^{(v)}}$ with $\Gamma(\tilde{X}^{(v)}, mL_0^{(v)})$ as vector spaces⁵. We give a decreasing filtration on each $V_{K^{(v)}}$ by setting $V_{K^{(v)}}^j = \Gamma(mL_{\gamma_{j,v}})$ for $j = 1, \dots, r_v$, and choose a $K^{(v)}$ -basis $\{s_{v,j,\ell}\}_{\ell \in I_{v,j}}$ for each $V_{K^{(v)}}^j$. As above we let W_\circ be the maximal destabilizing subspace and Z the base locus of the sections in W_\circ . Equation (5.6.1) is the statement that $\mu(V) > 1$, and so we conclude that $\mu(W_\circ) > 1$ too. We may therefore apply Theorem 5.5 and conclude that there are only finitely many solutions $y \in X(k) \setminus Z(k)$ to

$$(5.6.2) \quad |s_{v,j,\ell}(y)|_v^{\frac{1}{mR_v\gamma_{v,j}}} H_L(y) < 1 \text{ for all } v \in S, j = 1, \dots, r_v, \ell \in I_{v,j}.$$

Now suppose that (5.1.1) is false for this choice of Z . Then there exists a sequence $\{y_i\}$ of k -points of X , with no y_i contained in Z such that $\alpha_v(\{y_i\}, L) < \frac{1}{R_v}$ for each $v \in S$. This means that for all sufficiently small $\delta' > 0$, and each $v \in S$, $d_v(x_v, y_i)^{\frac{1}{R_v} - \delta'} H_L(y_i) \rightarrow 0$ as $i \rightarrow \infty$.

Since each $s_{v,j,\ell}$ is in $V_{K^{(v)}}^j$, each $s_{v,j,\ell}$ is in the $(m\gamma_{v,j})^{\text{th}}$ power of the maximal ideal of x_v , and so for any $\delta > 0$ and for large enough i (depending on δ) we have $|s_{v,j,\ell}(y_i)|_v \leq d_v(x_v, y_i)^{m\gamma_{v,j} - \delta}$. But then for large enough i

$$(5.6.3) \quad |s_{v,j,\ell}(y_i)|_v^{\frac{1}{mR_v\gamma_{v,j}}} H_L(y_i) \leq d_v(x_v, y_i)^{\frac{1}{R_v} - \left(\frac{\delta}{mR_v\gamma_{v,j}}\right)} H_L(y_i)$$

for all $v \in S$, $j = 1, \dots, r_v$, $\ell \in I_{j,v}$. For small enough $\delta > 0$ the right hand side of (5.6.3) tends to 0 as $i \rightarrow \infty$. This contradicts (5.6.2) and therefore assertion (5.1.1) holds. This proves Theorem 5.2. \square

6. ROTH THEOREMS

Let X be an irreducible n -dimensional variety over $\text{Spec}(k)$. In this section we present theorems giving lower bounds for $\alpha_x(L)$ independent of the field of definition of $x \in X(\bar{k})$, and in particular lower bounds in terms of $\epsilon_x(L)$. In the remaining sections of the paper we will deal with simultaneous approximation but in order to clarify the ideas we start by approximating with respect to a single place, either archimedean or non-archimedean. As in the beginning of the paper, we fix a place v_0 of k , an extension v of v_0 to \bar{k} , and compute α_x with respect to $d_v(\cdot, \cdot)$.

⁵If X is not normal, $\Gamma(X, mL) \otimes_k K^{(v)}$ may only be a proper subspace of $\Gamma(\tilde{X}^{(v)}, mL_0^{(v)})$. However, since the volume is a birational invariant, the asymptotic calculations go through without change and we omit further mention of this detail.

Theorem 6.1. *For any ample \mathbb{Q} -bundle L and any $x \in X(\bar{k})$ either*

(a) $\alpha_x(L) \geq \beta_x(L)$

or

(b) *There exists a proper subvariety $Z \subset X$, irreducible over \bar{k} , with $x \in Z(\bar{k})$ so that $\alpha_{x,X}(L) = \alpha_{x,Z}(L|_Z)$, i.e., “ $\alpha_x(L)$ is computed on a proper subvariety of X ”.*

Proof: If $\alpha_x(L) < \beta_x(L)$ then choose any $R > 0$ such that $\alpha_x(L) < \frac{1}{R} < \beta_x(L)$. Then $\beta_x(L)R > 1$ so by Theorem 5.2 in the case of a single place we conclude that there is a proper subvariety Z such that for all sequences $\{x_i\}$ of k -points with $\alpha_x(\{x_i\}, L) \leq \frac{1}{R}$, all but finitely many of the points lie in Z . We conclude that $\alpha_{x,Z}(L|_Z) = \alpha_x(L)$. To see that we may assume that Z is geometrically irreducible apply Lemma 2.18 to Z , and use Proposition 2.15(f) to replace Z by a component of the resulting variety Y . \square

By Corollary 4.4 we have $\beta_x(L) \geq \frac{n}{n+1}\epsilon_x(L)$. Thus Theorem 6.1 implies the weaker theorem:

Theorem 6.2. *(Schmidt type) Under the same hypothesis as Theorem 6.1, either*

(a) $\alpha_x(L) \geq \frac{n}{n+1}\epsilon_x(L)$

or

(b) $\alpha_x(L)$ is computed on a proper subvariety Z of X (irreducible over \bar{k} , as above).

This immediately yields

Theorem 6.3. *(Roth type) With the same hypotheses as above, $\alpha_x(L) \geq \frac{1}{2}\epsilon_x(L)$, with equality if and only if both α and ϵ are computed on a k -rational curve C such that (1) C is unibranch at x , (2) $\kappa(x) \neq k$, (3) $\kappa(x) \subset k_v$, and (4) $\epsilon_{x,C}(L|_C) = \epsilon_{x_v,X}(L)$.*

Proof of Theorem 6.3: If $\alpha_x(L) \geq \frac{n}{n+1}\epsilon_x(L)$ and $n > 1$ then this is stronger than $\alpha_x(L) \geq \frac{1}{2}\epsilon_x(L)$ so we are done. If not, then by Theorem 6.2 we pass to a smaller irreducible subvariety. Since the Seshadri constant can only go up when restricting to a subvariety (Proposition 2.15(c)), we are done by induction. Finally, in the case of equality we conclude that we must have gone all the way down to a curve C (irreducible over \bar{k}), and ϵ_x must also be computed on C , or the inequality would be strict (i.e., (4) above holds). Conditions (1), (2), and (3) then follow from Lemma 3.5.

Conversely, if C is a k -rational curve passing through x and satisfying (1), (2), and (3) above then Lemma 3.5 gives $\alpha_{x,C}(L|_C) = \frac{1}{2}\epsilon_{x,C}(L|_C)$. If in addition (4) holds then we have

$$\frac{1}{2}\epsilon_{x,X}(L) = \frac{1}{2}\epsilon_{x,C}(L|_C) = \alpha_{x,C}(L|_C) \geq \alpha_{x,X}(L),$$

where the last inequality is Proposition 2.15(c). By the first part of the theorem we always have $\alpha_{x,X}(L) \geq \frac{1}{2}\epsilon_{x,X}(L)$, and thus equality must hold. \square

Here is a form of Theorem 6.3 expressed in language closer to the usual statement of Roth's theorem.

Corollary 6.4. *With the same hypothesis as above, for any $\delta > 0$ there are only finitely many solutions $y \in X(k)$ to*

$$d_v(x, y) < H_L(y)^{-\left(\frac{2}{\epsilon_x(L)} + \delta\right)}.$$

Proof: This is immediate from Theorem 6.3 and Proposition 2.12. \square

Other variations on the deduction of Theorem 6.3 from Theorem 6.2 are possible; here are two examples.

Corollary 6.5. *If $\alpha_x(L) < \frac{m}{m+1}\epsilon_x(L)$ for some $m < n$ then $\alpha_x(L)$ is computed on a subvariety Z of dimension $< m$.*

Corollary 6.6. *If $x \in X(\bar{k})$, and no rational curve passes through x then $\alpha_x(L) \geq \frac{2}{3}\epsilon_x(L)$. Equivalently, for any $\delta > 0$ there are only finitely many solutions $y \in X(k)$ to*

$$d_v(x, y) < H_L(y)^{-\left(\frac{3}{2\epsilon_x(L)} + \delta\right)}.$$

Remark. Theorems 6.1 and 6.2 were stated for a variety X irreducible over k since if X were reducible, and $x \in X(\bar{k})$ not on a component of maximal dimension n , the estimate $\beta_x(L) \geq \frac{n}{n+1}\epsilon_x(L)$ would not hold (the volume only measures top-dimensional asymptotics). However by using Propositions 2.15(f) and 3.4(f) to reduce to the irreducible components of X it follows that Theorem 6.3 and Corollaries 6.4, 6.5, and 6.6 above still hold when X is reducible.

7. SIMULTANEOUS APPROXIMATION

In this section we apply Theorem 5.2 to study simultaneous approximation. As in §5 we let S be a finite set of places of k , each extended in some way to \bar{k} and X be an irreducible n -dimensional variety over $\text{Spec}(k)$. For each $v \in S$ we choose a point $x_v \in X(\bar{k})$, and let $d_v(\cdot, \cdot)$ be the distance function (as in §2) computed with respect to $v \in S$. Again, to simplify notation, we set α_v to be α_{x_v} computed with respect to d_v .

We are interested in understanding how well sequences of k -points can simultaneously approximate each x_v . An example of this, showing how Theorem 6.3 and Corollary 6.4 generalize to simultaneous approximation, is given in the introduction. We will also consider the case of sequences $\{x_i\}$ not contained in a subvariety Z , and obtain results along the lines of Theorem 6.2 or Corollary 6.5.

There is a general mechanism for proving such simultaneous approximation results due to Mahler. The basic idea is that these generalizations are equivalent to studying simultaneous approximations with weights. We next review these ideas, and then use Theorem 5.2 to deduce the appropriate weighted versions.

Definition 7.1. *A weighting function ξ is a function $\xi: S \rightarrow [0, 1]$ such that $\sum_{v \in S} \xi_v = 1$.*

Here, and in the rest of the paper, we use ξ_v for the value of ξ at v .

It will be useful to be able to reduce verifying a statement for infinitely many weighting functions to verifying a slightly stronger statement for only finitely many weighting functions. This is the purpose of the following lemma.

Lemma 7.2. *Let S be a finite set, and $\{\Delta'_v\}_{v \in S}$ and $\{\Delta_v\}_{v \in S}$ collections of positive real numbers with $\Delta'_v < \Delta_v$ for all $v \in S$. Then there exists a finite set Ξ of weighting functions $\xi': S \rightarrow [0, 1]$ so that given any function $\xi: S \rightarrow \mathbb{R}_{\geq 0}$ satisfying $\sum_{v \in S} \xi_v \geq 1$ there is a weighting function $\xi' \in \Xi$ satisfying $\xi'_v \Delta'_v \leq \xi_v \Delta_v$ for all $v \in S$.*

Proof: Let N be any positive integer so that $\min_{v \in S} \{\Delta_v / \Delta'_v\} - \#S/N \geq 1$, and Ξ the finite set of weighting functions $\xi': S \rightarrow [0, 1]$ such that $N\xi'_v$ is an integer for all $v \in S$ (i.e., all ξ'_v are rational with denominator dividing N). Given a function $\xi: S \rightarrow \mathbb{R}_{\geq 0}$ with $\sum_v \xi_v \geq 1$ set

$$\xi''_v = \frac{\left\lfloor \frac{N \cdot \Delta_v \cdot \xi_v}{\Delta'_v} \right\rfloor}{N} \text{ for each } v \in S.$$

Then $\xi''_v \leq \frac{\Delta_v}{\Delta'_v} \xi_v$, and so $\xi''_v \Delta'_v \leq \xi_v \Delta_v$ for each $v \in S$. Furthermore, each ξ''_v is rational and nonnegative with $N\xi''_v$ an integer. Since

$$\xi''_v \geq \frac{\frac{N \cdot \Delta_v \cdot \xi_v}{\Delta'_v} - 1}{N} = \frac{\Delta_v}{\Delta'_v} \xi_v - \frac{1}{N},$$

for each $v \in S$ we conclude that

$$\sum_{v \in S} \xi''_v \geq \sum_{v \in S} \left(\frac{\Delta_v}{\Delta'_v} \xi_v - \frac{1}{N} \right) \geq \left(\sum_{v \in S} \min \left\{ \frac{\Delta_v}{\Delta'_v} \right\} \xi_v \right) - \#S/N \geq \min_{v \in S} \left\{ \frac{\Delta_v}{\Delta'_v} \right\} - \#S/N \geq 1.$$

Therefore there exists a weighting function $\xi' \in \Xi$ with $\xi'_v \leq \xi''_v$ for all $v \in S$. \square

The following proposition shows the equivalence between statements on simultaneous approximation as in the introduction, and versions of simultaneous approximation with weights.

Proposition 7.3. *Let Z be a proper subvariety of X , and L an ample \mathbb{Q} -bundle. Then for any collection $\{R_v\}_{v \in S}$ of positive real numbers the following conditions are equivalent.*

$$(7.3.1) \left\{ \begin{array}{l} \text{For all weighting functions } \xi: S \rightarrow [0, 1] \text{ and all sequences } \{x_i\} \text{ of } X(k) \setminus Z(k) \\ \text{there is at least one } v \in S \text{ with } \xi_v \neq 0 \text{ such that } \alpha_v(\{x_i\}, L) \geq \frac{1}{R_v \xi_v}. \end{array} \right\}$$

$$(7.3.2) \quad \text{For all sequences } \{x_i\} \text{ of } X(k) \setminus Z(k), \sum_{v \in S} \frac{1}{R_v \alpha_v(\{x_i\}, L)} \leq 1.$$

$$(7.3.3) \left\{ \begin{array}{l} \text{For all weighting functions } \xi: S \rightarrow [0, 1] \text{ and any collection } \{\delta_v\}_{v \in S} \text{ of positive} \\ \text{real numbers, there are only finitely many solutions } y \in X(k) \setminus Z(k) \text{ to} \\ d_v(x_v, y)^{\frac{1}{R_v}} < H_L(y)^{-\xi_v(1+\delta_v)} \text{ for all } v \in S. \end{array} \right\}$$

$$(7.3.4) \left\{ \begin{array}{l} \text{For all } \delta > 0 \text{ there are only finitely many solutions } y \in X(k) \setminus Z(k) \text{ to} \\ \prod_{v \in S} d_v(x_v, y)^{\frac{1}{R_v}} < H_L(y)^{-(1+\delta)}. \end{array} \right\}.$$

Proof: (7.3.1) \implies (7.3.2): Given a sequence $\{x_i\}$ in $X(k) \setminus Z(k)$, set $D = \sum_{v \in S} \frac{1}{R_v \alpha_v(\{x_i\}, L)}$. If all $\alpha_v(\{x_i\}, L) = \infty$ then $D = 0$ and so the inequality in (7.3.2) holds. We may therefore assume that $D \neq 0$, i.e., that there is some $v \in S$ so that $\alpha_v(\{x_i\}, L) < \infty$. Define a weighting function by $\xi_v = \frac{1}{R_v \alpha_v(\{x_i\}, L) D}$ for each $v \in S$. By (7.3.1) there is a $v \in S$ with $\xi_v \neq 0$ so that the inequality in (7.3.1) holds. Writing out the definition of ξ_v and clearing denominators gives (7.3.2) (recall that $\alpha_v(\{x_i\}, L) > 0$ by Proposition 2.15(d)).

(7.3.2) \implies (7.3.1): If (7.3.1) is false then there is a weighting function ξ and a sequence $\{x_i\}$ in $X(k) \setminus Z(k)$ such that $\frac{1}{R_v \alpha_v(\{x_i\}, L)} > \xi_v$ for all $v \in S$ such that $\xi_v \neq 0$. Summing gives a contradiction to (7.3.2).

(7.3.3) \implies (7.3.4): Assume (7.3.4) is false and fix any $\delta > 0$. For each of the infinitely many solutions y_i in $X(k) \setminus Z(k)$ to inequality (7.3.4), define $\xi_{v,i}$ so that

$$d_v(x_v, y_i)^{\frac{1}{R_v}} = H_L(y_i)^{-\xi_{v,i}(1+\delta)}$$

for each $v \in S$. Taking the product and using the fact the y_i are solutions to the inequality in (7.3.4) we conclude that $\sum_{v \in S} \xi_{v,i} > 1$. Fix any positive δ' less than δ . Applying Lemma 7.2, with $\Delta'_v = 1 + \delta'$ and $\Delta_v = 1 + \delta$ for all $v \in S$ we obtain a finite set Ξ of weighting functions so that for any $\xi: S \rightarrow \mathbb{R}_{\geq 0}$ satisfying $\sum_v \xi_v \geq 1$, there is a $\xi' \in \Xi$ satisfying $\xi'_v(1+\delta') \leq \xi_v(1+\delta)$ for all $v \in S$. In particular, there is a $\xi'_i \in \Xi$ for each function ξ_i as above. Since Ξ is a finite set, by passing to a subsequence of $\{y_i\}$ there is a $\xi' \in \Xi$ which works for all i . Note that since L is ample, we may assume that $H_L(y_i) > 1$ for all i by omitting finitely many y_i . Choosing $\delta_v = \delta'$ for each v , we have infinitely many solutions to $d_v(x_v, y)^{\frac{1}{R_v}} < H_L(y)^{-\xi'_v(1+\delta_v)}$, for all $v \in S$, contradicting (7.3.3).

(7.3.4) \implies (7.3.1): Assume that (7.3.1) is false, so that there is a sequence $\{x_i\}$ in $X(k) \setminus Z(k)$ and a weighting function ξ such that $\alpha_v(\{x_i\}, L) < \frac{1}{R_v \xi_v}$ for each $v \in S'$, where $S' = \{v \in S \mid \xi_v \neq 0\}$. For $\delta > 0$ small enough we will still have $\alpha_v(\{x_i\}, L) < \frac{1}{R_v \xi_v(1+\delta)}$ for each $v \in S'$, and so by definition of α_v , $d_v(x_v, x_i)^{\frac{1}{R_v \xi_v(1+\delta)}} H_L(x_i) \rightarrow 0$ or equivalently $d_v(x_v, x_i)^{\frac{1}{R_v}} H_L(x_i)^{\xi_v(1+\delta)} \rightarrow 0$, as $i \rightarrow \infty$ for all $v \in S'$. Thus by omitting finitely many of the initial x_i we can make the product

$$\prod_{v \in S'} \left(d_v(x_v, x_i)^{\frac{1}{R_v}} H_L(x_i)^{\xi_v(1+\delta)} \right) = \left(\prod_{v \in S'} d_v(x_v, x_i)^{\frac{1}{R_v}} \right) H_L(x_i)^{(1+\delta)}$$

as small as desired. The product $\prod_{v \in S \setminus S'} d_v(x_v, x_i)^{\frac{1}{R_v}}$ is bounded since each distance function $d_v(\cdot, \cdot)$ is bounded. Hence after omitting finitely many of the initial x_i the rest satisfy

$$\prod_{v \in S} d_v(x_v, x_i)^{\frac{1}{R_v}} < H_L(x_i)^{-(1+\delta)}$$

contradicting (7.3.4).

(7.3.1) \implies (7.3.3): Assume that (7.3.3) is false. Then there is a weighting function ξ and a collection $\{\delta_v\}_{v \in S}$ so that the inequalities in (7.3.3) have infinitely many solutions. Let $S' = \{v \in S \mid \xi_v \neq 0\}$ and let $\{y_i\}$ be a sequence of these solutions ordered by height. Then $d_v(x_v, y_i)^{\frac{1}{R_v \xi_v(1+\delta_v)}} H_L(y_i) < 1$ for all $v \in S'$, so we conclude that $\frac{1}{R_v \xi_v(1+\delta_v)} \in A_{x_v}(\{y_i\}, L)$. Thus $\alpha_v(\{y_i\}, L) \leq \frac{1}{R_v \xi_v(1+\delta_v)} < \frac{1}{R_v \xi_v}$ for $v \in S'$, contradicting (7.3.1). \square

We now use Theorem 5.2 to establish cases where the equivalent conditions in Proposition 7.3 hold.

Theorem 7.4. *In each of the following two cases there is a proper subvariety $Z \subset X$ so that the equivalent conditions in Proposition 7.3 hold with respect to the given collection $\{R_v\}_{v \in S}$.*

- (a) Any choice of $\{R_v\}_{v \in S}$ such that $R_v > \frac{1}{\beta_{x_v}(L)}$ for each $v \in S$.
- (b) Any choice of $\{R_v\}_{v \in S}$ such that $R_v > \frac{n+1}{n \epsilon_{x_v}(L)}$ for each $v \in S$.

In the case $n = \dim X = 1$ equality in (a) and (b) is sufficient, and one may take $Z = \emptyset$.

Proof: By Corollary 4.4, $\beta_{x_v}(L) \geq \frac{n}{n+1}\epsilon_{x_v}(L)$, so the condition in (b) implies the condition in (a), and it therefore suffices to prove (a). Given such a collection $\{R_v\}_{v \in S}$ choose $\{R'_v\}_{v \in S}$ so that $R_v > R'_v > \frac{1}{\beta_{x_v}(L)}$ for each $v \in S$. Applying Lemma 7.2 with $\Delta'_v = R'_v$ and $\Delta_v = R_v$ for each $v \in S$, we obtain a finite set of weighting functions Ξ so that for any weighting function ξ there is $\xi' \in \Xi$ satisfying $\xi'_v R'_v \leq \xi_v R_v$ for all $v \in S$.

Temporarily fix $\xi' \in \Xi$ and set $S' = \{v \in S \mid \xi'_v \neq 0\}$. By our choice of R'_v we have $\sum_{v \in S'} \beta_{x_v}(L) \xi'_v R'_v > \sum_{v \in S'} \xi'_v = 1$. Applying Theorem 5.2 to the collection $\{\xi'_v R'_v\}_{v \in S'}$ we obtain a proper subvariety $Z_{\xi'}$ such that for any sequence $\{x_i\}$ in $X(k) \setminus Z_{\xi'}(k)$ there is at least one $v \in S'$ with $\alpha_v(\{x_i\}, L) \geq \frac{1}{\xi'_v R'_v}$.

Set Z to be the union of the finitely many $Z_{\xi'}$ over all $\xi' \in \Xi$. Given an arbitrary weighting function ξ and a sequence $\{x_i\}$ in $X(k) \setminus Z(k)$, let $\xi' \in \Xi$ be a weighting function such that $\xi'_v R'_v \leq \xi_v R_v$ for all $v \in S$. Then since $X(k) \setminus Z(k) \subseteq X(k) \setminus Z_{\xi'}(k)$ we conclude that there is some $v \in S$ with $\xi'_v \neq 0$ so that $\alpha_v(\{x_i\}, L) \geq \frac{1}{\xi'_v R'_v} \geq \frac{1}{\xi_v R_v}$.

Finally the statements about equality in the case $\dim X = 1$ follow as in the proof of Corollary 5.4. (After proving the equivalent version of Corollary 5.3.) \square

As in Theorem 6.3 inducting on dimension yields a version with $Z = \emptyset$.

Theorem 7.5. *Set $R_v = \frac{2}{\epsilon_{x_v}(L)}$ for each $v \in S$. Then the conditions in Proposition 7.3 hold with respect to the collection $\{R_v\}_{v \in S}$ and $Z = \emptyset$.*

Proof: We will show condition (7.3.3) holds for the collection $\{R_v\}_{v \in S}$ and with $Z = \emptyset$, i.e., given any weighting function ξ and any $\delta > 0$ we will show that there are only finitely many solutions $y \in X(k)$ to

$$(7.5.1) \quad d_v(x_v, y)^{\frac{\epsilon_{x_v}(L)}{2}} \leq H_L(y)^{-\xi_v(1+\delta)} \quad \text{for all } v \in S.$$

Suppose a weighting function ξ is given. When $\dim X = 1$ the result we want to prove is Theorem 7.4(b). If $\dim X = n > 1$ then $2 > \frac{n+1}{n}$ so by Theorem 7.4(b) again there is a proper subvariety $Z' \subset X$ such that there are only finitely many $y \in X(k) \setminus Z'(k)$ satisfying (7.5.1). Let Z_j be an irreducible component of Z' . By induction there are only finitely many solutions $y \in Z_j(k)$ to the equations

$$d_v(x_v, y)^{\frac{\epsilon_{x_v, Z_j}(L)}{2}} < H_L(y)^{-\xi_v(1+\delta)} \quad \text{for all } v \in S.$$

Since $\epsilon_{x_v, Z_j}(L|_{Z_j}) \geq \epsilon_{x_v, X}(L)$ this is a stronger statement than the one we are claiming, i.e., this implies that there are only finitely many solutions $y \in Z_j(k)$ to (7.5.1). Thus there are only finitely many solutions $y \in X(k)$ to (7.5.1). \square

Corollary 7.6. *For any sequence $\{x_i\}$ in $X(k)$*

$$(7.6.1) \quad \sum_{v \in S} \frac{\epsilon_{x_v}(L)}{\alpha_v(\{x_i\}, L)} \leq 2.$$

Equivalently, for any $\delta > 0$ there are only finitely many solutions $y \in X(k)$ to

$$\prod_{v \in S} d_v(x_v, y)^{\epsilon_{x_v}(L)} < H_L(y)^{-(2+\delta)}.$$

Proof: These are conditions (7.3.2) and (7.3.4) respectively when $Z = \emptyset$ and with the choice of $R_v = \frac{2}{\epsilon_{x_v}(L)}$ for all $v \in S$. These conditions hold by Theorem 7.5. \square

Equality. As in Theorem 6.3 it is useful to study the case of “equality” in Theorem 7.5. By “equality” we mean that there is a sequence $\{x_i\}$ so that (7.6.1) is an equality. Equivalently, in terms of condition (7.3.1), equality means that for the given sequence $\{x_i\}$ there a weighting function ξ such that

$$(7.6.2) \quad \alpha_v(\{x_i\}, L) = \frac{\epsilon_{x_v}(L)}{2\xi_v} \text{ for all } v \in S'.$$

where $S' = \{v \in S \mid \xi_v \neq 0\}$.

Theorem 7.7. *Suppose that $\{x_i\}$ is a sequence so that we have equality in (7.6.1). Let $S' = \{v \in S \mid \alpha_v(\{x_i\}, L) < \infty\}$ (note that S' is nonempty — otherwise equality in (7.6.1) is impossible). Then there is a k -rational curve C containing infinitely many x_i such that for all $v \in S'$: (1) C is unibranch at x_v (in particular, C contains x_v) (2) $\kappa(x_v) \neq k$, (3) $\kappa(x_v) \subset k_v$, and (4) $\epsilon_{x_v, C}(L|_C) = \epsilon_{x_v, X}(L)$.*

Conversely, given a k -rational curve C satisfying these conditions with respect to a non-empty subset $S' \subseteq S$, then for any weighting function $\xi: S' \rightarrow (0, 1]$ (extended by 0 to a weighting function on S) there is a sequence $\{x_i\}$ of points of $C(k)$ such that (7.6.2) holds.

For the converse direction of Theorem 7.7 we require a “simultaneous weighted Dirichlet” result on \mathbb{P}^1 , which seems to be generally known, but for which we could not find a reference. We first prove this result, which is slightly involved, below. The proof of Theorem 7.7 appears after Corollary 7.9.

We are indebted to Damien Roy for the following argument.

Theorem 7.8. *Let S be a finite set of places of k containing all the archimedean places. For each place v of S , let $e_v \in [0, 2)$ be a real number between 0 and 2, satisfying $e = \sum_{v \in S} e_v < 2$. For each v in S , let x_v be an algebraic element not in k of the completion k_v of k at v . Then there exist infinitely many elements $y \in k$ such that $\|y - x_v\|_v < H(y)^{-e_v}$ for all v in S .*

Proof: Let R be the ring of S -integers of k , and embed R in $V = \prod_{v \in S} k_v$ via the diagonal embedding. This embedding also induces an embedding of R^2 in V^2 . Let B be a large, positive real number, and for each $v \in S$ set $f_v = e_v/e$.

There is a convex subset D of V^2 of finite volume (i.e., Haar measure) with the property that D contains a complete set of representatives for the abelian group V^2/R^2 . For any positive real number N let A_N be the set of vectors $(\mathbf{a}, \mathbf{b}) \in V^2$ such that $\|a_v - x_v b_v\|_v < B^{-f_v}$ and $\|b_v\|_v < NB^{f_v}$ for all places v in S . Choose N large enough so that the volume of $\frac{1}{2}A_N$ is greater than the volume of D and set $A = A_N$ (note that the choice of N does not depend on B). We will show that A contains a nonzero element of R^2 by generalizing the proof of Minkowski’s famous result in the geometry of numbers, as found in [16, §1.4].

To see this, consider the sets $\frac{1}{2}A \cap (D + u)$ as u varies over elements of R^2 . They clearly cover the set $\frac{1}{2}A$, and for each u , we have $\frac{1}{2}A \cap (D + u) = ((\frac{1}{2}A - u) \cap D) + u$. Therefore, the volume of $\frac{1}{2}A \cap (D + u)$ is equal to that of $(\frac{1}{2}A - u) \cap D$. If the sets $(\frac{1}{2}A - u) \cap D$ were pairwise disjoint, then by summing over u , we would find that the volume of $\frac{1}{2}A$ is at most

the volume of D , in contradiction to our choice of A . We conclude that the sets $(\frac{1}{2}A - u) \cap D$ are not disjoint.

We may therefore find elements $u, v \in R^2$ and $a_1, a_2 \in A$ such that $\frac{1}{2}a_1 - u = \frac{1}{2}a_2 - v$. Since A is convex and closed under multiplication by -1 , it follows that $u - v$ is a nonzero element of $R^2 \cap A$, as desired. Let $(a_B, b_B) \in R^2$ be such an element (so $a_B, b_B \in R$, and for all $v \in S$ $\|a_B - x_v b_B\|_v < B^{-f_v}$ and $\|b_B\|_v < NB^{f_v}$). Since at least one of the x_v is not in k , at least one $\|a_B - x_v b_B\|_v \neq 0$, and as B goes to infinity we obtain infinitely many such pairs. Now, the height of a_B/b_B is at most $\prod \|b_B\|_v$ (since the v -adic valuation of a_B is essentially determined by those of b_B and x_v), so we deduce that $H([a_B:b_B]) \leq B \sum f_v = B$.

Since $e < 2$ we may choose $\delta > 0$ small enough so that $2e_v/e - \delta > e_v$ for each $v \in S$. Fix one such δ . By the Schmidt Subspace Theorem (see [1, Corollary 7.2.5]) applied to the linear forms $a - bx_v$ and b over the places v of S , it follows that there is a finite set of lines in k^2 which contain all pairs $(a, b) \in k^2$ satisfying $\|a - x_v b\|_v < H([a:b])^{-f_v}$ and $\|b\|_v < H([a:b])^{f_v - \delta}$ for all $v \in S$.

If this finite set of lines contains infinitely many of the $(a, b) \in R^2$ satisfying $\|a - x_v b\|_v < H([a:b])^{-f_v}$ and $\|b\|_v < H([a:b])^{f_v}$ constructed above, then there is an infinite set of such pairs lying on one of the lines. That is, there is a fixed $m \in k$ and an infinite set of pairs $(a, ma) \in R^2$ so that $\|a\|_v \cdot \|1 - x_v m\|_v = \|a - x_v ma\|_v < H([a:ma])^{-f_v} = H([1:m])^{-f_v}$. Since none of the x_v are in k , none of the $1 - x_v m$ are zero, and this implies that $\|a\|_v < C$ for all $v \in S$ and some constant C . Since $\|a\|_v \leq 1$ for all $v \notin S$ this implies that $H([a:1])$ is bounded, contradicting the fact that there are infinitely many different a .

Thus, there are an infinite number of pairs $(a, b) \in R^2$ which satisfy $\|a - x_v b\|_v < H([a:b])^{-f_v}$ and $\|b\|_v \geq H([a:b])^{f_v - \delta}$, and hence infinitely many $a/b \in k$ satisfying $\|a/b - x_v\|_v \leq H([a:b])^{-2f_v + \delta} = H([a:b])^{-2e_v/e + \delta} \leq H([a:b])^{-e_v}$. \square

Given a sequence $\{y_i\} \subseteq \mathbb{P}^1(k)$ set $\tau_v(\{y_i\}) = 1/\alpha_v(\{y_i\}, \mathcal{O}_{\mathbb{P}^1}(1))$. Since τ is the reciprocal of α , if $\tau' < \tau_v(\{y_i\})$ (respectively $\tau' > \tau_v(\{y_i\})$) then $\lim_{i \rightarrow \infty} d_v(x_v, y_i)^{1/\tau'} H(y_i) = 0$ (respectively $= \infty$). The content of Theorem 7.8 is that given any finite set S of places of k , and any collection $\{e_v\}_{v \in S}$ of elements of $[0, 2]$ with $\sum e_v < 2$, there is a sequence $\{y_i\}$ such that $e_v \leq \tau_v(\{y_i\})$ for all $v \in S$. By a simple diagonal argument we now see that if we choose the e_v so that $\sum e_v = 2$, we may achieve equality.

Corollary 7.9. (Simultaneous weighted Dirichlet): *Let S be a finite set of places of k , and $\{e_v\}_{v \in S}$ a collection of elements of $(0, 2]$ such that $\sum e_v = 2$. Then there is a sequence $\{y_i\}$ of k -points of \mathbb{P}^1 such that $e_v = \tau_v(\{y_i\})$ for all $v \in S$.*

Proof: Let n_0 be large enough so that $e_v - \frac{1}{n} > 0$ for all $n \geq n_0$ and all $v \in S$. By Theorem 7.8 for each $n \geq n_0$ there is a sequence $\{y_{i,n}\}_{i \geq 0}$ such that $e_v - \frac{1}{2n} \leq \tau_v(\{y_{i,n}\})$. Since $e_v - \frac{1}{n} < e_v - \frac{1}{2n}$, we have $\lim_{i \rightarrow \infty} d_v(x_v, y_{i,n})^{\frac{1}{e_v - 1/n}} H(y_{i,n}) = 0$ for all $v \in S$. For each fixed n , by choosing i large enough, we may pick $y_n = y_{i,n}$ so that $d_v(x_v, y_n)^{\frac{1}{e_v - 1/n}} H(y_n) < \frac{1}{n}$ and $d_v(x_v, y_n) < 1$ for all $v \in S$. In this way we construct a sequence $\{y_n\}_{n \geq n_0}$ which we simply call $\{y_n\}$.

Fix $\delta > 0$ small enough that $e_v - \delta > 0$ for each $v \in S$. For large n we have $e_v - \delta < e_v - \frac{1}{n}$ and hence

$$d_v(x_v, y_n)^{\frac{1}{e_v - \delta}} H(y_n) < d_v(x_v, y_n)^{\frac{1}{e_v - 1/n}} H(y_n) < \frac{1}{n}.$$

Therefore $\lim_{n \rightarrow \infty} d_v(x_v, y_n)^{\frac{1}{e_v - \delta}} H(y_n) = 0$ and so $e_v - \delta \leq \tau_v(\{y_n\})$. Letting δ go to zero we conclude that $e_v \leq \tau_v(\{y_n\})$ for each $v \in S$. By Roth's theorem for \mathbb{P}^1 (e.g., Corollary 7.6) $\sum_v \tau_v(\{y_n\}) \leq 2$. Since $\sum_v e_v = 2$ we conclude that $e_v = \tau_v(\{y_n\})$ for each $v \in S$. \square

Proof of Theorem 7.7: In the induction proving Theorem 7.5, in order to arrive at equality we must have gone all the way down to curve C , necessarily k -rational (since there are infinitely many rational points, and the approximation constants are finite). The first result then follows by Roth's theorem for \mathbb{P}^1 (with the appropriate modification for the singularity, as in Theorem 2.17 for a single point). The converse direction is Corollary 7.9 with the choice $e_v = 2\xi_v$ for all $v \in S$, combined with the appropriate modification for the singularity, again as in Theorem 2.17. \square

As in the case of a single place there are other variations on the deduction of Theorem 7.5 from Theorem 7.4.

Corollary 7.10. *For any positive integer $m < n$, if we choose R_v so that $R_v > \frac{m}{(m+1)\epsilon_{x_v}(L)}$ for each $v \in S$ then there is a subset Z of X with $\dim Z < m$ such that the equivalent conditions in Proposition 7.3 hold with respect to Z and $\{R_v\}_{v \in S}$.*

Corollary 7.11. *Suppose that there is no k -rational curve passing through any of the x_v , $v \in S$. Then the conditions of Proposition 7.3 hold with $Z = \emptyset$ and $R_v = \frac{3}{2\epsilon_{x_v}(L)}$ for all $v \in S$.*

Proof of Corollary 7.11: We prove that condition (7.3.3) holds with respect to this data, i.e., that given any weighting function ξ and any collection $\{\delta_v\}_{v \in S}$ of positive real numbers, there are only finitely many solutions $y \in X(k)$ to

$$(7.11.1) \quad d_v(x_v, y)^{\frac{2\epsilon_{x_v}(L)}{3}} < H_L(y)^{-\xi_v(1+\delta_v)} \quad \text{for all } v \in S.$$

Given the collection $\{\delta_v\}_{v \in S}$ set $\delta'_v = \frac{\delta_v}{2}$ and $R'_v = \frac{3}{2\epsilon_{x_v}(L)} + \delta'_v$ for all $v \in S$. By Corollary 7.10 with $m = 2$ there is a curve Z' , depending on $\{R'_v\}$, so that there are only finitely many solutions $y \in X(k) \setminus Z'(k)$ to (7.11.1). By hypothesis, there is no k -rational curve passing through any of the x_v , and so we conclude that there are only finitely many solutions $y \in Z'(k)$ to (7.11.1). Thus there are only finitely many solutions $y \in X(k)$ to (7.11.1). \square

As in Theorem 7.5 it is probably simplest to express Corollary 7.11 in terms of condition (7.3.2), i.e., as an inequality governing the position of the point $(\alpha_{v_1}(\{x_i\}, L), \dots, \alpha_{v_s}(\{x_i\}, L))$ in \mathbb{R}^s . Assuming the hypotheses of the Corollary, for any sequence $\{x_i\}$ of k points,

$$\sum_{v \in S} \frac{\epsilon_{x_v}(L)}{\alpha_v(\{x_i\}, L)} \leq \frac{3}{2}.$$

Remark. It is clear that it is possible to continue this type of argument if in each dimension m we knew the types of m -dimensional subvariety Z where ‘‘equality’’ occurs, i.e., where there is a sequence $\{x_i\}$ of points of $Z(k)$, with no subsequence contained in a proper subvariety of Z , satisfying

$$\sum_{v \in S} \frac{\epsilon_{x_v, Z}(L)}{\alpha_v(\{x_i\}, L)} = \frac{m+1}{m}.$$

One necessary condition on such a Z is that Z must be Seshadri exceptional (see §9) with respect to each point x_v where $\alpha_v(\{x_i\}, L) < \infty$. (Here Seshadri exceptional means as a subvariety of itself, not as a subvariety of X .) It would already be interesting to work out

the case of surfaces. For instance \mathbb{P}^2 is such a surface if none of the points x_v lie on k -rational lines.

Remark. In this section we have used a different constant R_v at each place when describing results on simultaneous approximation. By replacing each R_v with the largest (i.e., the worst) of the R_v we obtain a weaker statement, but with the advantage of the same constant at each place. Thus, for example, Theorem 7.5 implies the following product version.

Corollary 7.12. *Let $\epsilon = \min_{v \in S}(\epsilon_{x_v}(L))$. Then for any $\delta > 0$ there are only finitely many solutions $y \in X(k)$ to*

$$\prod_{v \in S} d_v(x_v, y) \leq H_L(y)^{-\left(\frac{2}{\epsilon} + \delta\right)}.$$

8. IMPROVEMENTS VIA UNRAMIFIED COVERS

Theorem 5.2 and an idea due to Robinson-Roquette [17] and Macintyre [13] (see also [19, p. 100 and §7.7]) allow us to give sharper versions of the theorems so far.

In this section by *unramified cover* we mean a finite surjective unramified map $\varphi: Y_1 \rightarrow Y_2$ in the category of varieties over $\text{Spec}(k)$, with both Y_1 and Y_2 irreducible.

Let $\varphi: Y \rightarrow X$ be an unramified cover and x be any point of $X(\bar{k})$. As we will see below, for any ample bundle L on X , $\min_{y \in \varphi^{-1}(x)}(\beta_y(\varphi^*L))$ and $\min_{y \in \varphi^{-1}(x)}(\epsilon_y(\varphi^*L))$ are at least as large as $\beta_x(L)$ and $\epsilon_x(L)$ respectively. We will define $\hat{\beta}_x$ and $\hat{\epsilon}_x$ as suprema over such unramified covers. The point of this section is that the theorems in §6 and §7 hold with β and ϵ replaced by $\hat{\beta}$ and $\hat{\epsilon}$. The basic idea is to lift a sequence $\{x_i\}$ on X to a sequence $\{y_i\}$ on Y and use the bounds there; however the lift involves a change of field, and this introduces a factor which seems to make the result strictly worse. Fortunately, by using simultaneous approximation on Y we can exactly cancel out this factor. In particular, even to get such a result for a single place of k we must use simultaneous approximation on the cover Y .

We first check that β and ϵ are weakly increasing in unramified covers; thus the theorems using $\hat{\beta}$ and $\hat{\epsilon}$ are stronger than the original ones.

Lemma 8.1. *Let $\varphi: Y \rightarrow X$ be an unramified cover, L an ample line bundle on X , x any point of $X(\bar{k})$, and $y \in \varphi^{-1}(x)$. Then*

- (a) $\beta_y(\varphi^*L) \geq \beta_x(L)$, and
- (b) $\epsilon_y(\varphi^*L) \geq \epsilon_x(L)$.

Proof: Let $\bar{X}_1, \dots, \bar{X}_r$ and $\bar{Y}_1, \dots, \bar{Y}_s$ be the irreducible components of $X \times_k \bar{k}$ and $Y \times_k \bar{k}$ containing x and y respectively. Each \bar{Y}_i maps to some \bar{X}_j , and this map expresses \bar{Y}_i as an unramified cover of \bar{X}_j . Since β and ϵ are defined as minima over irreducible components, establishing the conclusion of the lemma for each map $\bar{Y}_i \rightarrow \bar{X}_j$ establishes the lemma for $Y \rightarrow X$. Thus we are reduced to the case of studying unramified covers over an algebraically closed field. To reduce notation we continue to use X and Y as the names of the varieties, rather than \bar{X}_j and \bar{Y}_i , and φ as the name of the map.

Let $\pi_X: \tilde{X} \rightarrow X$ be the blow up of X at x with exceptional divisor E_x , and for any $\gamma \geq 0$ set $L_{X,\gamma} = \pi_X^*L - \gamma E_x$ and $f_X(\gamma) = \frac{\text{Vol}(L_{X,\gamma})}{\text{Vol}(L)}$. We similarly let $\pi_Y: \tilde{Y} \rightarrow Y$ be the blow up

of Y at y with exceptional divisor E_y , and for any $\gamma \geq 0$ we set $L_{Y,\gamma} = \pi_Y^* \varphi^* L - \gamma E_y$ and $f_Y(\gamma) = \frac{\text{Vol}(L_{Y,\gamma})}{\text{Vol}(\varphi^* L)}$.

We first prove (b). Let $\varphi^{-1}(x) = \{y_1, \dots, y_\ell\}$ with $y_1 = y$. Since φ is unramified, the fibre product $Y \times_X \tilde{X}$ is the blow up of Y at the points y_1, \dots, y_ℓ . Let ψ_X and ψ_Y be the maps from $Y \times_X \tilde{X}$ to \tilde{X} and \tilde{Y} respectively (the map to \tilde{Y} being the blow down at the points of $\varphi^{-1}(x)$ different from y). Thus we have the following commutative diagram of maps

$$\begin{array}{ccc} Y \times_X \times \tilde{X} & \xrightarrow{\psi_X} & \tilde{X} \\ \downarrow \psi_Y & & \downarrow \pi_X \\ \tilde{Y} & & X \\ \downarrow \pi_Y & \xrightarrow{\varphi} & \\ Y & & \end{array}$$

For $i = 2, \dots, \ell$ let E_i be the exceptional divisor of ψ_Y lying over y_i . The description of the fibre product as a further blowup of \tilde{Y} shows that for any γ we have the equality of line bundles

$$(8.1.1) \quad \psi_X^* L_{X,\gamma} = \psi_Y^* L_{Y,\gamma} - \gamma \left(\sum_{i=2}^{\ell} E_i \right).$$

If $0 \leq \gamma \leq \epsilon_x(L)$ then $L_{X,\gamma}$ is nef on \tilde{X} and so $\psi_X^* L_{X,\gamma}$ is nef on $Y \times_X \tilde{X}$. Equation (8.1.1) then implies that $L_{Y,\gamma}$ is nef on \tilde{Y} . This proves (b).

We will prove (a) by showing the inequality $f_Y(\gamma) \geq f_X(\gamma)$ for all $\gamma \geq 0$. Since both f_X and f_Y are continuous functions, it suffices to prove the inequality for rational γ .

Set $\mathcal{E} = \varphi_* \mathcal{O}_Y$ and let d be the generic rank of \mathcal{E} . By the projection formula, for any $m > 0$ we have $\varphi_*(\varphi^* mL) = mL \otimes_{\mathcal{O}_x} \mathcal{E}$, and so $H^0(Y, \varphi^* mL) = H^0(X, mL \otimes \mathcal{E})$. The volume measures the leading term in the asymptotic growth of global sections, and for this purpose tensoring with the (generic) rank d sheaf \mathcal{E} has the same effect as tensoring with d copies of \mathcal{O}_X . Therefore $\text{Vol}(\varphi^* L) = d \text{Vol}(L)$. Similarly, for any $\gamma \geq 0$ we have $\text{Vol}(\psi_X^* L_{X,\gamma}) = d \text{Vol}(L_{X,\gamma})$.

For any rational $\gamma \geq 0$, and any $m \geq 0$ sufficiently divisible so that $m\gamma$ is integral, multiplying (8.1.1) by m shows that global sections of $m\psi_X^* L_{X,\gamma}$ are a subspace of the global sections of $m\psi_Y^* L_{Y,\gamma}$ and so $\text{Vol}(\psi_Y^* L_{Y,\gamma}) \geq \text{Vol}(\psi_X^* L_{X,\gamma})$. Finally, since ψ_Y is birational, $\text{Vol}(L_{Y,\gamma}) = \text{Vol}(\psi_Y^* L_{Y,\gamma})$. We thus have

$$f_Y(\gamma) = \frac{\text{Vol}(\psi_Y^* L_{Y,\gamma})}{\text{Vol}(\varphi^* L)} \geq \frac{\text{Vol}(\psi_X^* L_{X,\gamma})}{\text{Vol}(\varphi^* L)} = \frac{d \text{Vol}(L_{X,\gamma})}{d \text{Vol}(L)} = f_X(\gamma),$$

and integrating gives $\beta_y(\varphi^* L) \geq \beta_x(L)$. \square

In the category of schemes over X , consider the full subcategory whose objects are the unramified covers $\varphi: Y \rightarrow X$ as above. If (Y_1, φ_1) and (Y_2, φ_2) are objects and $\psi: Y_1 \rightarrow Y_2$ a morphism in this category, then ψ expresses Y_1 as an unramified cover of Y_2 , and thus Lemma 8.1 applies. In particular, for any $y_2 \in Y_2(\bar{k})$, $\min_{y_1 \in \psi^{-1}(y_2)} (\beta_{y_1}(\varphi_1^* L)) \geq \beta_{y_2}(\varphi_2^* L)$ and similarly for ϵ .

Definition 8.2. Let X be an irreducible variety over $\text{Spec}(k)$, L an ample line bundle on X and $x \in X(\bar{k})$. We define

$$\hat{\beta}_x(L) = \sup_{\varphi: Y \rightarrow X} \min_{y \in \varphi^{-1}(x)} \beta_y(\varphi^* L) \quad \text{and} \quad \hat{\epsilon}_x(L) = \sup_{\varphi: Y \rightarrow X} \min_{y \in \varphi^{-1}(x)} (\epsilon_y(\varphi^* L)),$$

where the suprema are over the set of unramified covers $\varphi: Y \rightarrow X$.

In the arguments below it will be important to know we can find a single unramified cover which approximates finitely many of the $\hat{\beta}_x(L)$.

Lemma 8.3. Let X be an irreducible variety over $\text{Spec}(k)$, L an ample line bundle on X , and x_1, \dots, x_ℓ finitely many points of $X(\bar{k})$. Suppose that $\beta_1, \dots, \beta_\ell$ are positive real numbers with $\beta_i < \hat{\beta}_{x_i}(L)$ for $i = 1, \dots, \ell$. Then there exists an unramified cover $\varphi: Y \rightarrow X$ such that $\min_{y \in \varphi^{-1}(x_i)} (\beta_y(\varphi^*(L))) > \beta_i$ for $i = 1, \dots, \ell$.

Proof: By the definition of $\hat{\beta}$, for each i there is an unramified cover $\varphi_i: Y_i \rightarrow X$ such that $\min_{y \in \varphi_i^{-1}(x_i)} (\beta_y(\varphi_i^*(L))) > \beta_i$. Let Y be any irreducible component of the fibre product $Y_1 \times_X \dots \times_X Y_\ell$, and $\varphi: Y \rightarrow X$ the induced map. The natural projection maps of the fibre product induce maps $\psi_i: Y \rightarrow Y_i$ for each i , and ψ_i expresses Y as an unramified cover of Y_i . For any $y \in \varphi^{-1}(x_i)$, $\psi_i(y) \in \varphi_i^{-1}(x_i)$, and hence an application of Lemma 8.1 to the unramified cover ψ_i shows that Y has the desired property. \square

Remarks on Galois symmetries. (a) Suppose that $x \in X(\bar{k})$. Given any algebraic conjugate x' of x let $\sigma \in \text{Gal}(\bar{k}/k)$ be an element such that $\sigma(x) = x'$. Given any unramified cover $\varphi: Y \rightarrow X$, the action of σ on $Y(\bar{k})$ then takes points of $Y(\bar{k})$ lying over x to points lying over x' . We conclude that for any ample line bundle L on X , $\min_{y \in \varphi^{-1}(x)} \beta_y(\varphi^* L) = \min_{y' \in \varphi^{-1}(x')} \beta_{y'}(\varphi^* L)$. This Galois symmetry argument also shows that $\hat{\epsilon}_x(L) = \hat{\epsilon}_{x'}(L)$ and $\hat{\beta}_x(L) = \hat{\beta}_{x'}(L)$ for any ample line bundle L on X .

(b) Let v_0 be a place of k , v and v' two extensions of v_0 to \bar{k} , and $\sigma \in \text{Gal}(\bar{k}/k)$ such that $v' = v \circ \sigma$. Fix a point $x \in X(\bar{k})$ and set $x' = \sigma(x)$. Suppose that a sequence $\{x_i\} \subseteq X(k)$ converges to x with respect to a distance function d_v . If we define $d_{v'}$ by using the same embedding $X \hookrightarrow \mathbb{P}_k^r$ used to define d_v , then applying σ to formula (2.1.1) or (2.1.2) shows that $d_{v'}(x', x_i) = d_v(x, x_i)$ for all $i \geq 0$. More generally, if $d_{v'}$ is defined by using a different embedding of X then this result combined with Proposition 2.5 shows that $d_{v'}(x', x_i)$ and $d_v(x, x_i)$ are equivalent as $i \rightarrow \infty$. Summarizing, if $\{x_i\}$ converges to $x \in X(\bar{k})$ with respect to d_v , and if v' is a different extension of v_0 to \bar{k} , then $\{x_i\}$ will converge to an algebraic conjugate of x with respect to $d_{v'}$, with the same essential speed of convergence.

Lifting sequences. Let $\psi: Y' \rightarrow X$ be an unramified cover. By the theorem of Chevalley-Weil [1, Theorem 10.3.11] there is a finite extension F/k such that all points $\{y \in Y'(\bar{k}) \mid \psi(y) \in X(k)\}$ are defined over F (this field F is not unique, since any larger field will also work). Fix such a field F . It will be convenient for us that the covering variety is also a variety over $\text{Spec}(F)$. To do this we let Y be an irreducible component of $Y' \times_k F$, and $\varphi: Y \rightarrow X$ the induced map. Via the natural map $\text{Spec}(F) \rightarrow \text{Spec}(k)$, Y is a variety over $\text{Spec}(k)$, and φ is an unramified cover. Furthermore, all points of $Y(\bar{k})$ lying over points of $X(k)$ are again defined over F .

Given a sequence $\{x_i\}$ of points of $X(k)$, for each i we arbitrarily choose $y_i \in Y(F)$ with $\varphi(y_i) = x_i$. We call such a sequence $\{y_i\}$ a *lift* of $\{x_i\}$. This lift is somewhat haphazard, but by further passing to a subsequence we may obtain a lift with better properties.

Let v_0 be a place of k , extended to a place v on \bar{k} , and suppose that there is $x_v \in X(\bar{k})$ such that $d_v(x_v, x_i) \rightarrow 0$, i.e., that $\{x_i\}$ approximates x_v with respect to $d_v(\cdot, \cdot)$. Let w_0 be a place of F lying over v_0 , and w an extension of w_0 to \bar{F} . In this situation we define x_w , an algebraic conjugate of x , as follows. On $\bar{k} = \bar{F}$, w gives a place v' of \bar{k} extending v_0 , but which may not be equal to v . We then apply (b) of the ‘Remarks on Galois symmetries’ above to obtain an algebraic conjugate x_w of x . With respect to $d_{v'}$, $\{x_i\}$ converges to x_w .

Returning to the problem of improving the lift, since $Y(F_w)$ is compact by passing to a subsequence we may assume that the sequence $\{y_i\}$ has a limit $y_w \in Y(F_w)$. Since Y is a variety over $\text{Spec}(F)$, the place w gives a distance function $d_w(\cdot, \cdot)_F$ on Y . The topology on $Y(F_w)$ is that induced by $d_w(\cdot, \cdot)_F$, and so this means that $d_w(y_w, y_i)_F \rightarrow 0$ as $i \rightarrow \infty$. Furthermore, by continuity we have $\varphi(y_w) = x_w$, in particular, $y_w \in Y(\bar{k})$.

We will need a generalization obtained by repeating this procedure. Let T_v be the set of places of F over v_0 , each extended to a place of \bar{F} . As above, for each such $w \in T_v$ we obtain an algebraic conjugate x_w of x_v . (These conjugates are not necessarily distinct.) By applying the procedure above to each $w \in T_v$ in turn, we may find \bar{k} -points $y_w \in \varphi^{-1}(x_w)$ for each $w \in T_v$, and a subsequence of $\{y_i\}$ so that for each $w \in T_v$, $d_w(y_w, y_i)_F \rightarrow 0$ as $i \rightarrow \infty$.

Finally, given a finite set S of places of k extended to \bar{k} , we may repeat this process for each $v \in S$. We record the conclusion below.

Proposition 8.4. *Let $\psi: Y' \rightarrow X$ be an unramified cover, S a finite set of places of k each extended to \bar{k} , and F/k a finite extension so that for all $x \in X(k)$, all points of $\psi^{-1}(x)$ are defined over F . We replace Y' by a component Y of $Y' \times_k F$, and let $\varphi: Y \rightarrow X$ be the induced map. Suppose that $\{x_v\}_{v \in S}$ are a set of points of $X(\bar{k})$, and that $\{x_i\}$ is a sequence of k -points so that $d_v(x_v, x_i) \rightarrow 0$ for each $v \in S$. For each $v \in S$ let T_v be the set of places of F lying over $v_0 = v|_k$ each extended to a place of \bar{F} . For each such w we let x_w be the corresponding algebraic conjugate of x_v as defined above.*

Then by passing to a subsequence of $\{x_i\}$ we may find a lift $\{y_i\}$ of $\{x_i\}$ to Y , and for each $v \in S$ and $w \in T_v$ a \bar{k} -point $y_w \in \varphi^{-1}(x_w)$, such that $d_w(y_w, y_i)_F \rightarrow 0$ as $i \rightarrow \infty$.

We next compare the asymptotics of $d_v(x_v, x_i)_k$ with $d_w(y_w, y_i)_F$, and the resulting effect on α .

Lemma 8.5. *Let $\psi: Y' \rightarrow X$ be an unramified cover, v_0 a place of k , v an extension of v_0 to \bar{k} , and $\{x_i\}$ a sequence of points of $X(k)$ converging to $x \in X(\bar{k})$ with respect to $d_v(\cdot, \cdot)$. Let F/k be a finite extension so that for all $x \in X(k)$, all points of $\psi^{-1}(x)$ are defined over F . Let Y be a component of $Y' \times_k F$ and $\varphi: Y \rightarrow X$ the induced map. Let $\{y_i\}$ be a lift of $\{x_i\}$ to $Y(\bar{k})$, w_0 a place of F lying over v_0 , and w an extension of w_0 to \bar{F} . Suppose that $\{y_i\}$ converges to $y \in Y(\bar{k})$ with respect to $d_w(\cdot, \cdot)_F$. Finally, set $m_w = [F_w: k_v]$ ($= [F_{w_0}: k_{v_0}]$) and $e = [F: k]$.*

Then $d_w(y, y_i)_F$ is asymptotically equivalent to $d_v(x, x_i)_k^{m_w}$ as $i \rightarrow \infty$, and for any line bundle L on X , $\alpha_y(\{y_i\}, \varphi^ L)_F = \frac{e}{m_w} \alpha_x(\{x_i\}, L)_k$.*

Here, as in Proposition 2.13, the subscript k or F on $d_v(\cdot, \cdot)_k$ or $d_w(\cdot, \cdot)_F$ respectively indicates the field being used to normalize the distance function. The subscript on $\alpha(\cdot, \cdot)$ similarly indicates the field used to normalize the distance function and the height.

Proof: For each y_i in the sequence, we have $H_{\varphi^*L}(y_i)_F = H_L(x_i)_F = H_L(x_i)_k^e$. Thus once we show that $d_w(y, y_i)_F$ is asymptotically equivalent to $d_v(x, x_i)_k^{m_w}$ as $i \rightarrow \infty$ the equality $\alpha_y(\{y_i\}, \varphi^*L)_F = \frac{e}{m_w} \alpha_x(\{x_i\}, L)_k$ follows immediately as in the proof of Proposition 2.13.

The geometric point $x \in X(\bar{k})$ defines a morphism $\text{Spec}(\bar{k}) \rightarrow X$, whose image is a closed point $[x] \in X$. Let U' be an affine neighbourhood of $[x]$, and let u_1, \dots, u_r be elements of $\Gamma(U', \mathcal{O}_X)$ which generate the maximal ideal of $[x]$. Let K be the Galois closure (over k) of the field of definition of y , and set $x' = \varphi(y)$; as before x' is an algebraic conjugate of x . Since y maps to x' , and since x' and x have isomorphic fields of definition, K also contains the field of definition of x . Pulling back to $U'_K := U' \times_k K$, the functions u_1, \dots, u_r cut out all the geometric points lying over $[x]$. By passing to a smaller affine open $U \subset U'_K$ we may assume that u_1, \dots, u_r generate the maximal ideal of x . By Lemma 2.7, $d_v(x, x_i)_k$ is equivalent to $\max(\|u_1(x_i)\|_v, \dots, \|u_r(x_i)\|_v)$ as $i \rightarrow \infty$, and so $d_w(x, x_i)_k^{m_w}$ is equivalent to $\max(\|u_1(x_i)\|_v^{m_w}, \dots, \|u_r(x_i)\|_v^{m_w})$ as $i \rightarrow \infty$.

Let $V'_K = \varphi_K^{-1}(U'_K)$. Since φ is unramified, $\varphi^*u_1, \dots, \varphi^*u_r$ cut out the reduced points lying over all algebraic conjugates of x . Therefore if we restrict to a small enough affine open neighbourhood V of y in V'_K then $\varphi^*u_1, \dots, \varphi^*u_r$ generate the maximal ideal of y . Since the $\{y_i\}$ converge to y with respect to $d_w(\cdot, \cdot)_F$, we may apply Lemma 2.7 again to conclude that $d_w(y, y_i)_F$ is equivalent to $\max(\|(\varphi^*u_1)(y_i)\|_w, \dots, \|(\varphi^*u_r)(y_i)\|_w)$ as $i \rightarrow \infty$. Since $(\varphi^*u_j)(y_i) = u_j(\varphi(y_i)) = u_j(x_i)$ for each $j = 1, \dots, r$ and each i , we will be done if we show that $\|u_j(x_i)\|_w = \|u_j(x_i)\|_v^{m_w}$ for all i, j .

The value $u_j(x_i)$ is the value of the residue class of u_j in $\kappa(x_i)$. Since the x_i are defined over k , $\kappa(x_i) = k$ and so $u_j(x_i) \in k$ for all i, j . Since w_0 lies over v_0 (and w is normalized with respect to F), the restriction of $\|\cdot\|_w$ to $k \subset F$ is $\|\cdot\|_v^{m_w}$. This proves the lemma. \square

Applying Lemma 8.5 to the lift produced in Proposition 8.4 yields the following corollary.

Corollary 8.6. *Assume the setup and notation of Proposition 8.4, and let $\{y_i\}$ and $\{y_w\}_{w \in T_v, v \in S}$ be the lift and set of points provided by its conclusion. Then for any \mathbb{Q} -bundle L on X , every $v \in S$, and $w \in T_v$ we have $\alpha_{y_w}(\{y_i\}, \varphi^*L)_F = \frac{[F:k]}{[F_w:k_v]} \alpha_{x_v}(\{x_i\}, L)_k$.*

Here α_{y_w} is computed with respect to $d_w(\cdot, \cdot)_F$ and α_{x_v} with respect to $d_v(\cdot, \cdot)_k$.

We are now ready to establish the version of Theorem 5.2 with $\hat{\beta}$ in place of β .

Theorem 8.7. *Let X be an irreducible variety defined over k , S a finite set of places of k , each extended to \bar{k} , and for each $v \in S$ choose an $x_v \in X(\bar{k})$. Suppose that L is an ample \mathbb{Q} -bundle on X defined over k , and that $\{R_v\}_{v \in S}$ are a collection of positive real numbers such that*

$$(8.7.1) \quad \sum_{v \in S} \hat{\beta}_{x_v}(L) R_v > 1.$$

Then (5.1.1) and (5.1.2) hold with respect to the collection $\{R_v\}_{v \in S}$.

Proof: By condition (8.7.1) and Lemma 8.3 we may find an unramified cover $\psi: Y' \rightarrow X$ satisfying

$$(8.7.2) \quad \sum_{v \in S} \left(\min_{y_v \in \psi^{-1}(x_v)} (\beta_{y_v}(\psi^* L)) \right) R_v > 1.$$

Let F/k be a finite extension so that any sequence $\{x_i\}$ of k -points of X can be lifted to a sequence $\{y_i\}$ of points of Y' defined over F , and fix F for the rest of the proof. Let Y be an irreducible component of $Y' \times_k F$, and $\varphi: Y \rightarrow X$ the induced map. Since Y is a component of $Y' \times_k F$, by definition of β (Definition 4.3) we have $\beta_{y,Y}(\varphi^* L) \geq \beta_{y,Y'}\psi^* L$ for every point $y \in Y(\overline{k})$, and so by (8.7.2)

$$(8.7.3) \quad \sum_{v \in S} \left(\min_{y_v \in \varphi^{-1}(x_v)} (\beta_{y_v}(\varphi^* L)) \right) R_v > 1.$$

We will prove the theorem in the form of condition (5.1.1), that is, we will show that there is a proper subvariety $Z \subset X$, such that for all sequences $\{x_i\}$ of k -points of $X \setminus Z$ there is at least one $v \in S$ so that $\alpha_v(\{x_i\}, L) \geq \frac{1}{R_v}$. Here as in §5 for each $v \in S$ we use α_v to mean α_{x_v} , computed with respect to the distance $d_v(\cdot, \cdot)$.

For each $v \in S$ we set T_v to be the set of places of F lying over $v_0 := v|_k$, each extended to a place of \overline{F} . Each $v \in S$ and $w \in T_v$ determine an algebraic conjugate x_w of x_v as above.

Let $\{x_i\}$ be a sequence of k -points of X . If there is a $v \in S$ so that $d_v(x_v, x_i)$ does not go to 0 as $i \rightarrow \infty$, then $\alpha_v(\{x_i\}, L) = \infty$, and the statement to be proved is trivially satisfied. We may therefore restrict ourselves to studying sequences $\{x_i\}$ so that $d_v(x_v, x_i) \rightarrow 0$ as $i \rightarrow \infty$ for each $v \in S$, and we do so for the rest of the proof. We note again that passing to a subsequence can only possibly lower the value of α , so we may freely do so in proving the result.

Set $T = \bigsqcup_{v \in S} T_v$. By Proposition 8.4 if $\{x_i\}$ converges to each x_v with respect to $d_v(\cdot, \cdot)$, then for any lift $\{y_i\}$ of $\{x_i\}$ we may, after passing to a subsequence of $\{y_i\}$, choose a \overline{k} -point $y_w \in \varphi^{-1}(x_w)$ for each $w \in T$ so that $\{y_i\}$ converges to y_w with respect to $d_w(\cdot, \cdot)_F$.

Thus, up to passing to a subsequence, for any sequence $\{x_i\}$ of k -points of X there is a choice of $y_w \in \varphi^{-1}(x_w)$ for each $w \in T$ such that

$$(8.7.4) \quad \left\{ \begin{array}{l} \text{the sequence } \{x_i\} \text{ lifts to a sequence } \{y_i\} \text{ in } Y \text{ (necessarily defined over } F) \\ \text{which converges with respect to } d_w(\cdot, \cdot)_F \text{ to } y_w \text{ for each } w \in T. \end{array} \right\}$$

We will show that for each of the finitely many elements q of the product set $\prod_{w \in T} \varphi^{-1}(x_w)$, i.e., each of the finitely many choices of a $y_w \in \varphi^{-1}(x_w)$ for each $w \in T$, there exists a proper subvariety $Z_q \subset X$ (depending on these choices) so that for any sequence $\{x_i\}$ of k -points of $X \setminus Z_q$ satisfying (8.7.4) there is at least one $v \in S$ so that $\alpha_v(\{x_i\}, L) \geq \frac{1}{R_v}$. Taking Z to be the union over the finitely many such Z_q then yields the theorem.

We now assume that we have fixed $q = \{y_w\}_{w \in T} \in \prod_{w \in T} \varphi^{-1}(x_w)$ and prove the existence of such a Z_q . For each $w \in T$ set $g_w = \frac{[F:k]}{[F_w:k_v]}$, where $v \in S$ is such that $w \in T_v$. If $\{x_i\}$ is a sequence of k -points satisfying (8.7.4) above, and $\{y_i\}$ such a lift, Corollary 8.6 gives $\alpha_w(\{y_i\}, \varphi^* L)_F = g_w \alpha_v(\{x_i\}, L)_k$ for each $v \in S$ and $w \in T_v$. Here, as before, we use α_w to mean α_{y_w} computed with respect to $d_w(\cdot, \cdot)_F$, and the subscripts F and k to indicate the field used to normalize the distance and the heights.

For each $w \in T$ set $R'_w = R_v/g_w$, where again $v \in S$ is such that $w \in T_v$. Combining: (1) for each $v \in S$ we have $\sum_{w \in T_v} \frac{1}{g_w} = 1$; (2) for each $v \in S$, $w \in T_v$, $\min_{y \in \varphi^{-1}(x_v)} \beta_y(\varphi^* L) =$

$\min_{y' \in \varphi^{-1}(x_w)} \beta_{y'}(\varphi^*L)$ (see ‘Remarks on Galois symmetries’ (a)); and (3) inequality (8.7.3), we conclude that

$$\sum_{w \in T} \beta_{y_w}(\varphi^*L) R'_w > 1.$$

Working over $\text{Spec}(F)$, we now apply Theorem 5.2 to the collection $\{R'_w\}_{w \in T}$ and line bundle $\varphi^*(L)$, and let Z'_q be the resulting proper subvariety of Y . Set Z_q to be the image of Z'_q in X . Since Y is irreducible, Z'_q is of dimension strictly less than Y , and hence Z_q is again a proper subvariety of X .

Suppose that $\{x_i\}$ is a sequence of k -points of $X \setminus Z_q$ satisfying (8.7.4), and let $\{y_i\}$ be such a lift. Then $\{y_i\}$ is contained in the F -points of $Y \setminus Z'_q$ and thus by construction of Z'_q there is at least one $v \in S$ and $w \in T_v$ so that

$$\alpha_w(\{y_i\}, \varphi^*L)_F \geq \frac{1}{R'_w} = \frac{g_w}{R_v}.$$

Since $\alpha_w(\{y_i\}, \varphi^*L)_F = g_w \alpha_v(\{x_i\}, L)_k$ we conclude that $\alpha_v(\{x_i\}, L)_k \geq \frac{1}{R_v}$. Thus Z_q has the required property, and this completes the proof of Theorem 8.7. \square

Remarks. (a) Besides the fact that β is weakly increasing in unramified covers, the keys to the proof are (1) the fact that $\alpha_v(\{x_i\}, L)_k$ and $\alpha_w(\{y_i\}, \varphi^*L)_F$ differ by a factor of $1/g_w = \frac{[F_w:k_v]}{[F:k]}$, whenever $w_0 := w|_F$ lies over $v_0 := v|_k$, and (2) for each $v \in S$, $\sum_{w \in T_v} \frac{1}{g_w} = 1$, which allows us to get rid of this factor by using simultaneous approximation.

(b) If one of the $\hat{\beta}_{x_v}(L)$ is infinite then condition (8.7.1) holds for any collection $\{R_v\}_{v \in S}$ of positive numbers.

Proposition 8.8. *Let X be an irreducible variety of dimension n , L an ample line bundle on X , and $x \in X(\bar{k})$. Then*

- (a) $\hat{\beta}_x(L) \geq \frac{n}{n+1} \hat{\epsilon}_x(L)$, and
- (b) for any irreducible subvariety $Z \subseteq X$ and any $x \in Z(\bar{k})$, $\hat{\epsilon}_{x,Z}(L|_Z) \geq \hat{\epsilon}_{x,X}(L)$.

Proof: Let $\varphi: Y \rightarrow X$ be an unramified cover. By Corollary 4.4 we have $\beta_y(\varphi^*(L)) \geq \frac{n}{n+1} \epsilon_y(\varphi^*L)$ for each $y \in \varphi^{-1}(x)$. Thus $\min_{y \in \varphi^{-1}(x)} (\beta_y(\varphi^*L)) \geq \frac{n}{n+1} \min_{y \in \varphi^{-1}(x)} (\epsilon_y(\varphi^*L))$, and (a) follows after taking the supremum over such covers.

For part (b), let Z' be any irreducible component of $\varphi^{-1}(Z)$, where $\varphi: Y \rightarrow X$ is an unramified cover as above. The induced map $\psi: Z' \rightarrow Z$ expresses Z' as an unramified cover over Z , and for any $z \in \psi^{-1}(x)$ we have $\epsilon_{z,Z'}(\psi^*(L|_Z)) = \epsilon_{z,Z'}((\varphi^*L)|_{Z'}) \geq \epsilon_{z,Y}(\varphi^*L)$ by Proposition 3.4(c). Since $z \in \varphi^{-1}(x)$, this implies

$$\min_{z \in \psi^{-1}(x)} (\epsilon_{z,Z'}(\psi^*(L|_Z))) \geq \min_{z \in \psi^{-1}(x)} (\epsilon_{z,Y}(\varphi^*L)) \geq \min_{y \in \varphi^{-1}(x)} (\epsilon_y(\varphi^*L)).$$

Taking the suprema over unramified covers of Z and X we deduce (b). \square

Once Theorem 5.2 is established, the approximation results in §5—§7 follow from that theorem, Corollary 4.4, Propositions 3.4(c) and 2.15(f), as well as arguments common in Diophantine approximation. The necessary results about α , β , ϵ , and their asymptotic versions $\hat{\beta}$ and $\hat{\epsilon}$ needed to make these arguments are summarized in the following table.

Results about α , β and ϵ	Results about α , $\hat{\beta}$ and $\hat{\epsilon}$
<u>Theorem 5.2</u> $\sum_v \beta_{x_v}(L)R_v > 1 \implies (5.1.1) + (5.1.2)$	<u>Theorem 8.7</u> $\sum_v \hat{\beta}_{x_v}(L)R_v > 1 \implies (5.1.1) + (5.1.2)$
<u>Corollary 4.4</u> $\beta_x(L) \geq \frac{n}{n+1}\epsilon_x(L)$	<u>Proposition 8.8(a)</u> $\hat{\beta}_x(L) \geq \frac{n}{n+1}\hat{\epsilon}_x(L)$
<u>Proposition 3.4(c)</u> $\epsilon_{x,Z}(L _Z) \geq \epsilon_{x,X}(L)$	<u>Proposition 8.8(b)</u> $\hat{\epsilon}_{x,Z}(L _Z) \geq \hat{\epsilon}_{x,X}(L)$
<u>Proposition 2.15(f)</u> $\alpha_{x,X}(L) = \min(\alpha_{x,X_1}(L _{X_1}), \dots, \alpha_{x,X_r}(L _{X_r}))$	<u>Proposition 2.15(f)</u> $\alpha_{x,X}(L) = \min(\alpha_{x,X_1}(L _{X_1}), \dots, \alpha_{x,X_r}(L _{X_r}))$

By using Theorem 8.7 in place of Theorem 5.2 and Proposition 8.8(a,b) in place of Corollary 4.4 and Proposition 3.4(c) respectively, the arguments in §5—§7 hold with $\hat{\beta}$ and $\hat{\epsilon}$ used in place of β and ϵ . Explicitly, we have the following synthesis of the arguments in §5—§8.

Corollary 8.9. *Theorems 5.2, 6.1, 6.2, 6.3, 7.4, 7.5, and 7.7, Corollaries 5.3, 5.4, 6.4, 6.5, 6.6, 7.6, 7.10, 7.11, and 7.12 hold with β and ϵ replaced by $\hat{\beta}$ and $\hat{\epsilon}$.*

Remark. The larger the values of $\hat{\beta}$ and $\hat{\epsilon}$ the stronger these types of results are. In particular, this means that given any lower bounds for $\hat{\beta}$ and $\hat{\epsilon}$ the results listed in Corollary 8.9 hold with the lower bounds used in place of β or ϵ . One method of getting lower bounds for $\hat{\beta}$ and $\hat{\epsilon}$ which still takes into account the asymptotic behaviour of covers is to consider only étale Galois covers $\varphi: Y \rightarrow X$ with Y irreducible. By the transitivity of the Galois action, for any ample line bundle L on X , both $\epsilon_y(\varphi^*L)$ and $\beta_y(\varphi^*L)$ are independent of $y \in \varphi^{-1}(x)$, and thus we avoid worrying which point in the fibre achieves the minimum.

In particular, setting

$$\hat{\beta}_x^{\text{ét}}(L) = \sup_{\substack{\varphi: Y \rightarrow X \\ y \in \varphi^{-1}(x)}} \beta_y(\varphi^*L) \quad \text{and} \quad \hat{\epsilon}_x^{\text{ét}}(L) = \sup_{\substack{\varphi: Y \rightarrow X \\ y \in \varphi^{-1}(x)}} \epsilon_y(\varphi^*L).$$

where the suprema are over irreducible étale Galois covers $\varphi: Y \rightarrow X$, we obtain lower bounds $\hat{\beta}_x(L) \geq \hat{\beta}_x^{\text{ét}}(L)$ and $\hat{\epsilon}_x(L) \geq \hat{\epsilon}_x^{\text{ét}}(L)$ for all $x \in X(\bar{k})$ and ample L .

Example. Let X be an abelian variety and let $[m]: X \rightarrow X$ denote the multiplication by m map. For any ample line bundle L , $[m]^*L$ has the same numerical class as m^2L , and so $\epsilon_x([m]^*L) = m^2\epsilon_x(L)$ for any $x \in X(\bar{k})$. In particular, $\hat{\epsilon}_x^{\text{ét}}(L) = \infty$ and thus $\hat{\epsilon}_x(L) = \infty$. Therefore for any $x \in X(\bar{k})$, $\alpha_x(L) \geq \frac{1}{2}\hat{\epsilon}_x(L) = \infty$ by the unramified cover version of Theorem 6.3. (This gives another proof of example (c) on page 2 of the introduction.)

Remark. If X is normal then any unramified cover of X is étale, and any such cover can be dominated by a Galois étale cover. Thus if X is normal $\hat{\beta}^{\text{ét}}$ and $\hat{\epsilon}^{\text{ét}}$ agree with $\hat{\beta}$ and $\hat{\epsilon}$.

One of the themes of this article is the comparison of α and ϵ . In light of Corollary 8.9 it is natural to ask if $\hat{\epsilon}$ has same formal properties shared by α and ϵ (i.e. perhaps we have been writing the wrong article). We have not defined $\hat{\epsilon}$ when X is reducible, and we do it now by simply adopting one of the desired properties of $\hat{\epsilon}$ as the definition. If X is reducible

over k , $x \in X(\bar{k})$ and X_1, \dots, X_r the irreducible components passing through x then we set $\hat{\epsilon}_{x,X}(L) = \min(\hat{\epsilon}_{x,X_1}(L|_{X_1}), \dots, \hat{\epsilon}_{x,X_r}(L|_{X_r}))$.

Proposition 8.10. *Let X be a projective variety defined over k , $x \in X(\bar{k})$, and L a nef \mathbb{Q} -divisor on X . Consider the following assertions:*

- (a) *For any positive integer m , $\hat{\epsilon}_x(m \cdot L) = m \cdot \hat{\epsilon}_x(L)$.*
- (b) *$\hat{\epsilon}_x$ is a concave function of L : for any positive rational numbers a and b , and any nef \mathbb{Q} -divisors L_1 and L_2*

$$\hat{\epsilon}_x(aL_1 + bL_2) \geq a\hat{\epsilon}_x(L_1) + b\hat{\epsilon}_x(L_2).$$

- (c) *If Z is a subvariety of X then for any point $z \in Z(\bar{k})$ we have $\hat{\epsilon}_{z,Z}(L|_Z) \geq \hat{\epsilon}_{z,X}(L)$.*
- (d) *If L is very ample then $\hat{\epsilon}_x(L) \geq 1$, if L is ample then $\hat{\epsilon}_{x,X}(L) > 0$.*
- (e) *If x and y are points of varieties X and Y , with nef line bundles L_X and L_Y then*

$$\hat{\epsilon}_{x \times y, X \times Y}(L_X \boxplus L_Y) = \min(\hat{\epsilon}_{x,X}(L_X), \hat{\epsilon}_{y,Y}(L_Y)).$$

- (f) *Suppose that X is reducible and let X_1, \dots, X_r be the irreducible components containing x . Then $\hat{\epsilon}_{x,X}(L) = \min(\hat{\epsilon}_{x,X_1}(L|_{X_1}), \dots, \hat{\epsilon}_{x,X_r}(L|_{X_r}))$.*

Then (a), (c), (d), and (f) hold. We do not know if (b) and (e) hold in general, but they do hold when X is normal (respectively X and Y are normal).

Proof: Part (f) holds by definition of $\hat{\epsilon}$. It follows from the definition that establishing any of (a)—(e) for irreducible X implies the corresponding result for reducible X , so from now on we assume that X (or Y) is irreducible over k . Then parts (a) and (d) follow immediately from Proposition 3.4(a,d) and the definition of $\hat{\epsilon}$, and part (c) is Proposition 8.8(b).

The difficulty with (b) is that the definition of $\hat{\epsilon}$ involves the minimum over covers, and it is not clear that the minimum of all three of $\hat{\epsilon}(aL_1 + bL_2)$, $\hat{\epsilon}(L_1)$, and $\hat{\epsilon}(L_2)$ happen at the same point and can be compared. However for étale Galois covers, since we do not have to worry about the minimum, we can compare at any point and then it is clear the inequality holds by Proposition 3.4(b). Thus, in particular, (b) holds when X is normal.

Similarly, if X and Y are normal, so that again we may just consider étale Galois covers, (e) follows from Proposition 3.4(e) and the fact that such any such cover is a product of an étale Galois cover of X with an étale Galois cover of Y . (Specifically, let $\bar{X}_1, \dots, \bar{X}_r$ and $\bar{Y}_1, \dots, \bar{Y}_s$ be the irreducible components of $X \times_k \bar{k}$ and $Y \times_k \bar{k}$ respectively. Note that all \bar{X}_i and \bar{Y}_j are isomorphic over \bar{k} , and that $r = 1$ and $s = 1$ if X and Y are geometrically connected. For any étale Galois cover $\varphi: V \rightarrow X \times Y$, after passing to the algebraic closure, which we do when computing ϵ , each connected component of $V \times_k \bar{k}$ is an étale Galois cover of some $\bar{X}_i \times \bar{Y}_j$, and hence is a product of an étale Galois covers of \bar{X}_i and \bar{Y}_j . These étale Galois covers of \bar{X}_i and \bar{Y}_j may be descended to Galois covers of X and Y respectively.) \square

Remark. From the arguments for (b) and (e) above, it may seem that $\hat{\epsilon}^{\text{ét}}$ is a better substitute for $\hat{\epsilon}$, since for $\hat{\epsilon}^{\text{ét}}$ properties (b) and (e) hold for any variety, even non-normal ones. However, if X is not normal, it is not clear that property (c) holds for $\hat{\epsilon}^{\text{ét}}$. In the argument of Proposition 8.8(b) it was necessary to pass to a component of a cover of Z , and a component of an étale cover is not necessarily étale. This is one of the reasons for the definition of $\hat{\epsilon}$ as a supremum over unramified covers.

9. MORE ABOUT $\beta_x(L)$

In this section we discuss interpretations of and further results and remarks about $\beta_x(L)$. For simplicity we assume that X is irreducible and defined over an algebraically closed field.

Heuristic Interpretation of β . Let L be an ample \mathbb{Q} -bundle on X and $x \in X$. As in §4 we define a function $f(\gamma) = \text{Vol}(L_\gamma)/\text{Vol}(L)$ for $\gamma \geq 0$, and set $\gamma_{\text{eff}} = \gamma_{\text{eff},x}(L)$. The function f is decreasing with $f(0) = 1$ and $f(\gamma_{\text{eff}}) = 0$ (Figure 4b is a good illustration). By [12, Corollary C] or [2, Theorem A] the volume function is first-differentiable and hence so is f . The function $1 - f$ therefore satisfies the criteria to be a cumulative distribution function.

It is straightforward to say what the associated probability distribution is measuring. Suppose for the sake of discussion that L is an integral line bundle and base point free. For a fixed $\gamma > 0$, what is the probability that a randomly chosen section of $V = \Gamma(X, L)$ vanishes to order $\geq \gamma$ at x ? Since the set of sections vanishing to order $\geq \gamma$ at x forms a proper subspace W_γ of V , under the usual probability measure the chance is zero. However if we instead decide the ratio $\dim W_\gamma/\dim V$ is a good measure of the chance that a section of V lies in W_γ , and further decide that we should really ask the question asymptotically, that is, assign the limit $\dim W_{m\gamma}/\dim \Gamma(X, mL)$ as $m \rightarrow \infty$ as the probability of the event, then we arrive exactly at $f(\gamma)$. Therefore (under this strange distribution) $1 - f(\gamma)$ is the probability that a section vanishes to order $\leq \gamma$, and $-f'(\gamma)$ the probability density function for vanishing to order exactly γ .

The first computation one usually does when given a probability measure is to compute the expected value. Since $-f'$ is supported on $[0, \gamma_{\text{eff}}]$, and since $f(\gamma_{\text{eff}}) = 0$, integration by parts gives

$$\mathbb{E}(\gamma) = - \int_0^{\gamma_{\text{eff}}} \gamma f'(\gamma) d\gamma = -\gamma f(\gamma) \Big|_{\gamma=0}^{\gamma=\gamma_{\text{eff}}} + \int_0^{\gamma_{\text{eff}}} f(\gamma) d\gamma = -0 + 0 + \beta_x(L) = \beta_x(L).$$

This gives an interpretation of $\beta_x(L)$: under the probability distribution above $\beta_x(L)$ is the expected order of vanishing at x of a section of L .

The idea that the probability an element of a vector space V lies in a subspace W should be $\dim W/\dim V$ is counter to our intuition under the uniform measure, however it is exactly this type of probability measure which is used by Faltings-Wüstholz in the proof of their approximation theorem (see [7, §4]). Thus, with the exception of the passage to the limiting distribution, which is simply to get better control over the behaviour of the line bundle, $-f'$ is the probability measure used in the proof of the Faltings-Wüstholz approximation theorem. It is therefore completely natural that the expected order of vanishing at x governs approximation results as in Theorem 5.2.

Other results. In Corollary 4.4 we showed the inequalities $\beta_x(L) \geq \frac{n}{n+1} \sqrt[n]{\frac{\text{Vol}(L)}{\text{mult}_x X}} \geq \frac{n}{n+1} \epsilon_x(L)$, and we have used this to deduce approximation theorems involving ϵ from those involving β . If the inequalities are strict then replacing β by $\frac{n}{n+1}\epsilon$ produces a weaker result. It is therefore natural to ask when these inequalities are equalities.

Theorem 9.1. *Let X be an n -dimensional irreducible variety, $x \in X$ and L an ample \mathbb{Q} -bundle on X . Then the following conditions are equivalent.*

(a) $\beta_x(L) = \frac{n}{n+1} \sqrt[n]{\frac{\text{Vol}(L)}{\text{mult}_x X}}$

- (b) $\sqrt[n]{\frac{\text{Vol}(L)}{\text{mult}_x X}} = \epsilon_x(L)$
- (c) $\beta_x(L) = \frac{n}{n+1}\epsilon_x(L)$
- (d) $\epsilon_x(L) = \gamma_{\text{eff},x}(L)$.

Proof: To simplify the notation somewhat, set $\beta_x = \beta_x(L)$, $\omega_x = \sqrt[n]{\frac{\text{Vol}(L)}{\text{mult}_x X}}$, $\epsilon_x = \epsilon_x(L)$, and $\gamma_{\text{eff}} = \gamma_{\text{eff},x}(L)$.

(a) \implies (b): The estimate $\beta_x \geq \frac{n}{n+1}\omega_x$ resulted from integrating the lower bound $\text{Vol}(L_\gamma)/\text{Vol}(L) \geq 1 - \frac{\text{mult}_x(X)}{\text{Vol}(L)}\gamma^n$ over $[0, \omega_x]$. The equality in (a) is therefore equivalent to the two statements:

(9.1.a.1) $\text{Vol}(L_\gamma) = \text{Vol}(L) - (\text{mult}_x X)\gamma^n$ for $\gamma \in [0, \omega_x]$, and

(9.1.a.2) $\gamma_{\text{eff}} = \omega_x$.

Here (as usual) $L_\gamma = \pi^*L - \gamma E$ and $\pi: \tilde{X} \rightarrow X$ is the blow up of X at x with exceptional divisor E . We will see that (9.1.a.1) implies (b). We first recall an extension of the idea of volume to arbitrary cohomology groups. For any line bundle M on an n -dimensional variety Y , and any $0 \leq i \leq n$ we set

$$\hat{h}^i(M) = \lim_{m \rightarrow \infty} \frac{\dim H^i(Y, mM)}{m^n/n!}$$

so that $\hat{h}^0(M) = \text{Vol}(M)$. As in the case of the volume, the groups \hat{h}^i depend only on the numerical class of M , make sense for \mathbb{Q} -divisors, and for fixed i extend to continuous functions on $\text{NS}(Y)_{\mathbb{R}}$ (see [9, p. 1477]). We will also need a slight variation of this idea. As in §4 for any rational $\gamma > 0$ and m such that $m\gamma$ is an integer we denote by $m\gamma E$ the subscheme defined by the $(m\gamma)$ -th power of the defining equation for E . For any $0 \leq i \leq n$ we set

$$\hat{h}^i(\mathcal{O}_{\gamma E}) = \lim_{m \rightarrow \infty} \frac{\dim H^i(\tilde{X}, \mathcal{O}_{m\gamma E})}{m^n/n!}$$

where the limit runs over all m such that $m\gamma$ is an integer. Note that “ $\hat{h}^i(\mathcal{O}_{\gamma E})$ ” is being defined as an atomic symbol — we are not giving any meaning to $\mathcal{O}_{\gamma E}$ as a scheme. Since $\mathcal{O}_E(-E)$ is ample on E , it follows from Serre vanishing and (4.1.2) that $\hat{h}^i(\mathcal{O}_{\gamma E}) = 0$ for all $i > 0$. Combined with this, the argument in the proof of Lemma 4.1 actually shows that $\hat{h}^0(\mathcal{O}_{\gamma E}) = (\text{mult}_x X)\gamma^n$.

The asymptotic cohomology groups are birational invariants. Since L is ample, $h^i(X, mL) = 0$ for all $m \gg 0$, and hence (pulling back to \tilde{X}) $\hat{h}^i(L_0) = 0$ for all $i > 0$. The long exact sequence associated to (4.1.1) then implies that for any rational $\gamma \geq 0$, $\hat{h}^i(L_\gamma) = 0$ for all $i \geq 2$ and that

$$\text{Vol}(L_\gamma) - \hat{h}^1(L_\gamma) = \text{Vol}(L) - \hat{h}^0(\mathcal{O}_{\gamma E}) = \text{Vol}(L) - (\text{mult}_x X)\gamma^n.$$

Thus (9.1.a.1) is equivalent to the statement that $\hat{h}^1(L_\gamma) = 0$ for all $0 \leq \gamma \leq \omega_x$.

Let A be any ample bundle on \tilde{X} . By [4, Theorem A] L_γ is ample if and only if $\hat{h}^i(L_\gamma - tA) = 0$ for all $i > 0$ and all sufficiently small t . Let s be any number $0 < s < \epsilon$ so that $A = L_s$ is ample on \tilde{X} . Then $L_\gamma - tA = (1-t)\pi^*L - (\gamma-ts)E = (1-t)L_{\frac{\gamma-ts}{1-t}}$. The asymptotic cohomology

groups are homogeneous of degree n , so

$$\hat{h}^i(L_\gamma - tL_s) = \hat{h}^i\left((1-t)L_{\frac{\gamma-ts}{1-t}}\right) = (1-t)^n \hat{h}^i\left(L_{\frac{\gamma-ts}{1-t}}\right)$$

for all $i \geq 0$. If $0 < \gamma < \omega_x$, then for small enough t we have $0 \leq \frac{\gamma-ts}{1-t} < \omega_x$ too, and hence by (9.1.a.1) and the equation above $\hat{h}^i(L_\gamma - tA) = 0$ for all $i > 0$.

Summarizing, condition (9.1.a.1) and Theorem A of [4] imply that L_γ is ample for all $0 < \gamma < \omega_x$. Thus $\omega_x \leq \epsilon_x$. The opposite inequality, $\epsilon_x \leq \omega_x$, is [11, Proposition 5.1.9] (this already appeared in the proof of Corollary 4.2) and thus $\epsilon_x = \omega_x$, i.e., (b) holds.

(b) \implies (c)+(d): Since $\text{Vol}(L_\gamma) = \text{Vol}(L) - (\text{mult}_x X)\gamma^n$ for $\gamma \in [0, \epsilon_x]$, and since condition (b) is that $\omega_x = \epsilon_x$, we have

$$(9.1.b.1) \quad \text{Vol}(L_\gamma) = \text{Vol}(L) - (\text{mult}_x X)\gamma^n \text{ for } \gamma \in [0, \omega_x].$$

Condition (9.1.b.1) shows that $\text{Vol}(L_\gamma) > 0$ for $0 \leq \gamma < \omega_x$, and that $\text{Vol}(L_{\omega_x}) = 0$, hence ω_x is the boundary of the effective cone, i.e.,

$$(9.1.b.2) \quad \gamma_{\text{eff}} = \omega_x.$$

Given these two conditions,

$$\beta_x = \int_0^{\gamma_{\text{eff}}} \text{Vol}(L_\gamma) / \text{Vol}(L) d\gamma = \int_0^{\omega_x} 1 - \frac{\text{mult}_x X}{\text{Vol}(L)} \gamma^n d\gamma = \frac{n}{n+1} \omega_x = \frac{n}{n+1} \epsilon_x.$$

Thus (c) holds. Since (d) is condition (9.1.b.2) it is also clear that (b) implies (d).

(c) \implies (a)+(b): This is clear from the inequalities $\beta_x \geq \frac{n}{n+1} \omega_x \geq \frac{n}{n+1} \epsilon_x$.

(d) \implies (b): This is immediate from the inequalities $\gamma_{\text{eff}} \geq \omega_x \geq \epsilon_x$. \square

Remark. Condition (b) of Theorem 9.1 seems the easiest one to check in practice. Condition (d) is also tractable; it is the statement that along the ray $\pi^*L - \gamma E$ ($\gamma \geq 0$), the point where the ray exits the nef cone is the same point where the ray exits the effective cone.

Seshadri Exceptional Subvarieties. Recall that by [11, Proposition 5.1.9] for any irreducible subvariety $V \subseteq X$ of positive dimension passing through x we have the inequality

$$(9.1.1) \quad \epsilon_x(L) \leq \left(\frac{c_1(L)^{\dim V} \cdot V}{\text{mult}_x V} \right)^{\frac{1}{\dim V}},$$

and that there are irreducible subvarieties V for which (9.1.1) is an equality (including possibly $X = V$). An irreducible subvariety V is called *Seshadri exceptional* (with respect to x and L) if (9.1.1) is an equality, and if V is not properly contained in a larger subvariety having the same property. Condition (b) of Theorem 9.1 is that X itself is Seshadri exceptional.

Further properties of $\beta_x(L)$. As in previous sections, it is interesting to work out some formal properties of β_x , in particular to ask whether the list of properties in Propositions 2.15 and 3.4 hold. We do not know the status of all the properties listed there, and simply record some elementary observations. (The letters match those of Propositions 2.15 and 3.4.)

Proposition 9.2. *$x \in X$, L an ample line bundle on X , then*

$$(a) \quad \beta_x(mL) = m\beta_x(L).$$

(c) *If Z is a subvariety of X , $x \in Z$, it is not necessarily true that $\beta_{x,Z}(L|_Z) \geq \beta_{x,X}(L)$.*

- (d) If L is ample then $\beta_x(L) > 0$.
- (f) Suppose that X is reducible and let X_1, \dots, X_r be the irreducible components containing x . Then $\beta_{x,X}(L) = \min(\beta_{x,X_1}(L|_{X_1}), \dots, \beta_{x,X_r}(L|_{X_r}))$.

Proof: Property (f) holds by definition of β_x (Definition 4.3), and (d) is clear from the estimate $\beta_x(L) \geq \frac{n}{n+1}\epsilon_x(L)$ and Proposition 3.4(d). For part (a), fix $m > 0$ and let $f_L(\gamma)$ and $f_{mL}(\gamma)$ be the functions $f_L(\gamma) = \text{Vol}(L_\gamma)/\text{Vol}(L)$ and $f_{mL}(\gamma) = \text{Vol}((mL)_\gamma)/\text{Vol}(mL)$ respectively. On an n -dimensional variety one has $\text{Vol}(mM) = m^n \text{Vol}(M)$ for every big line bundle M and $m > 0$ and hence

$$f_{mL}(m\gamma) = \text{Vol}((mL)_{m\gamma})/\text{Vol}(mL) = \text{Vol}(mL_\gamma)/\text{Vol}(mL) = \frac{m^n}{m^n} \text{Vol}(L_\gamma)/\text{Vol}(L) = f_L(\gamma).$$

It follows from this equation or directly from the definition that $\gamma_{\text{eff},x}(mL) = m\gamma_{\text{eff},x}(L)$. Integrating (and using the previous equation) we conclude that $\beta_x(mL) = m\beta_x(L)$.

Finally, to see that β_x may strictly decrease under restriction, recall that $\beta_x(\mathcal{O}_{\mathbb{P}^n(1)}) = \frac{n}{n+1}$ for any point $x \in \mathbb{P}^n$ (see the example on page 25). Hence if Z is an m -dimensional linear subspace of $X = \mathbb{P}^n$ passing through x (with $m < n$) and $L = \mathcal{O}_{\mathbb{P}^n}(1)$ then $\beta_{x,Z}(L|_Z) < \beta_x(L)$. \square

Remark. The fact that ϵ_x is weakly increasing under restriction has been crucial for our inductive arguments. The fact β_x may decrease under restriction to a subvariety is one reason why this article is focussed on ϵ_x , and why it was important to estimate β_x in terms of ϵ_x .

10. A SPECIAL CASE OF VOJTA'S MAIN CONJECTURE

Vojta's Main Conjecture (Conjecture 3.4.3 of [21]) predicts how the height of rational points grow as they approach a simple normal crossings divisor $D \subset X$. One can also investigate the prediction for other subvarieties of X , with the result being stronger for larger subvarieties. Theorem 7.4 easily implies many special cases of the Main Conjecture, albeit ones where the subvariety in question is a collection of points (this is natural since the results of this paper are geared towards approximating points). Despite the fact that this is a weaker version than the classical case in which D is a divisor, many of the cases established below were previously unknown.

We refer the reader to [21, §3] for a statement and discussion of the Main Conjecture, and simply state the relevant result in the language of this paper.

Theorem 10.1. *Let k be a number field, X an irreducible n -dimensional variety over $\text{Spec}(k)$ such that $-K_X$ is ample, D a finite subset of $X(\bar{k})$, and S a finite set of places of k . If $\epsilon_x(-K_X) > \frac{n+1}{n}$ for every $x \in D$, then Vojta's Main Conjecture is true for X and D . Specifically, for every $\delta > 0$ and any big divisor A , there is a closed subset $Z \subset X$ such that for all k -rational points $P \in X(k) \setminus Z(k)$, we have:*

$$\sum_{v \in S, x \in D} -\log d_v(x, P) + h_{K_X}(P) < \delta h_A(P) + O(1)$$

Proof: If we can show the inequality for one big divisor A , then it will immediately follow for an arbitrary big divisor A , by adjusting δ and Z . Thus, we may assume that $A = -K_X$. Furthermore, note that for any place $v \in S$, there is at most one point in D for which $-\log d_v(x, P)$ contributes more than a bounded amount to the sum. Therefore, we may

apply Theorem 7.4(b) (in the equivalent form of (7.3.4)) with $R_v = 1$ for each $v \in S$ to see that there is a proper subset Z so that for any $\delta > 0$ the equation

$$\prod_{v \in S, x \in D} d_v(x, P) > H_{-K_X}(P)^{-(1+\delta)}$$

holds for all but finitely many $P \in X(k) \setminus Z(k)$. Taking log then gives the result. \square

Remark. For k -points x of D it is sufficient that the weaker condition $\epsilon_x(-K_X) \geq 1$ hold. One uses the Liouville bound $\alpha_x(-K_X) \geq \epsilon_x(-K_X)$ (valid for points of $X(k)$ – see [15]) in a simultaneous approximation version similar to Corollary 7.6.

There are many examples of varieties X which satisfy the criterion of the theorem. For example for any variety of the form $X = G/P$ where G is a semi-simple algebraic group and P is a parabolic subgroup (e.g., \mathbb{P}^n or Grassmannians) one has $\epsilon_x(-K_X) \geq 2$ for all points $x \in X(\bar{k})$.

REFERENCES

- [1] Bombieri, E.; Gubler, W., *Heights in Diophantine Geometry*, New Mathematical Monographs 4, Cambridge University Press, Cambridge, 2006.
- [2] Boucksom, S.; Favre, C.; Jonsson, M., *Differentiability of volumes of divisors and a problem of Tessier*, J. Algebraic Geom. **18** (2009), no. 2, 279–308.
- [3] Choi, K.K.; Vaaler, J.D., *Diophantine Approximation in Projective Space*, Number theory (Ottawa, 1996), CRM Proc. Lecture Notes, **19**, Amer. Math. Soc., Providence, RI, 1999, 55–65.
- [4] de Fernex, T.; K\"{u}ronya, A.; Lazarsfeld, R., *Higher cohomology of divisors on a projective variety*, Math. Ann. **337** (2007), no. 2, 443–455.
- [5] Demailly, J.-P., *Singular Hermitian metrics on positive line bundles*, Complex Algebraic Varieties (Bayreuth, 1990), Lect. Notes in Math., vol 1507, 1992, 87–104.
- [6] Dieudonn\'{e}, J.; Grothendieck, A. *\'{E}l\'{e}ments de G\'{e}om\'{e}trie Alg\'{e}brique. III. \'{E}tude cohomologique des faisceaux coh\'{e}rents.*, Inst. Hautes \'{E}tudes Sci. Publ. Math., No. 11, 1961.
- [7] Faltings, G.; W\"{u}stholz, G., *Diophantine approximations on projective spaces*, Invent. Math. **116** (1994), 109–138.
- [8] Hindry, M.; Silverman, J., *Diophantine geometry. An introduction.*, Graduate Texts in Mathematics, 201. Springer-Verlag, New York, 2000.
- [9] K\"{u}ronya, A., *Asymptotic cohomological functions on projective varieties*, Amer. J. Math. **128** (2006), no. 6, 1475–1519.
- [10] Lang, S., *Diophantine Geometry*, Interscience Tracts in Pure and Applied Mathematics, No. 11, John Wiley & Sons, New York-London, 1962.
- [11] Lazarsfeld, R., *Positivity in Algebraic Geometry I*, Springer-Verlag, 2004.
- [12] Lazarsfeld, R.; Musta\c{t}\u0107, M.; *Convex bodies associated to linear series*, Ann. Sci. \'{E}c. Norm. Sup\'{e}r. (4) **42** (2009), no. 5, 783–835.
- [13] Macintyre, A. *Non-standard number theory*, Proc. International Congress Math. Helsinki, Vol. 1 (1978), 253–262.
- [14] McKinnon, D., *A conjecture on rational approximations to rational points*, J. Algebraic Geom., **16** (2007), 257–303.
- [15] McKinnon, D.; Roth, M., *An analogue of Liouville’s theorem and an application to cubic surfaces*, submitted.
- [16] Neukirch, J. *Algebraic Number Theory*, Springer-Verlag, 1999.
- [17] Robinson, A.; Roquette, P. *On the finiteness theorem of Siegel and Mahler concerning Diophantine equations*, J. Number Theory **7** (1975), 121–176.
- [18] Schmidt, W., *Diophantine Approximation*, Lecture Notes in Mathematics 785, Springer, 1980.
- [19] Serre, J.-P., *Lectures on the Mordell-Weil Theorem*, Vieweg, 1997.
- [20] Soul\'{e}, C.; Abramovich, D.; Burnol, J.-F.; Kramer, J., *Lectures on Arakelov Geometry*, Cambridge University Press, 1994.

- [21] Vojta, P., *Diophantine Approximations and Value Distribution Theory*, Lecture Notes in Mathematics vol. 1239, Springer-Verlag, 1987.

DEPARTMENT OF PURE MATHEMATICS, UNIVERSITY OF WATERLOO, WATERLOO, ON N2L 3G1,
CANADA

E-mail address: dmckinnon@math.uwaterloo.ca

DEPT. OF MATHEMATICS AND STATISTICS, QUEENS UNIVERSITY, KINGSTON, ONTARIO, CANADA

E-mail address: mikeroth@mast.queensu.ca