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 $\alpha$ 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 $X$ be an irreducible projective variety defined over a number field $k$. 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 $X$, also defined over $k$, and $x$ a point of $X(\overline{k})$. By slightly modifying the usual definition of approximation exponent on $\mathbb{P}^1$ (and inspired by a definition from [12] 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 [4] 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 $\alpha_x$ is interpreted through Arakelov theory. Moreover, the classic approximation results on $\mathbb{P}^1$ — the theorems of Liouville and Roth — generalize as inequalities between $\alpha_x$ and $\epsilon_x$ 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.
Approximation on $\mathbb{A}^1$. For a point $x \in \mathbb{R}$ the approximation exponent of $x$ is defined as the smallest real number $\tau_x$ such that for any $\delta > 0$ the inequality
\[ |x - a/b| \leq \frac{1}{b^{\tau_x + \delta}} \]
has only finitely many solutions $a/b \in \mathbb{Q}$ (see [7, 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 $\tau_x > 1$. One therefore seeks upper bounds on $\tau$. 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 2.9 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 2.12(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 $\alpha$ given by previously known results on Diophantine approximation. We work over an arbitrary number field $k$.

(a) If $X = \mathbb{P}^1$, $x \in X(\overline{k})$, $L = \mathcal{O}_{\mathbb{P}^1}(1)$, then $\alpha_x(L) \geq \frac{1}{2}$.

(b) If $X = \mathbb{P}^n$, $x \in X(\overline{k})$, $L = \mathcal{O}_{\mathbb{P}^n}(1)$, then either $\alpha_x(L) \geq \frac{m}{n+1}$ or there is a smaller linear space $Z \cong \mathbb{P}^m \subset \mathbb{P}^n$, with $m < n$, defined over $k$ and containing $x$.

(c) If $X$ is an abelian variety, $x \in X(\overline{k})$, 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 one obtains exact values for $\alpha_x$. In the case of $\mathbb{P}^1$, if $x \in \mathbb{P}^1(\overline{k} \cap k_v) \setminus \mathbb{P}^1(k)$ then $\alpha_x(\mathcal{O}_{\mathbb{P}^1}(1)) = \frac{1}{2}$. In the case of $\mathbb{P}^n$, if $x \in \mathbb{P}^n(\overline{k} \cap k_v) \setminus \mathbb{P}^n(k)$, and $m$ is the smallest value so that there exists a linear subspace of dimension $m$ defined over $k$ and passing through $x$, then $\alpha_x(\mathcal{O}_{\mathbb{P}^n}(1)) = \frac{m}{m+1}$. Finally example (c) is [16, p. 98; second theorem].

The basic interpretation of $\alpha_x(L)$ is as the cost in complexity required to get closer to $x$. When $\alpha_x$ is finite this indicates that the complexity has polynomial growth in the reciprocal of the distance, and $\alpha_x$ is the exponent. In example (c) the complexity grows roughly exponentially in the reciprocal of the distance (see [16, 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 $\alpha$, here in order to emphasise the similarity between $\alpha_x$ and $\epsilon_x$, and to use one of the properties in the discussion below. Both $\alpha$ and $\epsilon$ make sense for $\mathbb{Q}$-bundles (and in fact for $\mathbb{R}$-bundles). Fix $x \in X(\overline{k})$, then

(a) for any ample $\mathbb{Q}$-bundle $L$, and any $m \in \mathbb{Q}_{>0}$,
\[
\alpha_x(mL) = m\alpha_x(L) \quad \text{and} \quad \epsilon_x(mL) = m\epsilon_x(L);
\]

(b) $\alpha_x$ and $\epsilon_x$ are concave functions of the line bundle. For any ample $\mathbb{Q}$-bundles $L_1$ and $L_2$, and any $a, b \in \mathbb{Q}_{\geq 0}$,
\[
\alpha_x(aL_1 + bL_2) \geq a \alpha_x(L_1) + b \alpha_x(L_2) \quad \text{and} \quad \epsilon_x(aL_1 + bL_2) \geq a \epsilon_x(L_1) + b \epsilon_x(L_2).
\]

These and other parallel properties appear in Propositions 2.12 and 3.4.

Examples.

(a) If $X = \mathbb{P}^n$, $x \in X$, 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 $x \in X$. One thesis of this paper is that $\epsilon_x$ affects approximation results. On varieties where $\epsilon_x$ 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 $\epsilon$ into account. This will be a feature of the results below.

Roth theorems.\(^1\) If $X$ is defined over $k$, and $x \in X(\overline{k})$ 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 $\alpha_x(\mathcal{O}_{\mathbb{P}^1}(1)) \geq \frac{1}{2}$, 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, and we do not prove it here (see [13] for a proof).

Our main concern is proving general “Roth-type” theorems. By this we mean lower bounds on $\alpha_x(L)$ that are: (1) independent of the field of definition of $x$, and (2) 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 $x \in X(\overline{k})$, either $\alpha_x(L) \geq c_n \epsilon_x(L)$ or there is a proper subvariety $Z$, containing $x$ and defined over $k$, 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 $x \in X(\overline{k})$ 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 of theorem.

\(^1\)All 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, “Roth theorem II”): Let $X$ be an irreducible $n$-dimensional variety defined over $k$. For any ample $\mathbb{Q}$-bundle $L$ and any $x \in X(k)$ either

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

or

(b) there exists a proper subvariety $Z \subset X$ defined over $k$, irreducible over $\overline{k}$, and containing $x$ so that $\alpha_{x,Z}(L) = \alpha_{x,Z}(L|Z)$, 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): There is a proper subvariety $Z \subset X$ defined over $k$ so that for each $\delta > 0$ there are only finitely many solutions $y \in X(k) \setminus Z(k)$ to

$$d_v(x,y) < H_L(y)^{-\frac{m+1}{n+1}(\epsilon_x^2 + \delta)}.$$

As far as approximating a point goes, Theorem 6.2 also generalizes the Schmidt subspace theorem. 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 theorem III”): For all varieties $X$ defined over $k$ (possibly reducible), all $x \in X(k)$ 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 than there is a 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

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

Heuristic explanation. Given an ample line bundle $L$ and $x \in X(k)$ 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 $M$ is a very ample line bundle, then $\epsilon_x(M) \geq 1$ for all $x' \in X$ (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) $mL$ may not be an integral (or even 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 $x \in X$ a point not on a line. As stated above $\epsilon_x(L) = \frac{2}{3}$. However $\frac{2}{3}L$ is not an integral line bundle (it has degree $\frac{2}{3}$ on every line), nor is there any smaller multiple of $L$ which is both integral and very ample (we ask for smaller because this would give a better value of the exponent).

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 §6, 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 $k$. 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 $x_v \in X(k)$. To simplify notation, we set $\alpha_v$ to be $\alpha_{x_v}$ 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, $\sum_{v \in S} \frac{\epsilon_x(L)}{\alpha_v(\{x_i\}, L)} \leq 2$, and

2. for any $\delta > 0$ there are only finitely many $y \in X(k)$ such that

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

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 $x \in X(k)$ we define $\tilde{\epsilon}_x^\text{ét}(L)$ by

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

where the limit is over all irreducible étale covers $\varphi: Y \to X$. This limit exists and it is always true that $\tilde{\epsilon}_x^\text{ét}(L) \geq \epsilon_x(L)$ (this is a consequence of Lemma 8.1). In §8 we show that all the previous theorems, for one place or simultaneous places, hold with $\epsilon_x(L)$ replaced by $\tilde{\epsilon}_x^\text{ét}(L)$ (see Corollary 8.9). Since $\tilde{\epsilon}^\text{ét}$ 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 $\tilde{\epsilon}_x^\text{ét}(L) = \infty$ (see the example on page 46). Thus Theorem 6.3 applied with $\tilde{\epsilon}_x^\text{ét}$ in place of $\epsilon$ 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 limit over irreducible unramified covers) and the reader is referred there for more detailed statements.

In his 1962 book *Diophantine Geometry*, Lang ([9, 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 is (paraphrasing) 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 $\beta_x(L)$ (which we also label a “Roth theorem”, despite the fact that $\epsilon_x(L)$ is not involved).

**Theorem (6.1, “Roth theorem I”):** Let $X$ be an irreducible variety defined over $k$. Then for any ample $\mathbb{Q}$-bundle $L$ and any $x \in X(\overline{k})$ either

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

or

(b) $\alpha_x(L)$ 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 $x_v \in X(\overline{k})$ 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 determine the best possible result (for the type of approximation theorem we are investigating) one can deduce from Faltings-Wüstholz. 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 [6, §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 $\beta_{x}$ — 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 is 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 $\beta_{x}$ and $\epsilon_{x}$ replaced by their limits $\beta_{x}$ and $\epsilon_{x}$ 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” we mean a (possibly reducible, possibly singular) projective variety. For varieties defined over $k$, “irreducible” means irreducible over $k$, and without further specification, line bundles $L$ are also defined over $k$. We use additive notation for line bundles since this is in line with the behaviour of $\alpha_{x}$, $\epsilon_{x}$, and $\beta_{x}$. 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 defined over $k$, a point $x \in X(\overline{k})$ is a map $\text{Spec}(\overline{k}) \rightarrow X$ of $k$-schemes. Such a point gives rise to a point of $X \times_k \overline{k}$, 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 $x$. 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 $k$, $\pi$ a uniformizer of the corresponding maximal ideal, and $\kappa$ 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 $d_{v}(\cdot, \cdot)$, with domain $X(\overline{k}) \times X(\overline{k})$, to height functions $H_{L}(\cdot)$, and to partially evaluated distance functions $d_{v}(x, \cdot)$, the last two with domain $X(\overline{k})$. We also use this terminology for open subsets $U \subset X$ and restrictions of the distance or height functions to $U$, and say that the functions are “equivalent on $U$”.

**Acknowledgements.** We thank Chris Dionne, Laurence Ein, Robert Lazarsfeld, Victor Lozovanu, and Damien Roy for helpful discussions. We also wish to acknowledge an intellectual debt to Michael Nakamaye who has long advocated that Seshadri constants are diophantine.
2. Approximation by rational points

Let $k$ be a number field, and $X$ a projective variety defined over $k$. We begin by explaining the distance functions in the archimedean and non-archimedean cases.

**Distance Functions: Archimedean case.** Fix an archimedean place $v$ of $k$, and an extension of $v$ to $\overline{k}$, which we also denote by $v$. We choose a distance function on $X(\overline{k})$ by choosing an embedding $X \rightarrow \mathbb{P}^n_k$ defined over $k$, and pulling back (via $v$) the distance function on $\mathbb{P}^n(\mathbb{C})$ given by the Fubini-Study metric on $\mathbb{P}^n$. We denote this distance by $\text{dist}(\cdot, \cdot)$.

The decision to use the Fubini-Study metric via an embedding to give a distance function is essentially an arbitrary one, and any equivalent distance function will give the same results. However, we do indicate two features of this choice. First, $\text{dist}$ is essentially an arbitrary one, and any equivalent distance function will give the same results. Thus choosing the distance function induced by the Fubini-Study metric and an arbitrary $P$ containing $X$ of $P$ is equivalent to choosing the distance function induced by the Fubini-Study metric and an arbitrary $P$ centered at $P$.

We now prove the existence of such an open set $U$ around an arbitrary point $P$ of $X$. The variety $X$ may be singular, and by “analytic open set”, and “holomorphic” (and similar terms) we mean in the sense of complex analytic spaces.

Suppose that $P = [p_0 : \cdots : p_n]$ is a point of $\mathbb{P}^n$ and $p_0 \neq 0$. Then any points $p$, $q$ not on the hyperplane $z_0 = 0$ can be written in the form $p = P + x$, $q = P + y$ with $x = [0 : x_1 : \cdots : x_n]$ and $y = [0 : y_1 : \cdots : y_n]$. A slightly awkward but standard local computation shows that in a sufficiently small neighbourhood of $P$ the Fubini-Study distance is equivalent to the Euclidean distance function $\sqrt{\sum_{i=1}^n ||x_i - y_i||^2}$.

Now let $P$ be a point of $X$, $U$ a sufficiently small open set around $P$, and $d$ and $d'$ the restrictions to $U$ of the Euclidean distance functions obtained from the embeddings giving $\text{dist}$ and $\text{dist}'$ — that is, distance functions obtained from an affine chart of the embedding centered at $P$. By the previous paragraph, up to shrinking $U$, it is sufficient to prove that there exists a constant $C$ so that $d'(x, y) \leq C d(x, y)$ on $U \times U$.

Suppose that the embedding $U \hookrightarrow \mathbb{A}^n$ inducing $d$ is given by the holomorphic functions $z_1, \ldots, z_n$ on $U$. Let $p_j : U \times U \rightarrow U$, $j = 1, 2$ be the two projections. On $U \times U$ define the functions $u_i$, $i = 1, \ldots, n$ by $u_i = pr_1^*(z_i) - pr_2^*(z_i)$ and set $\mathbf{u} = (u_1, \ldots, u_n)$. Then we have the equality of functions $d^2 = \sum ||u_i||^2 = \mathbf{u} \cdot \mathbf{u}$ on $U \times U$.

Similarly, suppose that the embedding $U \hookrightarrow \mathbb{A}^m$ inducing $d'$ is given by the holomorphic functions $z_1', \ldots, z_m'$, and set $w_j = pr_1^*(z_j') - pr_2^*(z_j')$ for $j = 1, \ldots, m$. Writing $\mathbf{w} = (w_1, \ldots, w_m)$ we have $(d')^2 = \sum_j ||w_j||^2 = \mathbf{w} \cdot \mathbf{w}$ on $U \times U$.
The functions $u_1, \ldots, u_n$ cut out the diagonal on $U \times U$ (since they are the restrictions from $\mathbb{A}^n \times \mathbb{A}^n$ of the functions defining the diagonal of $\mathbb{A}^n$). Similarly $w_1, \ldots, w_m$ also cut out the diagonal on $U \times U$, and so we have the equality of ideals $(u_1, \ldots, u_n) = \mathcal{I}_\Delta = (w_1, \ldots, w_m)$.

In particular, each $w_j$ is in the ideal generated by the $u_i$. Since the $w_j$ are the entries of $w$ this means that there exist $v_1, \ldots, v_n$, each an $m$-tuple of holomorphic functions on $U$, such that $w = \sum_i u_i v_i$. But then

$$(d')^2 = w' \cdot w = \sum_{i,j=1}^n \overline{u}_i u_j (\overline{v}_i \cdot v_j).$$

Thus, if we define the $n \times n$ matrix $M$ of $C^\infty$ functions on $U \times U$ by $M_{ij} = \overline{v}_i \cdot v_j$ we have $w' \cdot w = \overline{w} M u$. The matrix-valued function $M$ is Hermitian, and by the Gershgorin theorem the magnitude of the largest eigenvalue is no more than $\sum_{i,j} \| M_{ij} \|$. Pick a compact set $S$ contained in $U$ such that $P$ is in the (nonempty) interior, and let $C$ be the maximum of $\sum \| M_{ij} \|$ on $S \times S$. Shrinking $U$ so that it lies within $S$ we then have $w' \cdot w = \overline{w} M u \leq C (\overline{w} \cdot u)$ on $U \times U$. In other words, we've shown that $d' \leq \sqrt{C} d$ on $U \times U$. □

In this paper, it will be convenient to renormalize this notion of distance depending on the archimedean place $v$. Set $m_v = 1$ if $v(k)$ is real and $m_v = 2$ if $v(k)$ is complex. For $x, y \in X(k)$ we define $d_v(x, y) = \text{dist}(x, y)^{m_v}$. This normalization means that $d_v(\cdot, \cdot)$ is in general no longer a distance function. Note that functions $d_v$ defined via different projective embeddings are again equivalent.

**Distance Functions: Non-archimedean case.** Fix a non-archimedean place $v$ of $k$, and an extension of $v$ to $\overline{k}$, which we also denote by $v$. Let $k_v$ be the completion of $k$ with respect to $v$ and $\mathcal{O}_{k,v}$ the completion of $\mathcal{O}_k$ at the maximal ideal corresponding to $v$. Similarly, for any finite extension $F/k$ let $F_v$ and $\mathcal{O}_{F,v}$ denote the completions of $F$ and $\mathcal{O}_F$ with respect to $v$. The extension of the original place to $\overline{k}$ is equivalent to the data of an embedding $\overline{k} \hookrightarrow F_v$, and induces inclusions $\mathcal{O}_{k,v} \hookrightarrow \mathcal{O}_{F,v}$ and $k_v \hookrightarrow F_v \hookrightarrow \overline{k}_v$ for all finite extensions $F/k$.

Let $\mathfrak{x}$ be a projective integral model of $X$ over $\text{Spec}(\mathcal{O}_k)$ and set $\mathfrak{x}_v$ to be the base-change of $\mathfrak{x}$ to $\mathcal{O}_{k,v}$. We define a distance function $d_v(\cdot, \cdot)$ on points of $X(\overline{k})$ based on this integral model. We first define $d_v$ in the case that both points are in $X(k)$.

Suppose that $x, y \in X(k)$, then $x$ and $y$ give rise to sections $\sigma_x$ and $\sigma_y$ of $\mathfrak{x}_v$ over $\text{Spec}(\mathcal{O}_{k,v})$. Our distance function will measure the contact between these two sections. If $x = y$ we set $d_v(x, y) = 0$. If $x \neq y$, then let $Z$ be the scheme of intersection of $\sigma_x$ and $\sigma_y$ 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 $x, y \in X(\overline{k})$, 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 $\text{Spec}(\mathcal{O}_{F,v})$. As before, $x$ and $y$ give rise to sections $\sigma_x$ and $\sigma_y$ of $\mathfrak{x}_{F,v}$ over $\text{Spec}(\mathcal{O}_{F,v})$. If $x = y$ then set $d_v(x, y) = 0$. Otherwise let $Z$ be the scheme of intersection, and set $d_v(x, y) = 1/(\# \Gamma(Z, \mathcal{O}_Z))$. If $F' \mid F$ is any finite extension, and $Z'$ the scheme of intersection of the corresponding sections $\sigma'_x$ and $\sigma'_y$ of $\mathfrak{x}_{F',v}$ then $\# \Gamma(Z', \mathcal{O}_{Z'}) = \# \Gamma(Z, \mathcal{O}_Z)$. It follows that $d_v$ is well defined.

We now prove a slightly weaker version of Proposition 2.1 in the non-archimedean case.
Proposition 2.2. Let $d_v$ and $d'_v$ be two distance functions coming from different integral models of $X$, and assume that $X$ is irreducible. Then $d_v$ is equivalent to $d'_v$.

Proof: Let $\mathfrak{X}$ and $\mathfrak{X}'$ be the two (projective) integral models. By replacing $\mathfrak{X}'$ by the closure of the diagonal embedding of $X$ in $\mathfrak{X} \times_{\mathcal{O}_h} \mathfrak{X}'$ we are reduced to the case that we have a birational morphism $\mathfrak{X}' \to \mathfrak{X}$ between the integral models which is the identity on the generic fibre.

Base changing to $\hat{\mathcal{O}}_{k,v}$ we get a projective birational morphism $\pi: \mathfrak{X}' \to \mathfrak{X}_v$ which is again the identity on the generic fibre. By [5, Corollaire (2.3.5)] this means that $\pi$ is the blow-up along a sheaf of ideals $\mathcal{I}_\pi$ on $\mathfrak{X}$ (this is where we use irreducibility). Let $W$ be the scheme defined by $\mathcal{I}_\pi$. Since $\pi$ is an isomorphism on the generic fibre, $W$ is set-theoretically supported on the special fibre, and hence there is an integer $\ell$ such that $W$ is contained in the $(\ell - 1)^{st}$ order neighbourhood of the special fibre. Equivalently, setting $m$ to be the maximal ideal of $\hat{\mathcal{O}}_{k,v}$ and $\varphi: \mathfrak{X}_v \to \text{Spec}(\hat{\mathcal{O}}_{k,v})$ to be the structural morphism, there is an $\ell$ such that $\varphi^*m^\ell \subseteq \mathcal{I}_\pi$. Let $\kappa(v) = \hat{\mathcal{O}}_{k,v}/m$. We will show that

\begin{equation}
(2.2.1) \quad \frac{d'_v(x,y)}{\kappa(v)} \leq d_v(x,y) \leq d'_v(x,y) \quad \text{for all } x,y \in X(\overline{k}).
\end{equation}

If $x = y$ the inequality is obvious; therefore we assume throughout the rest of the discussion that $x \neq y$. As in the definition of $d_v$ we first study the case that the points $x, y$ are in $X(k)$.

Suppose that $x, y \in X(k)$ and let $\sigma_x, \sigma_y$ be the corresponding sections of $\mathfrak{X}_v$. We define $\text{ord}(\sigma_x, \sigma_y)$ (the “order of contact”) to be the largest $r \geq 0$ such that $\sigma_x = \sigma_y$ when restricted to $\text{Spec}(\hat{\mathcal{O}}_{k,v}/m^r)$. (So that $r = 0$ means that $\sigma_x$ and $\sigma_y$ don’t meet, $r = 1$ that $\sigma_x$ and $\sigma_y$ do meet but have different tangent directions, etc.) If $Z$ is the scheme of intersection of $\sigma_x$ and $\sigma_y$ then $Z \cong \text{Spec}(\hat{\mathcal{O}}_{k,v}/m^r(\sigma_x, \sigma_y))$ and thus $d_v(x,y) = (\#\kappa(v))^{-\text{ord}(\sigma_x, \sigma_y)}$.

Similarly let $\sigma'_x$ and $\sigma'_y$ be the corresponding sections of $\mathfrak{X}'_v$, and define $\text{ord}'(\sigma'_x, \sigma'_y)$ as above, so that $d'_v(x,y) = (\#\kappa(v))^{-\text{ord}'(\sigma'_x, \sigma'_y)}$. The assertion in (2.2.1) for points in $X(k)$ is then equivalent to the assertion that

\begin{equation}
(2.2.2) \quad \text{ord}'(\sigma'_x, \sigma'_y) \leq \text{ord}(\sigma_x, \sigma_y) \leq \text{ord}'(\sigma'_x, \sigma'_y) + \ell \quad \text{for all } x,y \in X(k).
\end{equation}

Since $\sigma_x = \pi \circ \sigma'_x$ and $\sigma_y = \pi \circ \sigma'_y$, it is clear that if $\sigma'_x$ and $\sigma'_y$ agree to order $r$ then $\sigma_x$ and $\sigma_y$ must agree to order $\geq r$, and this is precisely the first inequality in (2.2.2). To see the second inequality, restrict the ideal sheaf $\mathcal{I}_\pi$ to $\sigma_x$. The restriction will be a principal ideal (since $\hat{\mathcal{O}}_{k,v}$ is a DVR and $\sigma_x$ a section), and so will be some multiple $r$ of the maximal ideal. The condition that $\varphi^*m^\ell \subseteq \mathcal{I}_\pi$ guarantees that $r \leq \ell$. If $\text{ord}(\sigma_x, \sigma_y) > r$ then the restriction of $\mathcal{I}_\pi$ to $\sigma_y$ is also the $r^{th}$ power of the maximal ideal. The sections $\sigma'_x$ and $\sigma'_y$ are the proper transforms of $\sigma_x$ and $\sigma_y$ under the blowup, and to compute the proper transforms it is only necessary to understand $\mathcal{I}_\pi$ restricted to each of the sections. A straightforward local computation (omitted here) shows that if $\text{ord}(\sigma_x, \sigma_y) > r$ then $\text{ord}'(\sigma'_x, \sigma'_y) \geq \text{ord}(\sigma_x, \sigma_y) - r$, and thus $\text{ord}(\sigma_x, \sigma_y) \leq \text{ord}'(\sigma'_x, \sigma'_y) + r$. On the other hand, if $\text{ord}(\sigma_x, \sigma_y) \leq r$ then the inequality $\text{ord}(\sigma_x, \sigma_y) \leq \text{ord}'(\sigma'_x, \sigma'_y) + r$ is automatic, since $\text{ord}' \geq 0$. Thus in all cases we have $\text{ord}(\sigma_x, \sigma_y) \leq \text{ord}'(\sigma'_x, \sigma'_y) + r \leq \text{ord}'(\sigma'_x, \sigma'_y) + \ell$, and this proves (2.2.2) and hence (2.2.1) for all $x,y \in X(k)$.

In the general case that $x, y \in X(\overline{k})$, let $F/k$ be any finite extension so that $x$ and $y$ are defined over $F$, and $\mathfrak{X}_F, v$ and $\mathfrak{X}'_{F,v}$ the base changes of $\mathfrak{X}_v$ and $\mathfrak{X}'_v$ respectively to $\text{Spec}(\hat{\mathcal{O}}_{F,v})$. 

As before, \( x \) and \( y \) define sections \( \sigma_x, \sigma_y \) of \( \mathcal{F}_F_v \) and \( \sigma'_x, \sigma'_y \) of \( \mathcal{F}'_{F,v} \). Let \( n \) be the maximal ideal of \( \mathcal{O}_{F,v} \) and as above define functions \( \text{ord} \) and \( \text{ord}' \) measuring the order of contact between the sections, i.e., the maximum \( r \geq 0 \) so that the sections agree over \( \text{Spec}(\mathcal{O}_{F,v}/n^r) \). Set \( f_F \) to be the degree of \( \mathcal{O}_{F,v}/n \) over \( \kappa(v) \) so that \( \#(\mathcal{O}_{F,v}/n) = \#\kappa(v)^{f_F} \). If \( Z \) is the scheme of intersection between \( \sigma_x \) and \( \sigma_y \) then \( \#\Gamma(Z, \mathcal{O}_Z) = \#\kappa(v)^{f_F \text{ord}(\sigma_x, \sigma_y)} \) and similarly for the scheme of intersection between \( \sigma'_x \) and \( \sigma'_y \). Let \( e_F \) be the ramification degree of \( F_v \) over \( k_v \). Since \( [F_v:k_v] = e_F f_F \) we conclude that

\[
d_v(x,y) = (\#\kappa(v)^{-f_F \text{ord}(\sigma_x, \sigma_y)})^{\frac{1}{\text{ord}(\sigma_x, \sigma_y)}} = \#\kappa(v)^{-\frac{\text{ord}(\sigma_x, \sigma_y)}{e_F}},
\]

with a similar formula holding for \( d'_v \). Thus (2.2.1) amounts to the inequalities

(2.2.3) \[
\text{ord}'(\sigma'_x, \sigma'_y) \leq \text{ord}(\sigma_x, \sigma_y) \leq \text{ord}'(\sigma'_x, \sigma'_y) + \ell e_F.
\]

Let \( g_1: \mathfrak{X}_{F,v} \rightarrow \mathfrak{X}_v \) and \( g_0: \text{Spec}(\mathcal{O}_{F,v}) \rightarrow \text{Spec}(\mathcal{O}_{k,v}) \) be the base change maps, and set \( \varphi_F: \mathfrak{X}_{F,v} \rightarrow \text{Spec}(\mathcal{O}_{F,v}) \) to be the structural map. By combining the facts: (1) \( g_0^* \mathfrak{m} = \mathfrak{n}^{e_F} \), (2) \( \varphi \circ g_1 = g_0 \circ \varphi_F \), and (3) \( \varphi^* \mathfrak{n}^{e_F} \subseteq \mathfrak{I}_x \) we conclude that \( \varphi_F^* \mathfrak{n}^{e_F} \subseteq g_1^* \mathfrak{I}_x \). Since \( \mathfrak{X}_{F,v} \rightarrow \mathfrak{X}_v \) is the blow up along the sheaf of ideals \( g_1^* \mathfrak{I}_x \), we now argue as in the case \( x, y \in X(k) \) to conclude that (2.2.3) holds. Thus (2.2.1) holds for all \( x, y \in X(k) \), proving the proposition. \( \square \)

**Corollary 2.3.** Let \( d_v \) and \( d'_v \) be two distance functions coming from different integral models of \( X \), and fix \( x \in X(k) \). Then \( d_v(x, \cdot) \) is equivalent to \( d'_v(x, \cdot) \).

**Proof:** We again reduce to the case that we have a birational morphism \( \mathfrak{X}' \rightarrow \mathfrak{X} \) between the integral models which is the identity on the generic fibre. As in the proof of Proposition 2.2 this implies that \( d_v(x, y) \leq d'_v(x, y) \) for all \( y \in X(k) \). Let \( X_1, \ldots, X_r \) be the components of \( X \) containing \( x \), and \( X_{r+1}, \ldots, X_s \) the remaining components. By applying Proposition 2.2 to each of the components \( X_1, \ldots, X_r \) we see that there is a constant \( c \) so that \( c d'_v(x, y) \leq d_v(x, y) \) for all \( y \in X_i(k), i = 1, \ldots, r \).

We are therefore reduced to finding a constant which also works for all \( y \in X_i(k), i = r + 1, \ldots, s \). Fix one such \( i \). The value of \( d'_v(x, y) \) is at most 1, and so it suffices to show that \( d_v(x, y) \) is bounded from below for \( y \in X_i(k) \), and this is elementary: \( y \) cannot be too close to \( x \) without also being on one of the components containing \( x \). In more detail, let \( K \) be the field of definition of \( x \) and let \( \mathfrak{X}_i \) be the component of \( \mathfrak{X}_K \) corresponding to \( X_i \). Since \( x \) is not contained in \( X_i \), the intersection \( Z' = \sigma_x \cap \mathfrak{X}_i \) (where \( \sigma_x \) is the section corresponding to \( x \)) is a finite length scheme supported over the closed fibre. We now claim that \( d_v(x, y) \geq 1/\#(Z', \mathcal{O}_{Z'})^{\frac{1}{\text{ord}(\sigma_x, \sigma_y)}} \) for all \( y \in X_i(k) \). If \( y \in X_i(K) \) this is obvious, since the intersection \( \sigma_x \cap \mathfrak{X}_i \) is contained in \( Z' \). If \( y \notin X_i(K) \), then the same argument works after base changing to a larger field \( F \) where both \( x \) and \( y \) are defined, and taking into account the normalization used in defining \( d_v \) (and of course, how the intersection \( \sigma_x \cap \mathfrak{X}_i \) behaves under base change). \( \square \)

**Remark.** In the non-archimedean case, one may also define \( d_v \) by using a projective embedding \( X \hookrightarrow \mathbb{P}^m \) defined over \( k \). If \( x, y \in X(k) \), consider the corresponding projective coordinates \( \mathbf{x} = [x_0 : \ldots : x_m] \), \( \mathbf{y} = [y_0 : \ldots : y_m] \), and set \( d_v(x, y) = \frac{H_v(x_0 y_0)}{H_v(x_0) H_v(y_0)} \) where \( H_v \) is the local height at the place \( v \) (this is the definition given in [1, 2.8.16]). A short local calculation shows that this distance function is equal to the one defined above if one takes the integral model \( \mathfrak{X} \) to be the closure of \( X \) in \( \mathbb{P}^m_{\mathcal{O}_h} \).
Lemma 2.4. Let $x$ be a point of $X(\overline{k})$ and $K$ the field of definition of $x$. Let $U$ be a sufficiently small open affine of $X_K := X \times_k K$ or $X_{\overline{k}} := X \times_{\overline{k}} k$ containing $x$. Let $u_1, \ldots, u_r$ be elements of $\Gamma(U, \mathcal{O}_{X_k})$ (or $\Gamma(U, \mathcal{O}_{X_{\overline{k}}})$) which generate the maximal ideal of $x$. Then there are constants $c$ and $C$ such that

$$c d_v(x, y) \leq \min \left(1, \max \{|u_1(y)|_v, \ldots, |u_r(y)|_v\}\right) \leq C d_v(x, y)$$

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

Proof: In the archimedean case the lemma follows from the local equivalence of $d_v(\cdot, \cdot)$ with the appropriate power of the Euclidean distance function, as used in the proof of Proposition 2.1. To prove the non-archimedean case we start with a few reductions. First, we will only show the lemma for a particular choice of open set and generators of the maximal ideal, and leave it to the reader to check that shrinking the open set and taking a different choice of generators only change the function by a bounded amount. Second, since the norm $|\cdot|_v$ and the distance function $d_v(\cdot, \cdot)$ transform the same way under field extensions, we may assume that $x$ is defined over $k$. Finally, by Corollary 2.3 we may choose whichever integral model we wish when performing the calculation.

Given these reductions, choose an embedding $X \hookrightarrow \mathbb{P}^m$ so that $x$ is sent to $[1:0:\cdots:0]$ and let $\mathfrak{x}$ be the integral model obtained by taking the closure of $X$ in $\mathbb{P}^m_{O_k}$. Let $Z_0, \ldots, Z_m$ be homogenous coordinates on $\mathbb{P}^m$, and choose the open set $U$ of $X$ to be the set $Z_0 \neq 0$, and $u_i = Z_i/Z_0$ for $i = 1, \ldots, m$ as the generators of the maximal ideal of $x$. Using the fact that that $x$ is sent to $[1:0:\cdots:0]$ and the remark before this lemma, for this integral model

$$d_v(x, y) = \frac{\max(|Z_1(y)|_v,|Z_2(y)|_v,\ldots,|Z_m(y)|_v)}{\max(|Z_0(y)|_v,|Z_1(y)|_v,\ldots,|Z_m(y)|_v)}$$

for all $y \in X(\overline{k})$.

For $y \in U(\overline{k})$, this is equal to $\min(1, \max \{|u_1(y)|_v, \ldots, |u_m(y)|_v\})$. □

Remark. Let $x$ be a point of $X(\overline{k})$ 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 $K \not\equiv k_v$, 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 $K_v = k_v$.

Remark. From the definition we have the bound $d_v(x, y) \leq 1$ for all $x, y \in X(\overline{k})$ in the non-archimedean case, and $d_v(x, y) \leq (\frac{γ}{2})^2$ in the archimedean case. In the arguments below we will simply use the bound $(\frac{γ}{2})^2$, valid in all cases, in order to avoid distinguishing the particular case.

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

$$c H(x) < H'(x) < C H(x)$$

for all $x \in X(\overline{k})$ (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 $X$ defined over $k$ we may associate a height function $H_L$, well defined up to equivalence, in such a way that the map from $\text{Pic}(X)_k$ 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], [7, Part B], [9, Chap. III], or [16, 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 \( x = [x_0 : \cdots : x_n] \in \mathbb{P}^n(k) \), the height with respect to \( \mathcal{O}_{\mathbb{P}^n}(1) \) is

\[
H(x) = \prod_v \max(|x_0|_v, \ldots, |x_n|_v)
\]

where the product ranges over all the places \( v \) of \( k \).

Unless otherwise specified all height functions in this paper are multiplicative, relative to \( k \), and come from line bundles on \( X \) defined over \( k \).

**Approximation Constants.** We now define the main objects of study in this paper, inspired by similar definitions from [12]. We fix a single place \( v \), archimedean or non-archimedean, and a corresponding distance function \( d_v \) as described above.

**Definition 2.5.** Let \( X \) be a projective variety, \( x \in X(\mathbb{C}) \), \( L \) a line bundle on \( X \). For any sequence \( \{x_i\} \subset X(k) \) of distinct points with \( d_v(x,x_i) \to 0 \) (which we denote by \( \{x_i\} \to x \)), we set

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

**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 \( A(\{x'_i\}, L) \subset A(\{x_i\}, L) \).

**Definition 2.6.** If \( A(\{x_i\}, L) \) is empty we set \( \alpha_x(\{x_i\}, L) = \infty \). Otherwise we set \( \alpha_x(\{x_i\}, L) \) to be the infimum of \( A(\{x_i\}, L) \). 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 \( \delta > 0 \), \( d_v(x,x_i)^{\alpha_x(\{x_i\}, L) + \delta} H_L(x_i) \to 0 \) as \( i \to \infty \). 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 2.7.** 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 \( \alpha_x \) 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 \( \alpha_x \). In particular, by Proposition 2.1 and Corollary 2.3 the definition of \( \alpha_x \) does not depend on the choice of distance function \( d_v \), i.e., the choice of projective embedding or integral model in the archimedean and non-archimedean cases respectively).

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

\[
c H^{1-\delta} \leq H' \leq C H^{1+\delta}.
\]

The definitions of \( \alpha_x(\{x_i\}, L) \) and \( \alpha_x(L) \) only depend on the quasi-equivalence class of the height function. All heights from line bundles in \( \text{Pic}^0(X) \) are quasi-equivalent (see [16, p. 26];
the proof also applies to singular varieties). 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 2.12(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 \( \alpha_x(L) > 0 \), 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 \( x \in X(k) \). 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 [12], since it is the infimum of the set described above rather than the minimum, as in [12]. In [12] 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 [12] 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 [12] 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 \( P \in X(k) \) such that \( H_L(P) \leq c \) is finite. Note in particular that every ample line bundle has the Northcott property.

**Definition 2.8.** For any point \( x \in \overline{X(k)} \) and any line bundle \( L \) we set

\[
B_x(L) = \left\{ \gamma \in \mathbb{R}_{\geq 0} \middle| \text{for all } C > 0 \text{ the number of } x_i \in X(k) \text{ such that } d_v(x, x_i)^\gamma H_L(x_i) < C \text{ is finite.} \right\}
\]

**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 \( d_v(x, x_i) \leq \left( \frac{\pi}{2} \right)^{2\gamma} \), for any \( \gamma > 0 \) these infinitely many \( x_i \) also satisfy

\[
d_v(x, x_i)^\gamma H_L(x_i) \leq \left( \frac{\pi}{2} \right)^{2\gamma} H_L(x_i) < \left( \frac{\pi}{2} \right)^{2\gamma} C
\]

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

**Proposition 2.9.** Suppose that \( L \) has the Northcott property. Then \( \alpha_{\alpha}(L) = \sup(B_{\alpha}(L)) \).

**Proof:** Set \( \alpha_x = \alpha_x(L) \) and \( b_x = \sup(B_{\alpha}(L)) \). By definition of \( \alpha_x \), 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 \( \alpha_x + \delta \notin B_{\alpha}(L) \) and so \( \alpha_x + \delta \geq b_x \). Letting \( \delta \) 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 \( d_v(x, x_i)^{b_x+\delta} H_L(x_k) < C \). 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 \( H_L(x_i) \to \infty \) as \( i \to \infty \). By the boundedness of the product, we conclude that \( d_v(x, x_i) \to 0 \), and so \( \{x_i\} \) converges to \( x \). But then

\[
d_v(x, x_i)^{b_x+2\delta} H_L(x_i) < C \cdot d_v(x, x_i)^{\delta} \to 0
\]

and so \( b_x + 2\delta \in A(\{x_i\}, L) \). Thus \( b_x + 2\delta \geq \alpha_x \), and letting \( \delta \) 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 \( 0 \in B_{\alpha}(L) \) and hence \( \alpha_x(L) = \sup(B_{\alpha}(L)) \geq 0 \) by Proposition 2.9. 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 \( K \). We use this means that when computing \( \alpha_x \), we use the height \( H_L \) relative to \( K \) and normalize \( d_v \) relative to \( K \) (e.g., in the non-archimedean case we compute \( d_v(x, y) \) by passing to an extension \( F/K \) where both \( x \) and \( y \) are defined and take the \( \left[ F_v:K_v \right]^{\text{th}} \) root of the reciprocal of the number of elements in the ring of the intersection scheme). 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 2.10.** Suppose \( x \in X(\overline{k}) \), \( L \) a line bundle defined over \( k \), and \( \{x_i\} \to x \) a sequence of points in \( X(k) \) approximating \( x \). Let \( K \) be any finite extension of \( k \). Then \( \{x_i\} \to x \) 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 \( \alpha \).** We start by computing \( \alpha \) when \( x \in \mathbb{P}^n(k) \).
Lemma 2.11. 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 $x = [1:0: \ldots :0]$. 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 $Z_0, \ldots , Z_n$ be the coordinates on $\mathbb{P}^n$ and $\{x_i\}$ a sequence of k-points converging to $x$. Since $d_v(x, x_i) \to 0$ as $i \to \infty$ we conclude that $|Z_j(x_i)/Z_0(x_i)|_v \to 0$ for each $j = 1, \ldots , n$. By passing to a subsequence of the $x_i$, which can only possibly lower the value of $\alpha$, 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 $j \in \{1, \ldots , n\}$ so that $\max(|Z_1(x_i)|_v, \ldots , |Z_n(x_i)|_v) = |Z_j(x_i)|_v$. By Lemma 2.4 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 $\gamma \geq 0$

$$d_v(x, x_i)^\gamma H(x_i) = \left(\frac{|Z_j(x_i)|_v}{|Z_0(x_i)|_v}\right)^\gamma |Z_0(x_i)|_v \cdot \prod_{w \not= v} \max(|Z_0(x_i)|_w, \ldots , |Z_n(x_i)|_w),$$

where in the last step we have used the product formula. If $\gamma < 1$ then the lower bound above goes to infinity as $i \to \infty$, 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 $\prod_{w \text{ arch}} k_w = \mathbb{R}^r \times \mathbb{C}^s$. Let $D$ be the diameter of a fundamental domain of $\mathcal{O}_k$. Then there are infinitely many elements of $\mathcal{O}_k$ that lie in the cylinder $\{b \in \mathcal{O}_k \mid |b|_w \leq D \text{ for } w \not= v\}$. These elements $b_i$ satisfy $H([b_i: 1])d_v([1: 0], [b_i: 1]) \leq D^{s+2s}$, 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 multiplying by a 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 $H(x_i) = \prod_{w \text{ arch}} |b^i|_w = 1/|b|_v$, where the last equality follows from the product formula. Since $d_v(x, x_i) = |b|_v$, 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 $\alpha$.

Proposition 2.12. Let $X$ and $Y$ be projective varieties over $k$, $x \in X(\overline{k})$, and $L$ a line bundle on $X$.

(a) For any positive integer $m$, $\alpha_{x,X}(m \cdot L) = m \cdot \alpha_{x,X}(L)$.

(b) $\alpha_x$ 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 $X$ defined over $k$ then for any point $z \in Z(\overline{k})$ 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 $x \in X(\overline{k})$ 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 \( x \in X(k) \), \( y \in Y(k) \). Either
\[
\alpha_{x,y,xy}(L_X \boxplus L_Y) \geq \alpha_x(L_X) + \alpha_y(L_Y)
\]
or else
\[
\alpha_{x,y,xy}(L_X \boxplus L_Y) = \min(\alpha_{x,X}(L_X), \alpha_{y,Y}(L_Y)).
\]
If neither \( x \) nor \( y \) are defined over \( k \), then the first case must hold. If \( x \) and \( y \) are both defined over \( k \), then the second case must hold.

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

**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 \( \alpha_1 \) and \( \alpha_2 \) are finite. By the homogeneity of \( \alpha_x(\cdot) \) (i.e., part (a)) it is sufficient to prove (b) under the additional assumption 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 \( a\alpha_1 + b\alpha_2 - \delta > \alpha_x(\{x_i\}, aL_1 + bL_2) \). Taking the infimum over all sequences \( \{x_i\} \) of \( k \)-points we have \( \alpha_x(aL_1 + bL_2) \geq a\alpha_1 + b\alpha_2 \), which is the inequality in (b).

When one or both of \( \alpha_1 \) and \( \alpha_2 \) 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 \( \infty \geq \alpha_x(\{x_i\}, aL_1 + bL_2) \) 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 \( Z(k) \). 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 \( Z \) are 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 2.11 we conclude that \( \alpha_{x,X}(L) \geq \alpha_{x,\mathbb{P}^n}(O_{\mathbb{P}^n}(1)) = 1 \). If \( L \) is ample then some multiple \( mL \) is very ample, and so if \( x \in X(k) \) then \( \alpha_x(L) > \frac{1}{m} \) by the first part of this statement and homogeneity. Finally, if \( x \in X(k) \) let \( K \) be the field of definition of \( x \). We have just established that \( \alpha_x(L)_K > 0 \), hence by Proposition 2.10 we have \( \alpha_x(L)_k = \alpha_x(L)_K > \frac{m}{d} \alpha_x(L)_K > 0 \).
To prove claim (e), notice that the height function with respect to \( L_X \oplus L_Y \) is the product of the height functions of \( L_X \) and \( L_Y \). 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 \( \alpha_x \) and \( \alpha_y \), we must have

\[
\alpha_{x,y,x+y}(\{(x_i, y_i)\}, L_X \oplus L_Y) \geq \alpha_x(L_X) + \alpha_y(L_Y)
\]

as desired. Otherwise, one of \( \{x_i\} \) or \( \{y_i\} \) must be eventually constant, in which case we have

\[
\alpha_{x,y,x+y}(L_X \oplus L_Y) = \min(\alpha_{x,X}(L_X), \alpha_{y,Y}(L_Y)).
\]

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 that 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 \( A \). This follows immediately from the concavity in Proposition 2.12(b) and the facts that: (1) For any line bundle \( L \) and any ample line bundle \( A \), for large enough \( m \) the line bundle \( A_m = L + mA \) is ample. Thus, under the hypothesis on \( L \) and by concavity, there is an ample line bundle \( A_m \) such that \( \alpha_x(A_m) = \infty \). (2) The ample cone is an open convex cone. Thus any ample line bundle is a positive convex combination of \( A_m \) and some other (ample) line bundle. Note however that if \( X \) is uniruled over \( k \), then \( \alpha_x(L) \) is finite for any ample \( L \).

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

(c) If \( X \) is smooth and has a nontrivial Jacobian, any sequence \( \{x_i\} \) such that \( \alpha_x(\{x_i\}, L) \) is finite must lie in a fibre of the Albanese map. This follows from the fact that \( \alpha \) is infinite on Abelian varieties (see Example (c) in the introduction), the concavity of \( \alpha \), and the fact that distance can only decrease under the Albanese map (or any map).

**Lemma 2.13.** Let \( L = O_{\mathbb{P}^1}(d) \) and \( x \in \mathbb{P}^1(k) \). Then

\[
\alpha_x(L) = \begin{cases} 
\infty & \text{if } \kappa(x) \notin k_v, \\
\frac{d}{2} & \text{if } \kappa(x) = k \\
\frac{d}{2} & \text{otherwise.}
\end{cases}
\]

**Proof:** If \( \kappa(x) \notin k_v \), then there is no sequence of \( k \)-points converging (with respect to \( d_v(\cdot, \cdot) \)) to \( x \) (see the Remark on page 12), and hence \( \alpha_x(L) = \infty \). If \( x \in \mathbb{P}^1(k) \) then this is Lemma 2.11 and Proposition 2.12(a). If \( \kappa(x) \in k_v \) but \( \kappa(x) \neq k \) then \( \alpha_x(O_{\mathbb{P}^1}(1)) \geq \frac{1}{2} \) by Roth’s theorem for \( \mathbb{P}^1 \), while \( \alpha_x(O_{\mathbb{P}^1}(1)) \leq \frac{1}{2} \) by a Dirichlet-type argument. (This follows, for example, from Theorem 7.8.) Thus \( \alpha_x(O_{\mathbb{P}^1}(1)) = \frac{1}{2} \), and so \( \alpha_x(L) = \frac{d}{2} \) by Proposition 2.12(a) again. \( \square \)
Theorem 2.14. Let $C$ be any singular $k$-rational curve and $\varphi : \mathbb{P}^1 \to C$ the normalization map. Then for any ample line bundle $L$ on $C$, and any $x \in C(k)$ 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

$$r_q = \begin{cases} 
0 & \text{if } \kappa(q) \notin k_v \\
1 & \text{if } \kappa(q) = k \\
2 & \text{otherwise.}
\end{cases}$$

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

Proof: Given any sequence $\{x_i\} \to x$ 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 $\varphi$ 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 $\{q_i\} \to q$ 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 $\varphi$ is described by functions in the $m_q$-th power of the maximal ideal of $q$, and thus $d_v(x_i, \varphi(q_i))$ is equivalent to $d_v(q, q_i)^m_q$ as $i \to \infty$. Therefore, as in Proposition 2.10 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 2.13. \qed

Remark: This is similar to Theorem 2.8 of [12], except that it is actually correct. (The conclusion of Theorem 2.8 of [12] neglects the possibility that the $r_q$ defined in Theorem 2.14 might not be one.) Theorem 2.14 also uses the definition of $\alpha$ from this paper, rather than that of [12], and generalises the results to points defined over $k$.

Examples

(a) If $X = \mathbb{P}^N$, $L = \mathcal{O}_{\mathbb{P}^N}(d)$ then $\alpha_x(L) = d$ for all points $x$ in $\mathbb{P}^N(k)$. This follows from Lemma 2.11 and Proposition 2.12(a).

(b) If $X \subseteq \mathbb{P}^N$, $L = \mathcal{O}_{\mathbb{P}^N}(1)|_X$ and $Z \subseteq X$ is a linear space (i.e., of degree 1 in $\mathbb{P}^N$) defined over $k$, then for any point $z \in Z(k)$ we have $\alpha_{z,X}(L) = 1$, since we have $\alpha_{z,Z} = 1$ from part (a) and then applying Proposition 2.12 (c) and (d) gives $1 \geq \alpha_{z,X} \geq 1$.

(c) If $X = \mathbb{P}^1 \times \mathbb{P}^1$, $L = \mathcal{O}_{\mathbb{P}^1 \times \mathbb{P}^1}(a,b)$, with $a, b > 0$ then $\alpha_x = \min(a,b)$ for all $x \in X(k)$. This follows immediately from Proposition 2.12(e).

(d) Similarly if $X = \mathbb{P}^N_1 \times \cdots \times \mathbb{P}^N_r$, $L = \mathcal{O}_X(d_1, \ldots, d_r)$ with $d_i > 0$ then $\alpha_x(L) = \min(d_1, \ldots, d_r)$.

(e) 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 (c) 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 (e) 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 2.12(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 $\alpha$.

**Lemma 2.15.** Let $Z$ be variety defined over $k$, and set $Y$ to be the Zariski closure (over $k$) of the points of $Z(k)$. Then $Y$ is defined over $k$, each irreducible component of $Y$ is geometrically irreducible, and for any line bundle $L$ and any $x \in Z(k)$ we have $\alpha_{x,Z}(L) = \alpha_{x,Y}(L|_Y)$.

**Proof:** From the construction it is clear that $Y$ is defined over $k$. Let $Y_1, \ldots, Y_r$ be the irreducible components of $Y$ over $\overline{k}$; we will show that each $Y_i$ is actually defined over $k$. Let $Y_i$ be one such component. Since $Y$ is defined over $k$, all $\text{Gal}(\overline{k}/k)$ conjugates of $Y_i$ are also components of $Y$. Let $I \subseteq \{1, \ldots, r\}$ be the subset of indices such that each $Y_j$, $j \in I$, is a Galois conjugate of $Y_i$, and set $I' = \{1, \ldots, r\} \setminus I$. Any point $y \in Z(k)$ contained in $Y_i$ is also contained in $Y_j$ for $j \in I$. Therefore all points of $Z(k)$ contained in $Y' := (\cap_{j \in I} Y_j) \cup (\cup_{j' \in I'} Y_{j'})$. By construction $Y'$ is closed and defined over $k$. If $I \neq \{i\}$ then $Y'$ is a proper subset of $Y$. 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 $x \in Z(\overline{k})$. □

### 3. Seshadri Constants

In this section, we review some basic properties of Seshadri constants, first introduced and studied in [4]. Many foundational results on Seshadri constants are given in [10, 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 defined over $k$, $x$ a point of $X(\overline{k})$, 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 $\pi: \overline{X} \to X_{\overline{k}}$ is the blowup of $X_{\overline{k}} := X \times_k \overline{k}$ at $x$ with exceptional divisor $E$. Here, by abuse of notation, we also use $L$ for the base change of $L$ to $X_{\overline{k}}$.

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 $\gamma$ 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 $\epsilon_x$ for $\epsilon_{x,X}(L)$ if $X$ or $L$ are clear from the context.

Since the Seshadri constant only depends on $X_{\overline{k}}$, 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 defined over $k$.

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 $\epsilon$. We start by computing $\epsilon$ 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 $\pi: \mathbb{P}^n \rightarrow \mathbb{P}^n$ 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 $\mathbb{P}^n \rightarrow \mathbb{P}^{n-1}$ with fibres the proper transforms of lines in $\mathbb{P}^n$ passing through $x$. Thus $L_1$ is nef on $\mathbb{P}^n$. For any such fibre the degree of $L_\gamma$ 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 $\epsilon$ and $\alpha$ much further.

Proposition 3.4. Let $X$ be a projective variety, $x \in X(\mathbb{C})$, and $L$ a nef $\mathbb{Q}$-divisor on $X$.

(a) For any positive integer $m$, $\epsilon_{x,X}(m \cdot L) = m \cdot \epsilon_{x,X}(L)$.

(b) $\epsilon_x$ is a concave function of $L$: for any positive rational numbers $a$ and $b$, and any $\mathbb{Q}$-divisors $L_1$ and $L_2$

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

(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 $\epsilon_x(L) > 0$, if $L$ is ample then $\epsilon_x(L) > 0$.

(e) If $x$ and $y$ are points of varieties $X$ and $Y$, with nef line bundles $L_X$ and $L_Y$ then

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

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

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 $m$ 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 gives

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

To prove the other direction, we will use the description of $\epsilon_{x,y}$ from Proposition 3.2. Let $\pi_X$ and $\pi_Y$ 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 $\pi_X$. Then $\pi_X(C)$ is not equal to a point, and if $d$ is the generic degree of the map $C \to \pi_X(C)$ we have $\pi^* L_X \cdot C = d(L_X \cdot \pi_X(C))$, and $\text{mult}_{x,y}(C) \leq d \cdot \text{mult}_x(\pi_X(C))$.

Since $\epsilon_x$ is the Seshadri constant for $L_X$ at $x$, we have

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

where the first inequality follows from Proposition 3.2 applied to $\epsilon_x$, 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 $\pi_Y$ we have the corresponding inequality with $\epsilon_y$ in place of $\epsilon_x$. Since for any given curve $C$ one of these must be true we have

$$\min(\epsilon_x, \epsilon_y) \leq \inf_{x \times y \in X \times Y} \left\{ \frac{(\pi_X^* L_X + \pi_Y^* L_Y) \cdot C}{\text{mult}_{x,y}(C)} \right\} \frac{32}{\epsilon_{x,y}}$$

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 $\tilde{X}$. □

**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 $X \subseteq \mathbb{P}^n$, and $Z \subseteq \mathbb{P}^n$ is a linear space (i.e., of degree 1 in $\mathbb{P}^n$) then for any point $z \in Z$ we have $\epsilon_{x,Z}(L) = 1$, since as before we have $\epsilon_{z,x} = 1$ from part (a) of the example and applying Proposition 3.4(c) and (d) gives $1 \geq \epsilon_{z,x} > 1$.

(c) If $X = \mathbb{P}^1 \times \mathbb{P}^1$, $L = \mathcal{O}_{\mathbb{P}^1 \times \mathbb{P}^1}(a,b)$, with $a, b > 0$ 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., $\epsilon_x(aF_1 + bF_2) = \min(a, b)$.

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

As evidenced by our parallel statements in Proposition 2.12 and Proposition 3.4 (and Lemmas 2.11 and 3.3, and the examples) there is a great deal of formal similarity between \( \alpha_x \) and \( \epsilon_x \). See the discussion below on the Arakelov point of view for one heuristic reason for this similarity.

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

**Lemma 3.5.** Let \( C \) be an irreducible curve defined over \( k \), \( x \in C(k) \) and \( L \) and 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 \( \kappa(x) \subseteq k_x \).

**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 2.14. By that theorem we have \( \alpha_x(L) = \min_{q \in \varphi^{-1}(x)} \{ \frac{d}{r_q m_q} \} \) where \( d = \deg(L) \). By the definition of the Seshadri constant we have \( \epsilon_x(L) = \frac{d}{\mult_x C} = \frac{d}{\sum_{q \in \varphi^{-1}(x)} m_q} \). 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 2.14 this proves the lemma. \( \square \)

**Arakelov point of view.** Let \( X \) be a projective variety defined over \( k \) and \( x \) a point of \( X(k) \). By Proposition 2.9 for any ample line bundle \( L \) an equivalent description of \( \alpha_x = \alpha_x(L) \) is that for any \( \gamma < \alpha_x \) the set

\[
(3.5.1) \quad \left\{ x_i \in X(k) \mid d_v(x, x_i)^{\gamma} h_L(x_i) < 1 \right\}
\]

is finite, and \( \alpha_x \) is the largest number with this property. (The extra quantifier “\( C \)” in Definition 2.8 can be absorbed by the condition that the finiteness is supposed to hold for all \( \gamma < \alpha_x \). The only reason this (unnecessary) extra quantifier is used in Definition 2.8 is to simplify arguments.)

Taking log, the finiteness of (3.5.1) is equivalent to the finiteness of

\[
(3.5.2) \quad \left\{ x_i \in X(k) \mid h_L(x_i) + \gamma \log(d_v(x, x_i)) < 0 \right\}
\]

where \( h_L \) is the logarithmic height.

Let \( \tilde{X} \) be the blow up of \( X \) at \( x \) with exceptional divisor \( E \). The definition of the Seshadri constant \( \epsilon_x = \epsilon_x(L) \) is that for any \( \gamma < \epsilon_x \) the set

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

is empty, and \( \epsilon_x \) is the largest number with this property.

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

$$h_{L-\gamma E}(x_i) = (L - \gamma E) \cdot \sigma_i.$$ 

Assume that $x \in X(k)$ and choose an embedding $\varphi : X \to \mathbb{P}^n$ so that $x \mapsto [1:0: \cdots : 0]$. Let $Z_0, \ldots, Z_n$ be the coordinates on $\mathbb{P}^n$ and define functions $u_i$, $i = 1, \ldots, n$ on the open subset where $Z_0 \neq 0$ by $u_i = Z_i / Z_0$. For each place $w$ of $k$, define a function $e_w : X(k) \to \mathbb{R}_{>0}$ by

$$e_w(y) = \begin{cases} 1 & \text{if } Z_0(y) = 0, \\ \min(1, \max(|u_1(y)|_w, \ldots, |u_n(y)|_w)) & \text{if } Z_0(y) \neq 0. \end{cases}$$

A short local calculation shows that $-h_E(x_i) = \sum_w \log(e_w(x_i))$ and thus

$$(L - \gamma E) \cdot \sigma_i = h_{L-\gamma E}(x_i) = h_L(x_i) + \gamma \log(e_v(x_i)) + \gamma \left( \sum_{w \not\in v} \log(e_w(x_i)) \right).$$

From the definition $\log(e_w(x_i)) \leq 0$ for each $w$ and therefore

$$(3.5.3) \quad h_L(x_i) + \gamma \log(e_v(x_i)) < 0 \implies (L - \gamma E) \cdot \sigma_i < 0.$$ 

By Lemma 2.4, $\log(e_v(x_i))$ is equivalent to $\log(d_i(x, x_i))$ as $i \to \infty$. Treating $\log(e_v(x_i))$ as $\log(d_i(x, x_i))$, and ignoring for the moment that the implication in (3.5.3) is only one way, (3.5.2) shows that the definition for $\alpha_x$ is roughly that for each $\gamma < \alpha_x$

$$\left\{ \sigma_i \not\in \tilde{X} \mid \sigma_i \text{ an irreducible curve of degree one over } \text{Spec}(\mathcal{O}_k), \ (L - \gamma E) \cdot \sigma_i < 0 \right\}$$

is a finite set, and that $\alpha_x$ is the largest number with this property. 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 $\alpha_x$ look very much like the definition of $\epsilon_x$.

4. THE CONSTANT $\beta_x(L)$

The proof of the general versions 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 prove some basic facts to be used in the proof of Roth’s theorem. 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 $\pi : \tilde{X} \to X$ be the blow up at $x$ with exceptional divisor $E$, and for any $\gamma \in \mathbb{R}_{>0}$ 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 $h^0(mM) = \frac{\text{Vol}(M)}{m^n} m^n + O(m^{n-1})$ 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 [10, Corollary 2.2.45] $\text{Vol}$ depends only on the numerical class of $M$ and extends uniquely to a continuous function on the real Néron-Severi group. Let $\gamma_{\text{eff}} = \gamma_{\text{eff},x}(L) = \sup\{\gamma \in \mathbb{R}_{>0} \mid L_\gamma \text{ is effective}\}$. A line bundle $M$ is called big if $\text{Vol}(M) \neq 0$. Since
the big cone is the interior of the effective cone, $L_{\gamma}$ is big for every $\gamma \in [0, \gamma_{\text{eff}})$. Since $L_{\gamma_{\text{eff}}}$ is on the boundary of the effective cone (and by continuity of $\text{Vol}(\cdot)$), $\text{Vol}(L_{\gamma_{\text{eff}}}) = 0$. For any ample line bundle $L$, define a decreasing function (the “asymptotic relative volume function”) $f : [0, \gamma_{\text{eff}}] \rightarrow [0, 1]$ by

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

Finally, define

$$\beta_x(L) = \int_{\gamma_{\text{eff}},x}^{\gamma_{\text{eff}}} f(\gamma) \, d\gamma = \int_{0}^{\gamma_{\text{eff}}} f(\gamma) \, d\gamma$$

to be the area under $f$.

**Example.** On $\mathbb{P}^n$, and with $L = \mathcal{O}_{\mathbb{P}^n}(1)$, $\text{Vol}(L) = 1$, and for any $x \in \mathbb{P}^n$, $\gamma_{\text{eff},x}(L) = 1$ and $f(\gamma) = 1 - \gamma^n$, with area $\beta_x(L) = \frac{n}{n+1}$. This will turn out (via Theorem 6.1 or 6.2) 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).

**Example.** Let $X = \mathbb{P}^1 \times \mathbb{P}^1$ and $L = \mathcal{O}_X(d_1, d_2)$, then $\gamma_{\text{eff},x}(L) = d_1 + d_2$ for any $x \in X$. For any $\gamma \in [0, d_1 + d_2]$ the ratio $f(\gamma) = \text{Vol}(L_{\gamma})/\text{Vol}(L)$ can be visualized as the fraction of the rectangle $[0, d_1] \times [0, d_2]$ satisfying $u_1 + u_2 \geq \gamma$ (the shaded region shown below):  

\[y = f(\gamma)\]

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

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

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 to the shaded region in Figure 4a and will be explained below.)
Lemma 4.1. For any ample $L$ and $x \in X$ 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 $\gamma$. For $m$ large and sufficiently divisible (i.e., so that $m\gamma$ is an integer) we have the exact sequence of sheaves

\begin{equation}
0 \longrightarrow mL \xrightarrow{m\gamma E} mL_0 \longrightarrow mL_0|m\gamma E \longrightarrow 0
\end{equation}

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:

$$0 \longrightarrow \Gamma(\tilde{X}, mL_0) \longrightarrow \Gamma(\tilde{X}, mL_0) \longrightarrow \Gamma(m\gamma E, L_0|m\gamma E).$$

Since $h^0(mL_0) = h^0(mL) = \frac{\text{Vol}(L)}{\mathcal{I}h} m^n + O(m^{n-1})$ the lemma will follow if we show that $h^0(mL_0|m\gamma E) \leq \sum_{\ell=0}^{m\gamma-1} h^\ell(\mathcal{E}_E(-\ell E))$. 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 $\mathcal{I}_E$ be the ideal sheaf of $E$ on $\tilde{X}$. For any $\ell \geq 1$ we have $\mathcal{I}_E^\ell = \mathcal{O}_E(-\ell E)$, and thus the exact sequence of sheaves

\begin{equation}
0 \longrightarrow \mathcal{O}_E(-\ell E) \longrightarrow \mathcal{O}_{(\ell+1)E} \longrightarrow \mathcal{O}_{\ell E} \longrightarrow 0.
\end{equation}

This gives the inductive estimate

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

Choose an embedding $X \hookrightarrow \mathbb{P}^m$ and let $\mathbb{P}^m$ 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 $\mathbb{P}^m$, and $E = \tilde{X} \cap E'$. Furthermore, $E$ is an $(n-1)$-dimensional subvariety of $E'$ of degree (as a subvariety of $\mathbb{P}^{m-1}$) equal to $\text{mult}_x X$. We thus have

\begin{equation}
h^0(\mathcal{O}_E(-\ell E)) = \frac{\text{mult}_x X}{(n-1)!} \ell^{n-1} + O(\ell^{n-2}) \quad \text{for } \ell > 0
\end{equation}

since $\mathcal{O}_{E'}(-\ell E') = \mathcal{O}_{\mathbb{P}^m}(1)$, and so $h^0(\mathcal{O}_E(-\ell E)) = h^0(\mathcal{O}_{\mathbb{P}^m}(\ell)|E)$ is simply given by the Hilbert polynomial of $E$ for large $\ell$. 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., $H^0(m\ell_0|m\gamma E)$ fails to impose independent conditions on $H^0(m\ell_0)$ for $\gamma$ 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 $\beta_x(L) \geq \frac{n}{n+1} \sqrt{\text{Vol}(L)/\text{mult}_x X} \geq \frac{n}{n+1} \epsilon_x(L)$. In general, both these inequalities are strict.
Proof: Let $g(\gamma) = 1 - \frac{\text{mult}_x X}{\text{Vol}(L)} \gamma^n$. By Lemma 4.1 we have $f(\gamma) \geq g(\gamma)$ on $[0, \gamma_{\text{eff}}]$. Let

$$\omega = \sqrt[n]{\frac{\text{Vol}(L)}{\text{mult}_x X}}$$

(i.e., the solution to $g(\omega) = 0$). 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 [10, Proposition 5.1.9]. In the example of $X = \mathbb{P}^1 \times \mathbb{P}^1$, $L = O_X(d_1, d_2)$ (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 $\beta$ in general.

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

$$\beta_x(L) = \min(\beta_{x,X_1}(L|X_1), \ldots, \beta_{x,X_\ell}(L|X_\ell)),$$

where $X_1, \ldots, X_\ell$ are the irreducible components of $X = X \times_k \overline{k}$ 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 defined over $k$. Then for any $x \in X(k)$ 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 $\overline{X} = X \times_k \overline{k}$ with irreducible components $X_1, \ldots, X_\ell$. Then each component is $n$-dimensional, hence applying Corollary 4.2 we have $\beta_{x,X_i}(L|\overline{X}) \geq \frac{n}{n+1} \epsilon_{x,X_i}(L|\overline{X})$ for each $i = 1, \ldots, \ell$. 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, $x \in X(k)$ be any point and $K$ its field of definition. Set $X_K = X \times_k K$, $\pi_K: \overline{X}_K \to X_K$ 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 $\overline{X}_K$, $E_K$, and $L_{\gamma,K}$. Since $x$ is defined over $K$ it follows that $\overline{X}_K \times_K \overline{k} = \overline{X}_K$ and hence that $\dim_K H^0(\overline{X}_K, mL_{\gamma,K}) = \dim_K H^0(\overline{X}_K, mL_{\gamma,K})$ for all $m > 0$ and $\gamma \geq 0$ with $m\gamma$ an integer. Thus the dimension of $mL_{\gamma,K}$, and hence the asymptotic growth (i.e., the volume) of $L_{\gamma,K}$ may be computed “over $K$”. In particular, $\text{Vol}(L_{\gamma,K})/\text{Vol}(L) = \text{Vol}(L_{\gamma,K})/\text{Vol}(L)$ 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 approximation\(^2\) 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 $\overline{k}$. Let $X$ be an $n$-dimensional variety defined over $k$ and irreducible over $k$. For each $v \in S$ choose a point $x_v \in X(\overline{k})$, and let $d_v(\cdot, \cdot)$ be the distance

\(^2\)as well as Propositions 2.12(f) and 3.4(c), and Corollary 4.4...
function (as 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 \( \alpha \), and we will state the results of this section in terms of \( \alpha \) and in terms of the usual finiteness conditions; both versions are equivalent. In order to concisely state the results in terms of \( \alpha \) it is helpful to make the following definition.

**Definition 5.1.** We say that a sequence \( \{x_i\} \) is essentially contained in a subvariety \( Z \) if all but finitely many of the \( x_i \) are contained in \( Z \).

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:

\[
\begin{align*}
\text{(5.1.1)} \quad & \left\{ \begin{array}{l}
\text{There is a proper subvariety } Z \text{ defined over } k \text{ such that for all infinite sequences } \\
\{x_i\} \text{ of distinct points of } X(k) \text{ not essentially contained in } Z, \text{ there is at least one } v \in S \text{ so that } \\
\alpha_v(\{x_i\}, L) \geq \frac{1}{R_v}.
\end{array} \right. \\
\end{align*}
\]

and

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

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

\[
\sum_{v \in S} \beta_{x_v}(L)R_v > 1
\]

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 equality in (5.2.1) is sufficient. 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 \( Z \) defined over \( k \) 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 \( v \in V \). Since each \( \beta_{x_v}(L) > 0 \), and since each \( R'_v > R_v \), we conclude that
The bundle \( \sum_v \beta_{x_v}(L)R_v' > 1 \), and thus we may apply Theorem 5.2 to the collection \( \{R_v\}_{v \in S} \). By the theorem, there 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.

**Corollary 5.4.** If \( \dim X = 1 \) equality in (5.2.1) is sufficient, and one may take \( Z = \emptyset \).

**Proof:** We will establish the corollary in the form of assertion (5.1.2). Given a collection \( \{R_v\}_{v \in S} \) such that \( \sum_v \beta_{x_v}(L)R_v \geq 1 \) we need to show that for any collection \( \{\delta_v\}_{v \in S} \) with each \( \delta_v > 0 \) there are only finitely many solutions \( y \in X(k) \) to
\[
d_v(x_v, y) \leq H_L(y)^{-(R_v + \delta_v)} \text{ for all } v \in S.
\]

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 \( X \) defined over \( k \) 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
\[
V_{K^{(v)}} = V_{K^{(v)}}^0 \supseteq V_{K^{(v)}}^1 \supseteq \cdots \supseteq V_{K^{(v)}}^{r_v} = V_{K^{(v)}}^{r_v+1} = \{0\}
\]
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 \( W_{K^{(v)}}^j = W_{K^{(v)}}^j \cap W_{K^{(v)}}^{j+1} \) for \( j = 1, \ldots, r_v + 1 \). We define the \( v \)-th piece of the slope, \( \mu_v(W) \) by
\[
\mu_v(W) = \frac{1}{\dim W} \sum_{j=1}^{r_v} c_{v,j} \dim \left( \frac{W_{K^{(v)}}^j}{W_{K^{(v)}}^{j+1}} \right) = \frac{1}{\dim W} \sum_{j=1}^{r_v} c_{v,j} \left( \dim W_{K^{(v)}}^j - \dim W_{K^{(v)}}^{j+1} \right).
\]

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 \( \mu_o \) be the largest slope appearing, and among the subspaces of slope \( \mu_o \), let \( W_o \) be one of the largest dimension. A short calculation shows that if \( W' \) is a subspace with slope \( \mu_o \), then \( W' \subseteq W_o \), so \( W_o \) is the largest subspace of slope \( \mu_o \) both in dimension and in the partial ordering induced by inclusion. The bundle \( W_o \) is often called the “maximal destabilizing bundle”, or the “first step in the Harder-Narasimhan filtration”. We now fix \( W_o \) to be this subspace (rather than an arbitrary variable subspace). Note that \( W_o \neq \{0\} \).

Given the destabilizing bundle \( W_o \), set \( Z = \{ z \in X \mid s(z) = 0 \text{ for all } s \in W_o \} \). Since \( W_o \) is a nonzero subspace defined over \( k \), \( Z \) is a proper subvariety of \( X \) defined over \( k \).
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 \( y \in X(\kbar) \) 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 \( x \in U(\kbar) \), locally (with respect to \( d_v \)) near \( x \) the functions \( \|s(\cdot)\|_v \) and \( |g_s(\cdot)|_v \) are equivalent. In particular, if \( d_v(x_v, y_i) \to 0 \) as \( i \to \infty \) then the asymptotics of \( \|s(y_i)\|_v \) and \( |g_s(y_i)|_v \) are the same.

Finally, for each \( v \in S \) and \( j \in \{1, \ldots, r_v\} \) we choose a \( K^{(v)} \)-basis \( \{s_{v,j,\ell}\}_{\ell \in I_{v,j}} \) for \( W^{1}_{v, j} \). With this notation, the theorem [6, Theorem 9.1] of Faltings-Wüstholz is:

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

\[
\|s_{v,j,\ell}(y)\|_v < H_L(y)^{-\epsilon_{v,j}} \text{ for all } v \in S, \ j \in \{1, \ldots, r_v\}, \ \ell \in I_{v,j}.
\]

By definition of \( W_o \) we have the elementary estimate \( \mu(W_o) \geq \mu(V) \) and we will ensure the hypothesis \( \mu(W_o) > 1 \) 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 rational numbers \( 0 = \gamma_0 < \gamma_1 < \gamma_2 < \cdots < \gamma_r < \gamma_{\text{eff}} \) so that, if we define \( c_j \) by \( c_j = \gamma_j R \) and set \( \gamma_{r+1} = \gamma_{\text{eff}} \), we have

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

That is, the sum \( \sum c_j (f(\gamma_j) - f(\gamma_{j+1})) \) can be made arbitrarily close to \( \beta R \) by choice of rational \( \gamma_1, \ldots, \gamma_r \), and the choice \( c_j = R\gamma_j \) for \( j = 1, \ldots, r \).

**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 \( \gamma_1, \ldots, \gamma_r \in (0, \gamma_{\text{eff}}) \) we can clearly arrange for this approximation to be as close as we want to \( \beta \). \( \square \)

**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 best possible 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 2.15 that \( X \) is geometrically irreducible.

For each \( v \in S \) we let \( K^{(v)} \) be the field of definition of \( x_v, \pi_v: \overline{X^{(v)}} \to X^{(v)} \) 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
and so we conclude that $d(\gamma)$ that there are only finitely many solutions $\square$ Theorem 5.2.

For any $c_j \in V$ filtration on each $V_{\nu,j}$, we may also ensure that each $\dim_{K(\nu)} (mL_{\gamma_{\nu,j}})$ for $\nu \in S$. We may therefore apply Theorem 5.5 and conclude that there are only finitely many solutions $y \in X(k) \setminus Z(k)$ to

$$\sum_{\nu \in S} \left( \sum_{j=1}^{r_{\nu}} c_{\nu,j} (f_{\nu}(\gamma_{\nu,j}) - f_{\nu}(\gamma_{\nu,j+1})) \right) > 1,$$

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

For any $\gamma \geq 0$, $\dim_{K(\nu)} \Gamma(mL_{\gamma_{\nu,j}}) \Gamma(mL_{\gamma_{\nu,j}}) \rightarrow f_{\nu}(\gamma)$ as $m \rightarrow \infty$, and so by taking $m$ sufficiently large we may also ensure that each $\dim_{K(\nu)} \Gamma(mL_{\gamma_{\nu,j}}) \Gamma(mL_{\gamma_{\nu,j+1}})$ is sufficiently close to $f_{\nu}(\gamma_{j,v})$ so that

$$\sum_{\nu \in S} \frac{1}{\dim_{K(\nu)} \Gamma(mL)} \left( \sum_{j=1}^{r_{\nu}} c_{\nu,j} \left( \dim_{K(\nu)} \Gamma(mL_{\gamma_{\nu,j}}) \Gamma(mL_{\gamma_{\nu,j+1}}) \right) \right) > 1.$$  

(5.6.1)

Set $V = \Gamma(X, mL)$ and we identify $V_{K(\nu)}$ with $\Gamma(\tilde{X}(\nu), mL_{0}(\nu))$. We give a decreasing filtration on each $V_{K(\nu)}$ by setting $V_{K(\nu)}^{j} = \Gamma(mL_{\gamma_{\nu,j+1}})$ for $j = 1, \ldots, r_{\nu}$, and choose a $K(\nu)$-basis $\{s_{\nu,j,\ell}\} \in I_{\nu,j}$ for each $V_{K(\nu)}^{j}$. As above we let $W_{\nu}$ be the maximal destabilizing subspace and $Z$ the base locus of the sections in $W_{\nu}$. Equation (5.6.1) is the statement that $\mu(V) > 1$, and so we conclude that $\mu(W_{\nu}) > 1$ 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

$$\|s_{\nu,j,\ell}(y)\|_{V_{\nu}} \Gamma(mL_{\gamma_{\nu,j}}) H_{L}(y) < 1 \text{ for all } v \in S, j = 1, \ldots, r_{\nu}, \ell \in I_{\nu,j}.$$  

(5.6.2)

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 $\alpha_{v}(\{y_{i}\}, L) < \frac{1}{R_{\nu}}$ for each $v \in S$. This means that for all sufficiently small $\delta > 0$, and each $v \in S$, $d_{v}(x_{v}, y_{i})^{\frac{1}{R_{\nu}} - \delta} H_{L}(y_{i}) \rightarrow 0$ as $i \rightarrow \infty$.

Since each $s_{\nu,j,\ell}$ is in $V_{K(\nu)}^{j}$, each $s_{\nu,j,\ell}$ is in the $(m \gamma_{\nu,j})^{\delta}$ power of the maximal ideal of $x_{v}$, and so for any $\delta > 0$ and for large enough $i$ (depending on $\delta$) we have $\|s_{\nu,j,\ell}(y_{i})\|_{V} \leq d_{v}(x_{v}, y_{i})^{m \gamma_{\nu,j} - \delta}$. But then for large enough $i$

$$\|s_{\nu,j,\ell}(y_{i})\|_{V} \Gamma(mL_{\gamma_{\nu,j}}) H_{L}(y_{i}) \leq d_{v}(x_{v}, y_{i})^{\frac{1}{R_{\nu}} - \frac{\delta}{m \gamma_{\nu,j}}} H_{L}(y_{i})$$  

(5.6.3)

for all $v \in S, j = 1, \ldots, r_{\nu}, \ell \in I_{\nu,j}$. 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. $\square$

---

3If $X$ is not normal, $\Gamma(X, mL) \otimes_{k} K(\nu)$ may only be a proper subspace of $\Gamma(\tilde{X}(\nu), mL_{0}(\nu))$. However, since the volume is a birational invariant, the asymptotic calculations go through without change and we omit further mention of this detail.
6. Roth theorems

Let $X$ be an $n$-dimensional variety defined over $k$, and irreducible over $k$. In this section we present theorems giving lower bounds for $\alpha_x(L)$ independent of the field of definition of $x \in X(\overline{k})$, 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 $v$ of $k$, an extension of $v$ to $\overline{k}$, and compute $\alpha_x$ with respect to $d_v(\cdot, \cdot)$.

**Theorem 6.1.** (Roth’s theorem I) For any ample $\mathbb{Q}$-bundle $L$ and any $x \in X(\overline{k})$ either

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

or

(b) There exists a proper subvariety $Z \subset X$ defined over $k$, irreducible over $\overline{k}$, and containing $x$ so that $\alpha_{x,Z}(L) = \alpha_{x,Z}(L|_Z)$, 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 $Z$ defined over $k$ 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 2.15 to $Z$, and use Proposition 2.12(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.** (Roth’s theorem II) 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 $X$ (defined over $k$ and irreducible over $\overline{k}$, as above).

This immediately yields

**Theorem 6.3.** (Roth’s Theorem III) With the same hypotheses as above, $\alpha_x(L) \geq \frac{1}{2}\epsilon_x(L)$, with equality if and only if both $\alpha$ and $\epsilon$ 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) $\epsilon_{x,C}(L|_C) = \epsilon_{x,C}(L) = \epsilon_{x,C}(L)$.

**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 2.12(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 $C$ (defined over $k$ and irreducible over $\overline{k}$), and $\epsilon_x$ 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 2.12(c). By the first part of the theorem we always have $\alpha_{x,X}(L) \geq \frac{1}{2}L^d$, 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
\[ d_v(x,y) < H_L(y)^{-\left(\frac{3}{\epsilon_{x,X}(L)} + \delta\right)}. \]

**Proof:** This is immediate from Theorem 6.3 and Proposition 2.9. \(\square\)

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 $x \in X(\overline{k})$, and no rational curve passes through $x$ then $\alpha_x(L) \geq \frac{3}{5} \epsilon_x(L)$. Equivalently, for any $\delta > 0$ there are only finitely many solutions $y \in X(k)$ to
\[ d_v(x,y) < H_L(y)^{-\left(\frac{3}{\epsilon_{x,X}(L)} + \delta\right)}. \]

**Remark.** Theorems 6.1 and 6.2 were stated for a variety $X$ irreducible over $k$ since if $X$ were reducible, and $x \in X(\overline{k})$ 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 2.12(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 $\overline{k}$ and $X$ be an $n$-dimensional variety defined and irreducible over $k$. For each $v \in S$ we choose a point $x_v \in X(\overline{k})$, 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 $\alpha_v$ to be $\alpha_{x_v}$, 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 $\xi$ 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 $\xi_v$ for the value of $\xi$ 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 $\Xi$ 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 $\sum_{v \in S} \xi'_v \Delta'_v \leq \xi_v \Delta_v$ 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 $\Xi$ 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 $\xi'_v$ are rational with denominator dividing $N$). Given a function $\xi: S \rightarrow \mathbb{R}_{\geq 0}$ with $\sum_{v \in S} \xi_v \geq 1$ set

$$
\xi''_v = \frac{\left[\frac{N \Delta'_v \xi_v}{\Delta_v}\right]}{N} \quad \text{for each } v \in S.
$$

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

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

for each $v \in S$ we conclude that

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

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

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 $X$ defined over $k$, 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) \hspace{1cm} \begin{cases} 
\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}{\text{RC}_v}. 
\end{cases}

(7.3.2) \hspace{1cm} \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) \hspace{1cm} \begin{cases} 
\text{For all weighting functions } \xi: S \rightarrow [0, 1] \text{ and any collection } \{\delta_v\}_{v \in S} \text{ of positive real numbers, there are only finitely many solutions } y \in X(k) \setminus Z(k) \text{ to } 
\end{cases}

$$
d_v(x_v, y) = H_L(y)^{-\xi_v(1+\delta_v)} \text{ for all } v \in S.
$$
(7.3.4) \[
\begin{aligned}
\text{For all } \delta > 0 \text{ there are only finitely many solutions } y \in X(k) \setminus Z(k) \text{ to } \\
\prod_{v \in S} d_v(x_v, y)^{\frac{1}{\alpha_v}} < H_L(y)^{-(1+\delta)}.
\end{aligned}
\]

Proof: (7.3.1) $\implies$ (7.3.2): Given a sequence $\{x_i\}$ in $X(k) \setminus Z(k)$, set $D = \sum_{v \in S} \frac{1}{\alpha_v(x_i, L)}$. 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 $\xi_v = \frac{1}{\sum_{\alpha_v(x_i, L)}}$ 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 $\xi_v$ and clearing denominators gives (7.3.2) (recall that $\alpha_v(x_i, L) > 0$ by Proposition 2.12(d)).

(7.3.2) $\implies$ (7.3.1): If (7.3.1) is false then there is a weighting function $\xi$ and a sequence $\{x_i\}$ in $X(k) \setminus Z(k)$ such that $\frac{1}{\sum_{\alpha_v(x_i, L)}} > \xi_v$ 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
\[
d_v(x_v, y_i)^{\frac{1}{\alpha_v}} = H_L(y_i)^{-(1+\delta)}
\]
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 $\delta'$ less than $\delta$. Applying Lemma 7.2, with $\Delta_\xi = 1 + \delta'$ and $\Delta_v = 1 + \delta$ for all $v \in S$ we obtain a finite set $\Xi$ of weighting functions so that for any $\xi: S \longrightarrow \mathbb{R}_{\geq 0}$ 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 $\xi_i$ as above. Since $\Xi$ is a finite set, by passing to a subsequence of $\{y_i\}$ there is a $\xi'_i \in \Xi$ which works for all $i$. Choosing $\delta_v = \delta'$ for each $v$, we have infinitely many solutions to $d_v(x_v, y)^{\frac{1}{\alpha_v}} < H_L(y)^{-\xi'_v(1+\delta_v)}$, 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 $\xi$ such that $\alpha_v(\{x_i\}, L) < \frac{1}{\sum_{\alpha_v(x_i, L)}}$ for each $v \in S'$, where $S' = \{v \in S \mid \xi_v \neq 0\}$. For $\delta > 0$ small enough we will still have $\alpha_v(\{x_i\}, L) < \frac{1}{\sum_{\alpha_v(x_i, L)}}$ for each $v \in S'$, and so by definition of $\alpha_v$, $d_v(x_v, x_i)^{\frac{1}{\sum_{\alpha_v(x_i, L)}}} H_L(x_i) \to 0$ or equivalently $d_v(x_v, x_i)^{\frac{1}{\sum_{\alpha_v(x_i, L)}}} H_L(x_i)^{\xi_v(1+\delta)} \to 0$, as $i \to \infty$ for all $v \in S'$. Thus by omitting finitely many of the initial $x_i$ we can make the product
\[
\prod_{v \in S'} d_v(x_v, x_i)^{\frac{1}{\alpha_v}} H_L(x_i)^{\xi_v(1+\delta)} = \left(\prod_{v \in S'} d_v(x_v, x_i)^{\frac{1}{\alpha_v}}\right) H_L(x_i)^{(1+\delta)}
\]
as small as desired. The product $\prod_{v \in S', S'} d_v(x_v, x_i)^{\frac{1}{\alpha_v}}$ 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
\[
\prod_{v \in S} d_v(x_v, x_i)^{\frac{1}{\alpha_v}} < H_L(x_i)^{-(1+\delta)}
\]
contradicting (7.3.4).

(7.3.1) $\implies$ (7.3.3): Assume that (7.3.3) is false. Then there is a weighting function $\xi$ 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
Proof: We will show condition (7.3.3) holds for the collection \( \{v_i\} \) with respect to the collection \( \{v_j\} \) given any weighting function \( \xi \) and any \( \delta > 0 \) we will show that there are only finitely many solutions \( y \in X(k) \) to

\[
d_v(x_v, y) = \frac{1}{\alpha_v(y)} H_L(y) < 1 \quad \text{for all } v \in S', \quad \text{so we conclude that } \frac{1}{R_v \xi_v(1 + \delta)} \in A_x((\{y_i\}, L)).
\]

Thus \( \alpha_v((\{y_i\}, L)) \leq \frac{1}{\alpha_v \xi_v(1 + \delta)} < \frac{1}{R_v} \) for \( v \in S' \), contradicting (7.3.1). □

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 \( Z \subset X \) defined over \( k \) 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 \( R_v > \frac{1}{\xi_v(1 + \delta)} \) 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, \( \beta_{x_v}(L) \geq \frac{n}{n+1} \xi_{x_v}(L) \), 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 \( \Xi \) so that for any weighting function \( \xi \) 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 \( Z_{\xi'} \) defined over \( k \) such that for any sequence \( \{x_i\} \) in \( X(k) \setminus Z_{\xi'}(k) \) there is at least one \( v \in S' \) with \( \alpha_v((\{x_i\}, L), L) \geq \frac{1}{\xi'_v R'_v} \).

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

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.) □

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

**Theorem 7.5.** Set \( R_v = \frac{1}{\xi_v(L)} \) 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 \( \xi \) and any \( \delta > 0 \) we will show that there are only finitely many solutions \( y \in X(k) \) to

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

Suppose a weighting function \( \xi \) is given. When \( \dim X = 1 \) the result we want to prove is Theorem 7.4(b). If \( \dim X = n > 1 \) then \( 2 > \frac{n}{n+1} \) so by Theorem 7.4(b) again there is a proper subvariety \( Z' \subset X \) defined over \( k \) 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
\[
d_v(x_v, y) \left( \frac{\epsilon_{x_v, Z_j}(L)}{2} \right)^{\epsilon_{x_v, Z_j}(L)} < H_L(y)^{-\xi_v(1+\delta)} \quad \text{for all } v \in S.
\]
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

\[
\prod_{v \in S} d_v(x_v, y)^{\epsilon_{x_v}(L)} < H_L(y)^{-2(\delta \cdot \xi_v)}.
\]

**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 \( \xi \) such that

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

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) \( \kappa(x_v) \leq k_v \), 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 nonempty subset \( S' \subseteq S \), then for any weighting function \( \xi : S' \to (0, 1] \) (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 \( |y - x_v|_v < H(x)^{-e_v} \) 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 $(a, b) \in V^2$ such that $|a_v - x_v b_v|_v < B^{-f_v}$ and $|b_v|_v < N B^{f_v}$ for all places $v$ in $S$. Choose $N$ large enough so that the volume of $1\frac{1}{2} A_N$ 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 [14, §1.4].

To see this, consider the sets $1\frac{1}{2} A \cap (D + u)$ as $u$ varies over elements of $R^2$. They clearly cover the set $1\frac{1}{2} A$, and for each $u$, we have $1\frac{1}{2} A \cap (D + u) = ((1\frac{1}{2} A - u) \cap D) + u$. Therefore, the volume of $1\frac{1}{2} A \cap (D + u)$ is equal to that of $(1\frac{1}{2} A - u) \cap D$. If the sets $(1\frac{1}{2} A - u) \cap D$ were pairwise disjoint, then by summing over $u$, we would find that the volume of $1\frac{1}{2} A$ is at most the volume of $D$, in contradiction to our choice of $A$. We conclude that the sets $(1\frac{1}{2} A - u) \cap D$ are not disjoint.

We may therefore find elements $u, v \in R^2$ and $a_1, a_2 \in A$ such that $1\frac{1}{2} a_1 - u = 1\frac{1}{2} a_2 - v$. Since $A$ is convex and closed under multiplication by $-1$, it follows that $u - v$ is a nonzero element of $R^2 \setminus A$, 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 $|b_B|_v < N B^{f_v}$). 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 $2 e_v/e - \delta > e_v$ for each $v \in S$. Fix one such $\delta$. By the Schmidt Subspace Theorem (see [1, Corollary 7.2.5]) applied to the linear forms $a - b x_v$ 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 $-x_v m$ 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 $|b|_v > H((a:b))^{f_v - \delta}$, and hence infinitely many $a/b \in k$ satisfying $|a/b - x_v|_v \leq H((a:b))^{-2f_v + \delta} = H((a:b))^{-2e_v/e + \delta} \leq H((a:b))^{-e_v}$. □

Given a sequence $\{y_v\} \in \mathbb{P}^1(k)$ set $\tau_v(\{y_v\}) = 1/\alpha_v (\{y_v\}, \mathcal{O}_{\mathbb{P}^1}(1))$. Since $\tau$ is the reciprocal of $\alpha$, if $\tau' < \tau_v(\{y_v\})$ (respectively $\tau' > \tau_v(\{y_v\})$) then $\lim_{v \to \alpha_v} d_v(x_v, y_v)^{1/r_v} H(y_v) = 0$ (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 $\{y_v\}$ such that $e_v \leq \tau_v(\{y_v\})$ 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 \{\epsilon_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 $\epsilon_v = \tau_v(\{y_i\})$ 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}\}_{n \geq 0} such that $e_v - \frac{1}{2n} < \tau_v(\{y_{i,n}\})$. Since $e_v - \frac{1}{n} < e_v - \frac{1}{2n}$, we have $\lim_{i \to \infty} d_v(x_v, y_{i,n})^{-\epsilon_v/2n} H(y_{i,n}) = 0$ for all $v \in S$. For each fixed $n$, by choosing $i$ large enough, we may pick $y_n = y_{i,n}$ so that $d_v(x_v, y_n)^{-\epsilon_v/2n} H(y_n) < \frac{1}{n}$ and $d_v(x_v, y_n) < 1$ for all $v \in S$. In this way we construct a sequence \{y_n\}_{n>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

$$d_v(x_v, y_n)^{-\epsilon_v/2n} H(y_n) < d_v(x_v, y_n)^{-\epsilon_v/2n} H(y_n) < \frac{1}{n}.$$

Therefore $\lim_{n \to \infty} d_v(x_v, y_n)^{-\epsilon_v/2n} H(y_n) = 0$ and so $e_v - \delta < \tau_v(\{y_n\})$. Letting $\delta$ go to zero we conclude that $e_v < \tau_v(\{y_n\})$ 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 $v \in S$. □

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 $C$ defined over $k$, necessarily 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 2.14 for a single point). The converse direction is Corollary 7.9 with the choice $e_v = 2\epsilon_v$ for all $v \in S$, combined with the appropriate modification for the singularity, again as in Theorem 2.14. □

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 $R_v > \frac{m}{(m+1)^{m+1} \epsilon_v(L)}$ 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 $R_v = \frac{3}{2 \epsilon_v(L)}$ 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 $\xi$ and any collection \{\delta_v\}_{v \in S} of positive real numbers, there are only finitely many solutions $y \in X(k)$ to

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

Given the collection \{\delta_v\}_{v \in S} set $\delta_v' = \frac{\delta_v}{2}$ and $R_v' = \frac{3}{2 \epsilon_v(L)} + \delta_v'$ 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). □
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 \((\alpha_v, \{x_i\}, L)\) 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 and McIntyre (see [16, p. 100] and [16, §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 \(\varphi: Y \to X\) between irreducible varieties. An *unramified cover defined over \(k\)* means in addition that \(\varphi\), \(X\) and \(Y\) are defined over \(k\), and that \(X\) and \(Y\) are irreducible over \(k\).

Let \(\varphi: Y \to X\) be an unramified cover and \(x\) be any point in \(X\). 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 a limits over such unramified covers. The point of this section is that the theorems in §6 and §7 hold with \(\beta\) and \(\epsilon\) replaced by their limiting versions. 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 \(\beta\) and \(\epsilon\) 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 $\varphi: Y \to X$ be an unramified cover, $L$ an ample line bundle on $X$, $x$ any point of $X(\overline{k})$, 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 $\overline{X}_1, \ldots, \overline{X}_r$ and $\overline{Y}_1, \ldots, \overline{Y}_s$ be the irreducible components of $X \times_k \overline{k}$ and $Y \times_k \overline{k}$ containing $x$ and $y$ respectively. Each $\overline{Y}_i$ maps to some $\overline{X}_j$, and this map expresses $\overline{Y}_i$ as an unramified cover of $\overline{X}_j$. Since $\beta$ and $\epsilon$ are defined as minima over irreducible components, establishing the conclusion of the lemma for each map $\overline{Y}_i \to \overline{X}_j$ establishes the lemma for $Y \to X$. 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 $\overline{X}_j$ and $\overline{Y}_i$, and $\varphi$ as the name of the map.

Let $\pi_X: \overline{X} \to X$ 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 $\pi_Y: \overline{Y} \to Y$ 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 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 $\gamma$.

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 $\text{Vol}(\varphi^*L_{X,\gamma}) = d \text{Vol}(L_{X,\gamma})$.

For any rational $\gamma \geq 0$, and any $m > 0$ sufficiently divisible so that $m\gamma$ is integral, global sections of $mL_{X,\gamma}$ pull back to global sections of $m\varphi^*L$ vanishing to order at least $m\gamma$ at all points in $\varphi^{-1}(x)$. In particular, these sections are a subspace of the global sections of $m\varphi^*L$ vanishing to order at least $m\gamma$ at $y$. We thus have

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

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

We now prove (b). Let $\varphi^{-1}(x) = \{y_1, \ldots, y_\ell\}$ with $y_1 = y$. Since $\varphi$ is unramified, the fibre product $Y \times_X \overline{X}$ is the blow up of $Y$ at the points $y_1, \ldots, y_\ell$. Let $\psi_X$ and $\psi_Y$ be the maps from $Y \times_X \overline{X}$ to $\overline{X}$ and $\overline{Y}$ respectively (the map to $\overline{Y}$ being the blow down at the points of $\varphi^{-1}(x)$ different from $y$). For $i = 2, \ldots, \ell$ let $E_i$ be the exceptional divisor of $Y \times_X \overline{X}$ lying over $y_i$. Then for any $\gamma, \psi_X^*L_{X,\gamma} = \psi_Y^*L_{Y,\gamma} - \gamma (\sum_{i=2}^\ell E_i)$.

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

Consider the full subcategory of $k$-schemes over $X$ whose objects are the unramified covers $\varphi: Y \to X$ defined over $k$ as above (and recall that this means that $Y$ is irreducible). If $(Y_1, \varphi_1)$ and $(Y_2, \varphi_2)$ are objects and $\psi: Y_1 \to Y_2$ a morphism in this category, then $\psi$ expresses $Y_1$ as an unramified cover of $Y_2$, and thus Lemma 8.1 applies. In particular, for any $y_2 \in Y_2$, $\min_{y_1 \in \psi^{-1}(y_2)}(\beta_{y_1}(\varphi_1^*L)) \geq \beta_{y_2}(\varphi_2^*L)$ and similarly for $\epsilon$. 
Definition 8.2. Let $X$ be an irreducible variety defined over $k$, $L$ an ample line bundle on $X$ defined over $k$ and $x \in X$. We define

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

where the limits are over the category of unramified covers $\varphi: Y \to X$ defined over $k$.

We allow $\infty$ as a limit; by Lemma 8.1 and the remarks above these limits exist.

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 $k$, $L$ an ample line bundle on $X$, and $x_1, \ldots, x_\ell$ finitely many closed points of $X$. Suppose that $\beta_1, \ldots, \beta_\ell$ are positive real numbers with $\beta_i < \hat{\beta}_{x_i}(L)$ for $i = 1, \ldots, \ell$. Then there exists an unramified cover $\varphi: Y \to X$ such that $\min_{y \in \varphi^{-1}(x_i)}(\beta_y(\varphi^*(L))) > \beta_i$ for $i = 1, \ldots, \ell$.

Proof: By the definition of $\hat{\beta}$, for each $i$ there is an unramified cover $\varphi_i: Y_i \to X$ 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 $Y_1 \times_X \cdots \times_X Y_\ell$, and $\varphi: Y \to X$ the induced map. The natural projection maps of the fibre product induce maps $\psi_i: Y \to Y_i$ for each $i$, and $\psi_i$ 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 $\psi_i$ shows that $Y$ has the desired property. $\square$

Lifting sequences. Let $\varphi: Y \to X$ be an unramified cover defined over $k$. By the theorem of Chevalley-Weil [1, Theorem 10.3.11] there is a finite extension $F/k$ such that all points $\{y \in Y(\kbar) \mid \varphi(y) \in X(k)\}$ are defined over $F$ (this field $F$ is not unique, since any larger field will also work). Given a sequence $\{x_i\}$ of points of $X(k)$ we arbitrarily choose $y_i \in \varphi^{-1}(x_i)$ for each $i$. We call such a sequence $\{y_i\}$ a lift of $\{x_i\}$, and $F$ the field of definition of $\{y_i\}$. This lift is somewhat haphazard, but by further passing to a subsequence we may obtain a lift with better properties.

Let $v$ be a place of $k$ (extended to $\kbar$) and suppose that there is $x \in X(\kbar)$ such that $d_v(x, x_i) \to 0$, i.e., that $\{x_i\}$ approximates $x$ with respect to $d_v(\cdot, \cdot)$. Since $Y(F_v)$ is compact by passing to a subsequence we may assume that the sequence $\{y_i\}$ has a limit $y \in Y(F_v)$. The topology on $Y(F_v)$ is that induced by $d_v(\cdot, \cdot)$, and so this means that $d_v(y, y_i) \to 0$ as $i \to \infty$. Furthermore, by continuity we have $\varphi(y) = x$, in particular, $y \in Y(\kbar)$.

We will need a generalization obtained by repeating this procedure. First assume that we have lifted $\{x_i\}$ to a sequence $\{y_i\}$ with field of definition $F$. Let $v$ be a place of $k$, and let $T_v$ be the set of places of $F$ dividing $v$. For each $w \in T_v$ choose an extension of $w$ to $\overline{F} = \kbar$, which we continue to call $w$. These extensions define distance functions $d_w(\cdot, \cdot)_F$ on varieties defined over $F$, well defined up to equivalence.

Assume again that there is an $x \in X(\kbar)$ so that $d_v(x, x_i) \to 0$. By applying the procedure above to each $w \in T_v$ in turn, we may find $y_w \in \varphi^{-1}(x)$ for each $w \in T_v$, and a subsequence of $\{y_i\}$ so that for each $w \in T_v$, $d_w(y_w, y_i)_F \to 0$ as $i \to \infty$.

Finally, given a finite set $S$ of places of $k$, we may repeat this process for each $v \in S$. We record the conclusion below.
Proposition 8.4. Let \( \varphi: Y \to X \) be an unramified cover defined over \( k \), and \( S \) a finite set of places of \( k \). Suppose that \( \{ x_v \}_{v \in S} \) are a set of points of \( X(\overline{k}) \), and that \( \{ x_i \} \) is a sequence of \( k \)-points so that \( d_v(x_v, x_i) \to 0 \) for each \( v \in S \).

Then by passing to a subsequence of \( \{ x_i \} \) we may find a lift \( \{ y_i \} \) of \( \{ x_i \} \) to \( Y \) with field of definition \( F \) and for each \( v \in S \) and \( w \in T_v \) we may find a \( y_w \in \varphi^{-1}(x_v) \) such that \( d_w(y_w, y_i)_F \to 0 \) as \( i \to \infty \). Here, as above, each \( T_v \) is the set of places of \( F \) dividing \( v \), and \( d_w(\cdot, \cdot)_F \) a distance function obtained by extending \( w \) to \( \overline{k} \).

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 \( \alpha \). It suffices to compare \( d_v(x_v, x_i)_k \) with \( d_v(y_w, y_i)_F \) (for some extension of \( v \) to \( \overline{k} \)), since each \( d_w(\cdot, \cdot)_F \) is a such an extension.

Lemma 8.5. Let \( \varphi: Y \to X \) be an unramified cover defined over \( k \), \( v \) a place of \( k \), and \( \{ x_i \} \) a sequence of points of \( X(\overline{k}) \) converging to \( x \in X(\overline{k}) \) with respect to \( d_v(\cdot, \cdot) \). Let \( \{ y_i \} \) be a lift of \( \{ x_i \} \) converging to \( y \in Y(\overline{k}) \) with respect to \( d_v(\cdot, \cdot)_F \), \( F \) the field of definition of \( \{ y_i \} \), and set \( m_v = [F_v:k_v] \) and \( e = [F:k] \).

Then \( d_v(y_i, y_i)_F \) is asymptotically equivalent to \( d_v(x_i, x_i)^{m_v} \) as \( i \to \infty \), and for any line bundle \( L \) on \( X \) (defined over \( k \)) \( \alpha_y(\{ y_i \}, \varphi^*L)_F = \frac{1}{m_v} \alpha_x(\{ x_i \}, L)_k \).

Here, as in Proposition 2.10, the subscript on \( d_v(\cdot, \cdot)_F \) indicates which field is being used to normalize the distance function and the subscript on \( \alpha(\cdot, \cdot)_F \) similarly indicates the field used to normalize the distance function and the height.

Proof: For each \( y_i \) in the sequence, we have \( H_{\varphi^*L}(y_i)_F = H_L(x_i)_F = H_L(y_i)_k \). Thus once we show that \( d_v(y, y)_F \) is asymptotically equivalent to \( d_v(x, x)^{m_v} \) as \( i \to \infty \) the equality \( \alpha_y(\{ y_i \}, \varphi^*L)_F = \frac{1}{m_v} \alpha_x(\{ x_i \}, L)_k \) follows immediately as in the proof of Proposition 2.10.

By our normalizations for computing distance we have \( d_v(x, x)_F = d_v(x, x)^{[F_v:k_v]} = d_v(x, x)_k^{m_v} \), and so it suffices to show that \( d_v(y, y)_F \) is asymptotically equivalent to \( d_v(x, x)_F \) as \( i \to \infty \). Base changing to \( F \), we may replace \( k \) by \( F \) and simply assume that \( \{ y_i \} \) is defined over \( k \).

Let \( U \) be an affine neighbourhood of \( x \) defined over \( k \) and let \( u_1, \ldots, u_r \) be elements of \( \Gamma(U, \mathcal{O}_X) \) which generate the maximal ideal of \( x \). By Lemma 2.4, \( d_v(x, x)_F \) is equivalent to \( \max(|u_1(x_i)|_v, \ldots, |u_r(x_i)|_v) \) as \( i \to \infty \).

Let \( V = \varphi^{-1}(U) \). Since \( \varphi \) is unramified, if we restrict to an affine open neighbourhood \( V' \) of \( y \) in \( V \) not containing the other points of \( \varphi^{-1}(x) \), then \( \varphi^*u_1, \ldots, \varphi^*u_r \) generate the maximal ideal of \( y \). Thus by Lemma 2.4 again \( d_v(y, y)_F \) is equivalent to \( \max(|\varphi^*u_1(y_i)|_v, \ldots, |\varphi^*u_r(y_i)|_v) \) as \( i \to \infty \). Since \( \varphi^*(u_j)(y_i) = u_j(x_i) \) for each \( j = 1, \ldots, r \) and each \( i \), this establishes the equivalence. \( \square \)

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 \( \alpha_{y_w}(\{ y_i \}, \varphi^*L)_F = \frac{[F:k]}{[F_v:k_v]} \alpha_x(\{ x_i \}, L)_k \).

Here \( \alpha_{y_w} \) is computed with respect to \( d_w(\cdot, \cdot)_F \) and \( \alpha_x \) with respect to \( d_v(\cdot, \cdot)_F \).

We are now ready to establish the version of Theorem 5.2 with \( \hat{\beta} \) in place of \( \beta \).
Theorem 8.7. Let $X$ be an irreducible variety defined over $k$, $S$ a finite set of places of $k$, and for each $v \in S$ choose an $x_v \in X(K)$. 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

\[(8.7.1) \quad \sum_{v \in S} \beta_{x_v}(L)R_v > 1.\]

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 $\varphi: Y \to X$ defined over $k$, with $Y$ irreducible, satisfying

\[(8.7.2) \quad \sum_{v \in S} \left( \min_{y_v \in \varphi^{-1}(x_v)} (\beta_{y_v}(\varphi^*L)) \right)R_v > 1,\]

We fix this cover for the rest of the argument. We will prove the theorem in the form of condition (5.1.1), that is, we will show that there is a proper subvariety $Z \subset X$ defined over $k$, such that for all sequences $\{x_i\}$ of $k$-points of $X$ not essentially contained in $Z$ there is at least one $v \in S$ so that $\alpha_v(\{x_i\}, L) > \frac{1}{R_v}$. Here as in §5 for each $v \in S$ we use $\alpha_v$ to mean $\alpha_{x_v}$, computed with respect to the distance $d_v(\cdot, \cdot)$.

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 $i \to \infty$, 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 $d_v(x_v, x_i) \to 0$ as $i \to \infty$ 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 $\alpha$, so we may freely do so in proving the result.

As in the discussion above, we may find a finite extension $F/k$ so that any sequence $\{x_i\}$ of $k$-points of $X$ can be lifted to a sequence $\{y_i\}$ of points of $Y$ with field of definition $F$. We may assume that $F/k$ is a Galois extension and we fix $F$ for the rest of the proof. By Proposition 8.4 if $\{x_i\}$ converges to each $x_v$ with respect to $d_v(\cdot, \cdot)$ we may, after passing to a subsequence of $\{y_i\}$, choose $y_w \in \varphi^{-1}(x_v)$ for all $v \in S$ and $w \in T_v$ so that $\{y_i\}$ converges to $y_w$ with respect to $d_w(\cdot, \cdot)$. (As above, for each $v \in S$, $T_v$ is the set of places of $F$ dividing $v$, and we have extended each $w$ to $\overline{k}$ in some way in order to define $d_w(\cdot, \cdot)_F$.)

Let $Y_1, \ldots, Y_r$ be the irreducible (over $F$) components of $Y \times_k F$, and $\varphi_j: Y_j \to X$, $j = 1, \ldots, r$ the induced maps ($Y_j$ is defined over $F$, and hence $\varphi_j$ is a map of $k$-schemes, but not one defined over $k$). By passing to a subsequence we can assume that any lift $\{y_i\}$ is contained in one of the components. For each $Y_j$ there are only finitely many choices of $y_w \in \varphi_j^{-1}(x_v)$ for each $v \in S$, $w \in T_v$. It thus suffices to prove that if we fix $Y_j$ and a choice of $y_w \in \varphi_j^{-1}(x_v)$ for each $v \in S$ and $w \in T_v$ then there exists a proper subset $Z_{j, \{y_w\}} \subset X$ defined over $k$ (depending on these choices) so that for any sequence $\{x_i\}$ of $k$-points of $X$ which

1. is not essentially contained in $Z_{j, \{y_w\}}$,
2. lifts to a sequence $\{y_i\}$ in $Y_j$ (necessarily defined over $F$) that converges with respect to $d_w(\cdot, \cdot)$ to $y_w$ for each $v \in S$, and $w \in T_v$

there is at least one $v \in S$ so that $\alpha_v(\{x_i\}, L) > \frac{1}{R_v}$. Taking $Z$ to be the union over the finitely many such $Z_{j, \{y_w\}}$ then yields the general case.
We now assume that we have fixed these choices. Set \( g_v = \left[ \frac{F_v}{k} \right] \) for each \( v \in S \). If \( \{x_i\} \) is a sequence of \( k \)-points satisfying (2) above, and \( \{y_i\} \) such a lift, then by Corollary 8.6 and the fact that \( F/k \) is Galois we know that \( \alpha_w(\{x_i\}, \varphi^* L)_F = g_v, \alpha_v(\{x_i\}, L)_k \) for each \( w \in T_v \).

Here we use \( \alpha_w \) to mean \( \alpha_{yw} \) computed with respect to \( d_w(\cdot, \cdot) \), and the subscripts \( F \) and \( k \) to indicate the field used to normalize the distance and the heights.

For each \( v \in S \) and \( w \in T_v \) set \( R'_w = R_v/g_v \) and let \( T = \bigcup_{v \in S} T_v \). We have \( \beta_{yw,y_j}(\varphi_j^*(L)) \geq \min_{y \in \varphi^{-1}(x_y)} (\beta_y(\varphi^* L)) \) for each \( v \in S \), \( w \in T_v \) (since \( y_w \in \varphi^{-1}(x_y) \) and \( y_j \in Y_j \)), and we also have \( \#T_v = g_v \) for each \( v \in S \). Thus (8.7.2) implies

\[
\sum_{w \in T} \beta_{yw,y_j}(\varphi_j^* L) R'_w > 1.
\]

We now apply Theorem 5.2 to the collection \( \{R'_w\}_{w \in T} \) and line bundle \( \varphi_j^*(L) \), and let \( Z' \) be the resulting proper subvariety of \( Y_j \) defined over \( F \). Set \( Z_j,\{y_i\} \) to be the image of \( Z' \) in \( X \) (\( Z_j,\{y_i\} \) is defined over \( k \) since \( \varphi_j \) is a morphism of \( k \)-schemes). Since \( Y_j \) is irreducible, \( Z' \) is of dimension strictly less than \( Y_j \), and hence \( Z_j,\{y_i\} \) is again a proper subvariety of \( X \).

Suppose that \( \{x_i\} \) is a sequence of \( k \)-points of \( X \) satisfying (1) and (2) above, and let \( \{y_i\} \) be such a lift. Then \( \{y_i\} \) is not essentially contained in \( Z' \), thus by construction of \( Z' \) there is at least one \( v \in S \) and \( w \in T_v \) so that

\[
g_v \alpha_v(\{x_i\}, L)_k = \alpha_w(\{y_i\}, \varphi_j^* L)_F \geq \frac{1}{R'_w} = g_v R_v.
\]

Thus \( \alpha_v(\{x_i\}, L)_k \geq \frac{1}{R'_w} \). This completes the proof of Theorem 8.7. \( \square \)

**Remark.** If one of the \( \hat{\beta}_{x_i}(L) \) is infinite then condition (8.7.1) holds for any collection \( \{R_w\}_{w \in S} \) of positive numbers.

**Proposition 8.8.** Let \( X \) be an irreducible variety of dimension \( n \), \( L \) an ample line bundle on \( X \), and \( x \) a closed point of \( X \). Then

(a) \( \hat{\beta}_x(L) \geq \frac{n}{n+1} \hat{\epsilon}_x(L) \), and

(b) for any irreducible subvariety \( Z \subset X \) and any \( x \in Z \), \( \hat{\epsilon}_{x,Z}(L|_Z) \geq \hat{\epsilon}_{x,X}(L) \).

**Proof:** Let \( \varphi: Y \to X \) 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 limit over such covers.

For part (b), let \( Z' \) be any irreducible component of \( \varphi^{-1}(Z) \), where \( \varphi: Y \to X \) is an unramified cover as above. The induced map \( \psi: Z' \to Z \) 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

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

Taking the limits 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 2.12(f), as well as arguments common in Diophantine approximation. The necessary results about \( \alpha, \beta, \epsilon, \) and their limiting versions needed to make these arguments are summarized in the following table.
Results about $\alpha$, $\beta$ and $\epsilon$

| Theorem 5.2 | $\sum_v \beta_x(L) R_v > 1 \implies (5.1.1) + (5.1.2)$ |
|-------------|------------------------------------------------------------------------------------------------|
| Corollary 4.4 | $\beta_x(L) \geq \frac{n}{n+1} \epsilon_x(L)$ |
| Proposition 3.4(c) | $\epsilon_x,\tilde{z}(L|z) \geq \epsilon_x, Z(L)$ |
| Proposition 2.12(f) | $\alpha_x, X(L) = \min(\alpha_x, x_1(L|X_1), \ldots, \alpha_x, x_r(L|X_r))$ |

Results about $\alpha$, $\beta$ and $\hat{\epsilon}$

| Theorem 8.7 | $\sum_v \beta_x(L) R_v > 1 \implies (5.1.1) + (5.1.2)$ |
|-------------|------------------------------------------------------------------------------------------------|
| Proposition 8.8(a) | $\beta_x(L) \geq \frac{n}{n+1} \hat{\epsilon}_x(L)$ |
| Proposition 8.8(b) | $\hat{\epsilon}_x,\tilde{z}(L|z) \geq \epsilon_x, Z(L)$ |
| Proposition 2.12(f) | $\alpha_x, X(L) = \min(\alpha_x, x_1(L|X_1), \ldots, \alpha_x, x_r(L|X_r))$ |

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 $\beta$ and $\epsilon$. 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 $\beta$ and $\epsilon$ replaced by $\hat{\beta}$ and $\hat{\epsilon}$.

**Remark.** The larger the values of $\beta$ and $\epsilon$ the stronger these types of results are. In particular, this means that given any lower bounds for $\beta$ and $\epsilon$ the results listed in Corollary 8.9 hold with the lower bounds used in place of $\beta$ or $\epsilon$. One method of getting lower bounds for $\beta$ and $\epsilon$ which still takes into account the asymptotic behaviour of covers is to consider only étale Galois covers $\varphi: Y \to X$ 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

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

where the limits are over irreducible étale Galois covers $\varphi: Y \to X$, 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 $x \in X$ and ample $L$.

**Example.** Let $X$ be an abelian variety and let $[m]: X \to X$ denote the multiplication by $m$ map. For any ample line bundle $L$, $[m]^* L$ has the same numerical class as $m^2 L$, and so $\epsilon_x([m]^* L) = m^2 \epsilon_x(L)$. In particular, $\hat{\epsilon}_x^\text{ét}(L) = \infty$ and thus $\hat{\epsilon}_x(L) = \infty$. Therefore $\alpha_x(L) \geq \frac{1}{2} \hat{\epsilon}_x(L) = \infty$ 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 $\alpha$ and $\epsilon$. In light of Corollary 8.9 it is natural to ask if $\hat{\epsilon}$ has same formal properties shared by $\alpha$ and $\epsilon$ (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 $k$, $x \in X(\overline{k})$ and $X_1, \ldots, X_r$ the irreducible components passing through $x$ then we set $\hat{\epsilon}_{x,X}(L) = \min(\hat{\epsilon}_{x,X_1}(L|_{X_1}), \ldots, \hat{\epsilon}_{x,X_r}(L|_{X_r}))$.

Proposition 8.10. Let $X$ be a projective variety defined over $k$, $x \in X(\overline{k})$, 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 $\mathbb{Q}$-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 $z \in Z$ 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

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

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

Then (a), (c), (d), (e), 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 $\overline{X}_1, \ldots, \overline{X}_r$ and $\overline{Y}_1, \ldots, \overline{Y}_s$ be the irreducible components of $X \times_k \overline{k}$ and $Y \times_k \overline{k}$ respectively. Note that all $\overline{X}_i$ and $\overline{Y}_j$ are isomorphic over $\overline{k}$, 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 $\epsilon$, each connected component of $V \times_k \overline{k}$ is an étale Galois cover of some $\overline{X}_i \times \overline{Y}_j$, and hence is a product of an étale Galois covers of $\overline{X}_i$ and $\overline{Y}_j$. These étale Galois covers of $\overline{X}_i$ and $\overline{Y}_j$ may be descended to Galois covers of $X$ and $Y$ respectively.)

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 limit 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 $\beta$.** 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 [11, 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 $\gg \gamma$ at $x$? Since the set of sections vanishing to order $\gg \gamma$ at $x$ forms a proper subspace $W_\gamma$ 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 $W_\gamma$, 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 $m \to \infty$ 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 $\gamma$.

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

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

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 [6, §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 $\beta_x(L) \geq \frac{n}{n+1} \sqrt{\frac{\text{Vol}(L)}{\text{mult}_{_X} X}} \geq \frac{n}{n+1} \epsilon_x(L)$, and we have used this to deduce approximation theorems involving $\epsilon$ from those involving $\beta$. If the inequalities are strict then replacing $\beta$ 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) $\beta_x(L) = \frac{n}{n+1} \sqrt{\frac{\text{Vol}(L)}{\text{mult}_{_X} X}}$
(b) $\sqrt{\frac{\text{Vol}(L)}{\text{mult}_X X}} = \epsilon_x(L)$
(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 $\beta_x = \beta_x(L)$, $\omega_x = \sqrt{\frac{\text{Vol}(L)}{\text{mult}_X X}}$, $\epsilon_x = \epsilon_x(L)$, 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:

1. $\text{Vol}(L_\gamma) = \text{Vol}(L) - (\text{mult}_X X) \gamma^n$ for $\gamma \in [0, \omega_x]$, and
2. $\gamma_{\text{eff}} = \omega_x$.

Here (as usual) $L_\gamma = \pi^* L - \gamma E$ and $\pi: \widetilde{X} \to X$ 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

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

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 [8, 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

$$\hat{h}^i(O_{\gamma E}) = \lim_{m \to \infty} \frac{\dim H^i(\widetilde{X}, O_{m\gamma E})}{m^n/n!}$$

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

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

$$\text{Vol}(L_\gamma) - \hat{h}^1(L_\gamma) = \text{Vol}(L) - \hat{h}^0(O_{\gamma E}) = \text{Vol}(L) - (\text{mult}_X X) \gamma^n.$$ 

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 $\widetilde{X}$. By [3, Theorem A] $L_\gamma$ 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 $\widetilde{X}$. Then $L_\gamma - tA = (1-t)\pi^* L - (\gamma - ts)E = (1-t)L_{\gamma - ts}$. The asymptotic cohomology
groups are homogeneous of degree \( n \), so
\[
\hat{h}^{i}(L_{\gamma} - tL_{x}) = \hat{h}^{i}\left((1 - t)L_{\frac{\gamma - t\gamma_{E}}{1 - t}}\right) = (1 - t)^{n}\hat{h}^{i}\left(L_{\frac{\gamma - t\gamma_{E}}{1 - t}}\right)
\]
for all \( i \geq 0 \). If \( 0 < \gamma < \omega_{x} \), then for small enough \( t \) we have \( 0 \leq \frac{\gamma - t\gamma_{E}}{1 - t} < \omega_{x} \) too, and hence by (9.1.a.1) and the equation above \( \hat{h}^{i}(L_{\gamma} - tA) = 0 \) for all \( i \geq 0 \).

Summarizing, condition (9.1.a.1) and Theorem A of [3] imply that \( L_{\gamma} \) is ample for all \( 0 < \gamma < \omega_{x} \). Thus \( \omega_{x} \leq \epsilon_{x} \). The opposite inequality, \( \epsilon_{x} \leq \omega_{x} \), is [10, 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) \( \text{Vol}(L_{\gamma}) = \text{Vol}(L) - (\text{mult}_{x} X)\gamma^{n} \) 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 \( \omega_{x} \) is the boundary of the effective cone, i.e.,

(9.1.b.2) \( \gamma_{\text{eff}} = \omega_{x} \).

Given these two conditions,
\[
\beta_{x} = \int_{0}^{\gamma_{\text{eff}}} \frac{\text{Vol}(L_{\gamma})}{\text{Vol}(L)} d\gamma = \int_{0}^{\omega_{x}} 1 - \frac{\text{mult}_{x} X}{\text{Vol}(L)} \frac{n}{n+1} \omega_{x} = \frac{n}{n+1} \epsilon_{x}.
\]
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 [10, Proposition 5.1.9] for any irreducible subvariety \( V \subset X \) of positive dimension passing through \( x \) we have the inequality

(9.1.1) \[
\epsilon_{x}(L) \leq \left( \frac{c_{1}(L)^{\dim V} \cdot V}{\text{mult}_{x} V} \right)^{\frac{1}{\dim V}},
\]
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 try and work out some formal properties of \( \beta_{x} \), in particular to ask whether the list of properties in Propositions 2.12 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 2.12 and 3.4.)

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

(a) \( \beta_{x}(mL) = m\beta_{x}(L) \).

(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 \( X_1, \ldots, X_r \) be the irreducible components containing \( x \). Then \( \beta_{x,X}(L) = \min(\beta_{x,X_1}(L|_{X_1}), \ldots, \beta_{x,X_r}(L|_{X_r})) \).

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

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

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 \( \beta_x \) may strictly decrease under restriction, recall that \( \beta_x(O_{\mathbb{P}^n}) = \frac{n}{n+1} \) 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 \( L = O_{\mathbb{P}^n}(1) \) then \( \beta_{x,Z}(L|_Z) < \beta_x(L) \).

\[\square\]

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

10. A special case of Vojta’s main conjecture

Vojta’s Main Conjecture (Conjecture 3.4.3 of [17]) 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 [17, §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, \( X/k \) an irreducible \( n \)-dimensional variety such that \( -K_X \) is ample, \( D \) a reduced and effective algebraic zero-cycle on \( X \), and \( S \) a finite set of places of \( k \). If \( \epsilon_x(-K_X) > \frac{n+1}{n} \) for every \( x \in \text{Supp}(D) \), 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:

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

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 \( \delta \) 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 the support of \( D \) 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

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

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

**Remark.** For $k$-points $x$ of $\text{Supp}(D)$ 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 [13]) 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 $x \in X$.

**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 Tressier*, J. Algebraic Geom. 18 (2009), no. 2, 279–308.

[3] de Fernex, T.; Küronya, A.; Lazarsfeld, R., *Higher cohomology of divisors on a projective variety*, Math. Ann. 337 (2007), no. 2, 443–455.

[4] Demailly, J.-P., *Singular Hermitian metrics on positive line bundles*, Complex Algebraic Varieties (Bayreuth, 1990), Lect. Notes in Math., vol 1507, 1992, 87–104.

[5] Dieudonné, J.; Grothendieck, A. *Éléments de Géométrie Algébrique. III. Étude cohomologique des faisceaux cohérents.*, Inst. Hautes Études Sci. Publ. Math., No. 11, 1961.

[6] Faltings, G.; Wüstholz, G., *Diophantine approximations on projective spaces*, Invent. Math. 116 (1994), 109–138.

[7] Hindry, M.; Silverman, J., *Diophantine geometry. An introduction.*, Graduate Texts in Mathematics, 201. Springer-Verlag, New York, 2000.

[8] Küronya, A., *Asymptotic cohomological functions on projective varieties*, Amer. J. Math. 128 (2006), no. 6, 1475–1519.

[9] Lang, S., *Diophantine Geometry*, Interscience Tracts in Pure and Applied Mathematics, No. 11, John Wiley & Sons, New York-London, 1962.

[10] Lazarsfeld, R., *Positivity in Algebraic Geometry I*, Springer-Verlag, 2004.

[11] Lazarsfeld, R.; Mustaţă, M.; *Convex bodies associated to linear series*, Ann. Sci. Éc. Norm. Supér. (4) 42 (2009), no. 5, 783–835.

[12] McKinnon, D., *A conjecture on rational approximations to rational points*, J. Algebraic Geom., 16 (2007), 257-303.

[13] McKinnon, D.; Roth, M., *An analogue of Liouville’s theorem and an application to cubic surfaces*, in preparation.

[14] Neukirch, J. *Algebraic Number Theory*, Springer-Verlag, 1999.

[15] Schmidt, W., *Diophantine Approximation*, Lecture Notes in Mathematics 785, Springer, 1980.

[16] Serre, J.P., *Lectures on the Mordell-Weil Theorem*, Vieweg, 1997.

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