Forum of Mathematics, Sigma (2023), **0**: 1–54
doi:10.1017/fms.2023.5

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Heights on stacks and a generalized Batyrev–Manin–Malle conjecture

Jordan S. Ellenberg[1], Matthew Satriano[2] and David Zureick-Brown[3]

[1]University of Wisconsin-Madison, WI, USA.
[2]University of Waterloo, Ontario, Canada.
[3]Emory University, GA, USA.

Received: 3 January 2022; **Revised:** 30 November 2022; **Accepted:** 25 December 2022

2020 Mathematics Subject Classification: 11G35, 11G50, 14G05, 14D23, 14G25, 14G40

Abstract

We define a notion of *height* for rational points with respect to a vector bundle on a proper algebraic stack with finite diagonal over a global field, which generalizes the usual notion for rational points on projective varieties. We explain how to compute this height for various stacks of interest (for instance: classifying stacks of finite groups, symmetric products of varieties, moduli stacks of abelian varieties, weighted projective spaces). In many cases, our uniform definition reproduces ways already in use for measuring the complexity of rational points, while in others it is something new. Finally, we formulate a conjecture about the number of rational points of bounded height (in our sense) on a stack $\mathcal{X}$, which specializes to the Batyrev–Manin conjecture when $\mathcal{X}$ is a scheme and to Malle's conjecture when $\mathcal{X}$ is the classifying stack of a finite group.

Contents

1. Introduction

Two subjects of central importance in arithmetic statistics are the enumeration of number fields of bounded discriminant (governed by *Malle's conjecture*) and the enumeration of rational points of bounded height on varieties (governed by the *Batyrev–Manin conjecture*).

More specifically, if G is a subgroup of S_n, denote by $N_G(B)$ the number of degree n number fields $K/\mathbf{Q}$ whose Galois closure has Galois group G and whose discriminant has absolute value at most B. Similarly, if X is a projective Fano variety, denote by $N_X(B)$ the number of rational points in $X(\mathbf{Q})$ whose height is at most B. Malle's conjecture predicts that $N_G(B)$ is asymptotic to $cB^{a(G)}(\log B)^{b(G)}$, where $a(G)$ and $b(G)$ are explicitly computable constants. The Batyrev–Manin conjecture predicts that $N_X(B)$ is asymptotic to $cB^{a(X)}(\log B)^{b(X)}$, where $a(X)$ and $b(X)$ are explicitly computable constants. (The prediction of c is much more delicate: see Peyre [56, Définition 2.1] for the Batyrev–Manin case, and Bhargava [10] for the Malle case, in the special case $G = S_n$. We make no attempt in the present paper to study the constants in our generalization of Batyrev–Manin–Malle, and we say only a bit about the powers of $\log B$; we confine our concrete predictions to the exponents a.)

The similarity between these two asymptotic predictions has not gone unremarked. The relation between the two conjectures becomes even closer upon making the observation that a Galois G-extension of $\mathbf{Q}$ actually *is* a rational point: not a rational point on a variety, but a rational point on an algebraic stack, in this case the classifying stack BG. It is thus natural to ask how one might formulate a conjecture about counting rational points of bounded height on a stack $\mathcal{X}$, which would specialize both to the Batyrev–Manin conjecture (when $\mathcal{X}$ is a Fano variety) and to Malle's conjecture (when $\mathcal{X}$ is the classifying stack of a finite group).

An obstacle appears immediately: There is no agreed-upon definition of the height of a rational point on a stack. The conventional definition of height, due to Weil, is a real-valued function on $X(\mathbf{Q})$, where X is a projective variety. It suffices to define height on $\mathbb{P}^n(\mathbf{Q})$ because, given the projective embedding $\iota\colon X \to \mathbb{P}^n$, we simply define $\mathrm{ht}_X(x)$ to be $\mathrm{ht}_{\mathbb{P}^n}(\iota(x))$ for every point $x \in X(\mathbf{Q})$. But a stack which, like BG, is *not* a scheme does not embed in projective space.

The goal of the present paper is to propose a definition of height for rational points on stacks over arbitrary global fields K, and, using this definition, to formulate a conjecture of Batyrev–Manin–Malle type for the number of rational points on a stack $\mathcal{X}$ of height at most B (under certain assumptions which guarantee this number is finite). Having made the definition, we find that our notion of height applies to many interesting stacks which are neither schemes nor classifying spaces of finite groups (e.g., weighted projective spaces, moduli spaces, symmetric powers of varieties). In many cases, our definition agrees with ad hoc notions of 'size' of a rational point which already appear in the literature.

We remark on some existing work concerning heights on stacks. One proposed definition for the height of a point on a Deligne–Mumford stack is given and used by Abramovich and Várilly-Alvarado in [2, 3, 1]; this notion of height is useful for moduli spaces but does not, for example, extend to an interesting height on BG. Beshaj, Gutierrez and Shaska [9] have a definition of height on weighted

projective space which agrees with ours in that case, as does the earlier preprint of Deng [23]. Starr and Xu [68, §1.4 of arXiv v1] have another definition whose relation to the one used in the present work is roughly that between the minimal slope in the Harder–Narasimhan filtration of a vector bundle and the slope of that vector bundle. And in very recent work, Nasserden and Xiao [54] offer an alternative definition for stacky curves, and Ratko Darda [20, Theorem 1.5.7.1] has proposed a definition for weighted projective stacks.

We have seen above that one cannot define the height of a rational point of a stack by imitating the standard definition for rational points on varieties. Before sketching our definition, we explain some further reasons for the difficulty of defining heights on stacks.

Failure of additivity

A central feature of the theory of heights on varieties is *additivity*. Given a proper variety X, we can define a height function $\mathrm{ht}_{\mathcal{L}}$ on $X(\mathbf{Q})$ corresponding to any line bundle $\mathcal{L}$ on X, and we have

$$\mathrm{ht}_{\mathcal{L}\otimes\mathcal{L}'}(x) = \mathrm{ht}_{\mathcal{L}}(x) + \mathrm{ht}_{\mathcal{L}'}(x) \tag{1.1}$$

for any pair $\mathcal{L}, \mathcal{L}'$ of line bundles on X and any $x \in X(\mathbf{Q})$.

It turns out there is no choice but to discard this useful feature when we extend the theory of heights to stacks. The following example shows why. Let $\mathcal{X} = B(\mathbf{Z}/2\mathbf{Z})$, and let $K = \mathbf{Q}$. A line bundle $\mathcal{L}$ on $\mathcal{X}$ is a representation of $\mathbf{Z}/2\mathbf{Z}$; we choose $\mathbf{C}$ to be the nontrivial one-dimensional representation. Then the tensor product of $\mathcal{L}$ with itself is the trivial line bundle; that is, $\mathcal{L}\otimes\mathcal{L} = \mathcal{O}$ in $\mathrm{Pic}(\mathcal{X})$. Thus, $\mathrm{ht}_{\mathcal{L}\otimes\mathcal{L}'}(x) = 0$ for all $x \in \mathcal{X}(\mathbf{Q})$. If our height functions satisfied equation (1.1), we would have $2\,\mathrm{ht}_{\mathcal{L}}(x) = 0$, and thus $\mathrm{ht}_{\mathcal{L}}$ would be identically 0, and thus uninteresting.[1]

Failure of valuative criterion of properness

Suppose $K = \mathbb{F}_q(t)$, and X_0/K is a projective variety. In this case, the height of a point $x \in X_0(K)$ has a very nice geometric interpretation. We may choose an projective integral model $X/\mathbb{P}^1$ whose generic fiber is X_0. By the valuative criterion of properness, we can extend x to a section $\overline{x}\colon \mathbb{P}^1 \to X$. Then the height of x is just the degree of the line bundle $\overline{x}^*\mathcal{O}_X(1)$ on $\mathbb{P}^1$. (Note that the height may depend on the choice of integral model.) When X is a proper stack instead of a projective scheme, the valuative criterion of properness does not allow us to 'spread out' a rational point in this fashion. For instance, an $\mathbb{F}_q(t)$-point of $B(\mathbf{Z}/2\mathbf{Z})$ is a quadratic extension of $\mathbb{F}_q(t)$. On the other hand, a map from $\mathbb{P}^1$ to $B(\mathbf{Z}/2\mathbf{Z})$ is an étale double cover of $\mathbb{P}^1$, which can only be the disjoint union of two copies of $\mathbb{P}^1$. In particular, the fiber of such a map over the generic point $\mathrm{Spec}\,\mathbb{F}_q(t)$ must correspond to the *trivial* quadratic extension $\mathbb{F}_q(t) \oplus \mathbb{F}_q(t)$.

Modification of Northcott property

A useful feature of the height on a variety X attached to an ample line bundle L is the *Northcott property*; the set of points x in $X(\overline{K})$ with $h_L(x) < B$ and which are defined over an extension K'/K of degree at most d is finite. We will often consider heights here which we want to consider 'positive' but which do not have this property. For example, when $x \in B(\mathbf{Z}/2\mathbf{Z})(K)$ is a point corresponding to an *everywhere unramified* G-extension of K, and L is a (the!) nontrivial line bundle on $B(\mathbf{Z}/2\mathbf{Z})$, we will see below that $h_L(x) = 0$. But there are infinitely many distinct degree-d extensions of $\mathbf{Q}$ which have everywhere unramified double covers, so the Northcott property cannot hold in its usual sense. What will typically be true, on the other hand, is that the heights of greatest interest to us will admit only finitely many points of bounded height over any *individual* global field. This is the notion of Northcott we will use in the present paper, though it does not quite follow the usual convention.

[1] One might suggest abandoning the requirement that height functions be real-valued instead of abandoning additivity. This feels like a bad idea to us: For one thing, if our goal is to count points of bounded height we want the target of the height function to carry a natural ordering.

Vector bundles

The usual height machine assigns a height function on $X(K)$ to any line bundle on X. For rational points on a stack $\mathcal{X}$, it turns out that this point of view is not quite sufficient for our purposes. Consider again the example of BG, where G is a finite group. The line bundles on BG are the one-dimensional representations of G; in particular, the line bundles only 'see' the abelianization of G, not all of G. When G is nonabelian, this turns out to imply that no height function coming from a line bundle on $\mathcal{X}$ can compute the discriminant of the G-extension L/K corresponding to a K-rational point. Rather, we need access to the entire representation theory of G, which is to say we need to study heights associated to vector bundles of higher rank on BG.

Our definitions of heights on stacks

We now sketch the main idea of our definition. Suppose K is a global field. If K is a function field, let C be the smooth projective curve with function field K; if K is a number field, let C be $\operatorname{Spec} \mathcal{O}_K$. Given a rational point $x\colon \operatorname{Spec} K \to \mathcal{X}$, we may not, as mentioned above, be able to extend x to a morphism from C to $\mathcal{X}$. However, it turns out that we *can* extend x to a map $\overline{x}\colon \mathcal{C} \to \mathcal{X}$, where $\mathcal{C}$ is a so-called *tuning stack* over C. When C is $\mathbb{P}^1/\mathbb{F}_q$, for instance, $\mathcal{C}$ is a 'stacky $\mathbb{P}^1$' which is generically isomorphic to $\mathbb{P}^1$ but has some points with nontrivial finite inertia groups. In general, the structure map $\pi\colon \mathcal{C} \to C$ will be a coarse moduli map.

Suppose $\mathcal{V}$ is a vector bundle on $\mathcal{X}$, which we take to be metrized at Archimedean places if K is a number field. Then $\overline{x}^*\mathcal{V}$ is a vector bundle on the tuning stack $\mathcal{C}$, and $\pi_*\overline{x}^*\mathcal{V}$ is a vector bundle on C, whose determinant is a line bundle on C. We now define

$$\mathrm{ht}_{\mathcal{V}}(x) = -\deg(\det(\pi_*\overline{x}^*\mathcal{V}^\vee)).$$

In the number field case, $-\det(\pi_*\overline{x}^*\mathcal{V}^\vee)$ is a *metrized* line bundle on C, and degree means Arakelov degree.

We note that the reason for the failure of additivity is now apparent: While the pullback $\overline{x}^*$ is compatible with tensor product of vector bundles, the pushforward π_* is not. Moreover, it really is crucial to include the push forward π_*; otherwise, line bundles on BG, which are all torsion in the Picard group, would all give trivial height functions!

In the Section 2, we define $\mathrm{ht}_{\mathcal{V}}$ rigorously and show that it does not depend on the choice of tuning stack. In Section 3, we compute several examples, which show that this notion captures arithmetic quantities of interest in many cases. In particular, we show that if

- G is a subgroup of S_n,
- $\mathcal{V}$ is the corresponding n-dimensional permutation representation of G,
- and x is a point of $BG(\mathbf{Q})$, corresponding to a degree-n extension $K/\mathbf{Q}$ whose Galois closure has Galois group G,

the height $\mathrm{ht}_{\mathcal{V}}(x)$ is precisely the discriminant of $K/\mathbf{Q}$; see Subsection 3.1. This realizes the goal of expressing the discriminant of a field extension as the height of a rational point on the classifying stack of a finite group.

We also work out in varying levels of detail several examples of natural stacks: stacks birational to $\mathbb{P}^1$, weighted projective spaces, symmetric powers of projective spaces and moduli stacks of abelian varieties.

Finally, we turn to conjectures about point-counting in Section 4. Using geometric intuition derived from the function field case, we propose a heuristic rate of growth for the function $N_{\mathcal{X},\mathcal{V}}(B)$, the number of rational points x of a stack $\mathcal{X}$ such that $\mathrm{ht}_{\mathcal{V}}(x) \leq B$. There is one further technical hurdle worthy of note in the introduction: in the case of the Batyrev–Manin conjecture for schemes X, the expected growth rate B^a is governed by the anticanonical height ht_{-K_X}; in the case of stacks, one cannot simply import the same formula since for many stacks of interest, for example, $\mathcal{X} = BG$, the anticanonical

bundle is trivial! Thus, we introduce a new function (see Definition 4.5) which replaces the anticanonical height function on stacks; it can be viewed as a suitable perturbation of the anticanonical height. Our point-counting conjecture 4.14 includes (the weak versions of) both the Batyrev–Manin conjecture and Malle's conjecture as special cases, but it makes many more predictions as well, which we hope will be the subject of future research.

1.1. Notation and conventions

Throughout this paper, we treat the arithmetic and the geometric settings in unison, letting C denote either $\operatorname{Spec} \mathcal{O}_K$ for a number field K or a smooth proper curve over a field k in which case we set $K = k(C)$. In the number field case, we implicitly assume that all vector bundles are metrized. Finally, if L/K is a finite extension of function fields corresponding to a map $f\colon C' \to C$, we let $\operatorname{disc}(L/K)$ be the degree of the ramification divisor.

2. Heights of rational points on stacks

Recalling our notation and conventions (Section 1.1), let K be either a number field or a function field of transcendence degree 1 over k. In the former case, let $C = \operatorname{Spec} \mathcal{O}_K$ and in the latter case, let C be the smooth proper curve over k with $K = k(C)$. Next, let $p\colon \mathcal{X} \to C$ be a normal proper Artin stack over C with finite diagonal. This implies by [18] that there is a coarse space morphism $q\colon \mathcal{X} \to X$.

A K-rational point $x \in \mathcal{X}(K)$ is a section

$$x\colon \operatorname{Spec} K \to \mathcal{X}$$

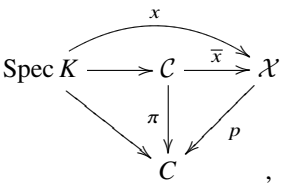
of p over the generic point $\eta := \operatorname{Spec} K$ of C, and an integral point is a section $\overline{x}\colon C \to \mathcal{X}$ of p. Now in the case of proper schemes, the valuative criterion tells us that every rational point extends uniquely to an integral point. However, this is no longer true for proper stacks; instead there exists a (possibly ramified) surjection $C' \to C$ such that the point $x'\colon \operatorname{Spec} k(C') \to \mathcal{X}$ extends to an integral point $C' \to \mathcal{X}$. It is precisely this phenomenon that leads to difficulties in defining heights on stacks.

Before discussing how to define heights of rational points on stacks, let us start by describing heights of integral points. This is actually rather simple and not different from the case of schemes. Given a vector bundle $\mathcal{V}$ on $\mathcal{X}$, we let the height $\operatorname{ht}_{\mathcal{V}}(\overline{x})$ of an integral point $\overline{x}\colon C \to \mathcal{X}$ be $-\deg\left(\overline{x}^*\mathcal{V}^\vee\right)$. (In the arithmetic setting, $\mathcal{V}$ is metrized, and we mean the Arakelov degree.) The notion of height of an integral point satisfies Weil's height machine, in that

$$\operatorname{ht}_{\mathcal{L}^{\otimes n}}(\overline{x}) = n \operatorname{ht}_{\mathcal{L}}(\overline{x})$$

for a line bundle $\mathcal{L}$. As mentioned above, for proper schemes there is no difference between rational points and integral points, so for schemes it is enough to define heights for integral points. For stacks, we must now deal with rational points that do not extend to integral points.

Let us now outline the general case of how we define heights of rational points on stacks. Given a rational point $x\colon C \dashrightarrow \mathcal{X}$, we know it extends to an integral point after allowing for a ramified extension of C. Unfortunately, there are many choices of such ramified extensions and so our first task is to construct a 'minimal' such extension; this extension is no longer a curve, but rather a stack, which we call a *tuning stack*. Precisely, we construct a commutative diagram



,

where $\pi\colon \mathcal{C} \to C$ is a birational coarse space map, and $\overline{x}\colon \mathcal{C} \to \mathcal{X}$ is a representable morphism of stacks which extends the rational point $x\colon \operatorname{Spec} K \to \mathcal{X}$. We can therefore think of $\mathcal{C}$ as being a 'stacky version' of C and can think of $\overline{x}$ as an integral point of $\mathcal{X}$. We then define the stable height of the rational point $x \in \mathcal{X}(K)$ with respect to $\mathcal{V}$ to be

$$\mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}(x) := -\deg(\overline{x}^*\mathcal{V}^\vee)$$

and define the unstable height (which we will refer to as simply the height) of the rational point $x \in \mathcal{X}(K)$ with respect to $\mathcal{V}$ to be

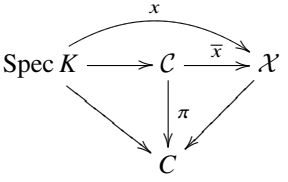
$$\mathrm{ht}_{\mathcal{V}}(x) := -\deg(\pi_*\overline{x}^*\mathcal{V}^\vee).$$

In Subsection 2.1, we show that tuning stacks exist and discuss their basic properties. We then turn to the study of heights in Subsection 2.2, and in Subsections 2.3 and 2.4 discuss some details of the practical computation of heights. In Appendix B, we gather technical facts about one dimensional normal Artin stacks with finite diagonal (i.e., the types of stacks that occur as tuning stacks).

2.1. Tuning stacks and tuning sheaves

Throughout we let K, C and $\mathcal{X}$ be as at the start of Section 2. Motivated by the *tuning module* of Yasuda–Wood [74, Definition 3.3], we begin by defining the central object of this subsection.

Definition 2.1. Given $x \in \mathcal{X}(K)$, we say that $(\mathcal{C}, \overline{x}, \pi)$ is a tuning stack for x if $\mathcal{C}$ is a normal Artin stack with finite diagonal, $\pi\colon \mathcal{C} \to C$ is a birational coarse space map, and the diagram



commutes. A morphism $(\mathcal{C}', \overline{x}', \pi') \to (\mathcal{C}, \overline{x}, \pi)$ of tuning stacks for x is a map $f\colon \mathcal{C}' \to \mathcal{C}$ such that $\pi \circ f = \pi'$ and $\overline{x} \circ f = \overline{x}'$. Finally, if $(\mathcal{C}, \overline{x}, \pi)$ is terminal among all tuning stacks, we say $\mathcal{C}$ is a universal tuning stack for x.

We show the existence of a universal tuning stack after some preliminaries.

Remark 2.2. Given a rational point $x\colon \operatorname{Spec} K \to \mathcal{X}$, there exists a nonempty open subset $U \subseteq C$ and a map $U \to \mathcal{X}$ over C that extends the morphism x. Since $\mathcal{X}$ is of finite type over C, this follows, for example, from [62, Proposition B.1].

Lemma 2.3. *Let $x \in \mathcal{X}(K)$, and suppose $(\mathcal{C}, \overline{x}, \pi)$ and $(\mathcal{C}', \overline{x}', \pi')$ are tuning stacks for x. Then the following hold.*

1. *If $\mathcal{C}' \overset{f}{\underset{g}{\rightrightarrows}} \mathcal{C}$ are two morphisms of tuning stacks, then f and g are isomorphic up to unique 2-isomorphism.*
2. *If $f\colon \mathcal{C}' \to \mathcal{C}$ is a representable morphism of tuning stacks, then f is an isomorphism.*
3. *If $\overline{x}$ and $\overline{x}'$ are representable, then any map $f\colon \mathcal{C}' \to \mathcal{C}$ of tuning stacks is an isomorphism.*

Proof. We start with (1). Since π and π' are birational, there is a nonempty open subset $U \subseteq C$ over which both π and π' are isomorphisms. Then $f|_U = g|_U$. Since $\mathcal{C}$ is normal and $\mathcal{C}'$ is separated, [29, Proposition A.1] tells us there is a unique 2-isomorphism $f \simeq g$.

We now turn to (2) and (3). Since $\overline{x}' = \overline{x} \circ f$, if $\overline{x}$ and $\overline{x}'$ are representable then [19, Corollary 2.2.7] shows f is also representable. Thus, (3) reduces to (2). To handle case (2), note that π and π' are

birational, proper, and quasi-finite, so f is as well. Then f is a birational, proper, quasi-finite morphism of normal stacks, hence an isomorphism by Zariski's main theorem. □

The next result makes use of relative normalization for morphisms of stacks. We refer the reader to [52, Definition 5.3].

Lemma 2.4. *Let $f\colon \mathcal{Y} \to \mathcal{Z}$ be a quasi-compact quasi-separated morphism of stacks with finite diagonal. Let $\mathcal{Y}' \to \mathcal{Z}$ be the relative normalization of f. If $\mathcal{Y}$ is normal, then $\mathcal{Y}'$ is normal.*

Proof. By definition of the relative normalization, f factors as $\mathcal{Y} \to \mathcal{Y}' := \underline{\mathrm{Spec}}_{\mathcal{Z}} \mathcal{O}' \to \mathcal{Z}$, where the sheaf $\mathcal{O}'$ is the integral closure of $\mathcal{O}_{\mathcal{Z}}$ in $f_*\mathcal{O}_{\mathcal{Y}}$ (i.e., the integral closure relative to the morphism of sheaves $\mathcal{O}_{\mathcal{Z}} \to f_*\mathcal{O}_{\mathcal{Y}}$ induced by the map f). Letting $Z \to \mathcal{Z}$ be a smooth cover by a scheme, we have a Cartesian diagram

$$\begin{array}{ccccc} \mathcal{W} & \longrightarrow & W' & \longrightarrow & Z \\ \downarrow & & \downarrow & & \downarrow \\ \mathcal{Y} & \longrightarrow & \mathcal{Y}' & \longrightarrow & \mathcal{Z}, \end{array}$$

where $\mathcal{W}$ may not be a scheme since we have not assumed f is representable. Since relative normalization commutes with smooth base change, $W' \to Z$ is the relative normalization of $\mathcal{W} \to Z$. Since $W' \to \mathcal{Y}'$ is a smooth cover, to show normality of $\mathcal{Y}'$ it suffices to prove W' is normal. We have therefore reduced to the case where $\mathcal{Z}$ is a scheme, which we will denote by Z.

We are now in the situation where $f\colon \mathcal{Y} \to Z$ and Z is a scheme. Notice that $\mathcal{Y}' \to Z$ is affine, and so $\mathcal{Y}' = Y'$ is a scheme. Since Z is a scheme, we know that f factors as $\mathcal{Y} \xrightarrow{\pi} Y \xrightarrow{g} Z$, where π is a coarse space map (which exists since $\mathcal{Y}$ has finite diagonal). By definition, $\mathcal{O}'$ is the integral closure of $\mathcal{O}_Z$ in $f_*\mathcal{O}_{\mathcal{Y}} = g_*\pi_*\mathcal{O}_{\mathcal{Y}} = g_*\mathcal{O}_Y$ where the last equality holds because π is Stein. Thus, $Y' \to Z$ is the relative normalization of $Y \to Z$. Since $\mathcal{Y}$ is normal, Y is as well so Y' is normal by [66, Tag 035L]. □

We are now ready to show the existence of universal tuning stacks. We thank Martin Olsson for suggesting this construction.

Proposition 2.5 (Universal tuning stacks exist). *Let $x \in \mathcal{X}(K)$. If $U \to \mathcal{X}$ is any extension of x as in Remark 2.2, then its relative normalization $\overline{x}\colon \mathcal{C} \to \mathcal{X}$ is a universal tuning stack, and it is independent of the choice of extension $U \to \mathcal{X}$.*

Proof. We abusively refer to the extended map $U \to \mathcal{X}$ as x. By definition of the relative normalization, x factors as

$$U \longrightarrow \mathcal{C} := \underline{\mathrm{Spec}}_{\mathcal{X}} \mathcal{O}' \xrightarrow{\overline{x}} \mathcal{X},$$

where the sheaf $\mathcal{O}'$ is the integral closure relative to the morphism of sheaves $\mathcal{O}_{\mathcal{X}} \to x_*\mathcal{O}_U$ induced by the map x. Lemma 2.4 shows that $\mathcal{C}$ is normal. Since $\overline{x}$ is representable, integral and of finite type, it follows from [66, Tag 01WJ] that it is finite. Then finiteness of the diagonal for $\mathcal{C}$ follows from finiteness of the diagonal for $\mathcal{X}$. Thus, $\mathcal{C}$ has a coarse space map $\pi\colon \mathcal{C} \to C'$. Since $\mathcal{C}$ is normal, C' is as well. The morphism $\overline{x}$ induces a map $q\colon C' \to C$.

We next show that $\mathcal{C} \to C$ is an isomorphism over U. Consider the Cartesian diagram

$$\begin{array}{ccccccc} U & \xrightarrow{\alpha} & \mathcal{C}_U & \xrightarrow{\beta} & \mathcal{X}_U & \xrightarrow{\gamma} & U \\ \| & & \downarrow & & \downarrow & & \downarrow \\ U & \longrightarrow & \mathcal{C} & \longrightarrow & \mathcal{X} & \longrightarrow & C. \end{array}$$

Since relative normalization commutes with smooth base change, $\beta\colon \mathcal{C}_U \to \mathcal{X}_U$ is the relative normalization of $\beta\circ\alpha\colon U \to \mathcal{X}_U$. Note that $\gamma\circ\beta\circ\alpha = \mathrm{id}_U$ is proper quasi-finite and $\gamma\colon \mathcal{X}_U \to U$ is separated, so $\beta\circ\alpha$ is proper quasi-finite, hence finite as it is representable. Thus, $\beta\alpha$ is integral so its relative normalization $\alpha\colon U \to \mathcal{C}_U$ is an isomorphism. As a result, $\gamma\circ\beta\colon \mathcal{C}_U \to U$ is an isomorphism.

Now that we have established $\mathcal{C} \to C$ is an isomorphism over U, it follows that $q\colon C' \to C$ is an isomorphism over U. So, q is a birational map of normal curves (or Dedekind schemes), hence an isomorphism. This shows that $\pi\colon \mathcal{C} \to C' \simeq C$ is a birational coarse space map, and hence $\mathcal{C}$ is a tuning stack.

Before turning to the claim concerning universality, we show that $\overline{x}\colon \mathcal{C} \to \mathcal{X}$ is independent of the choice of open subset U and extension $U \to \mathcal{X}$ of x. To see this, it suffices to show that if $i\colon V \to U$ is the inclusion of a nonempty open subset, then the relative normalizations of $x\colon U \to \mathcal{X}$ and $x\circ i\colon V \to \mathcal{X}$ are the same. Letting $\overline{x}\colon \mathcal{C} \to \mathcal{X}$ be the former normalization and $\overline{x}'\colon \mathcal{C}' \to \mathcal{X}$ be the latter one, by functoriality of the relative normalization we obtain a morphism $f\colon \mathcal{C}' \to \mathcal{C}$ of tuning stacks. Lemma 2.3 (3) shows f is an isomorphism.

To prove universality, let $(\mathcal{C}', \overline{x}', \pi')$ be another tuning stack. By Lemma 2.3 (1), we need only show the existence of a map $f\colon \mathcal{C}' \to \mathcal{C}$ of tuning stacks. We let $\mathcal{C}' \longrightarrow \mathcal{C}'' \xrightarrow{\overline{x}''} \mathcal{X}$ be the relative normalization of $\overline{x}'$. Since π and π' are birational, we can choose a nonempty open subset $U \subseteq C$ over which π and π' are isomorphisms. We have just showed that $\mathcal{C}$ is independent of the choice of U, so we have a commutative diagram

$$\begin{array}{ccccc} U & \longrightarrow & \mathcal{C}' & \longrightarrow & \mathcal{C}'' \\ \downarrow & & & \overset{g}{\nearrow} & \downarrow{\scriptstyle \overline{x}''} \\ \mathcal{C} & & \xrightarrow{\overline{x}} & & \mathcal{X}, \end{array}$$

where we obtain the morphism $g\colon \mathcal{C} \to \mathcal{C}''$ (shown as a dotted arrow above) from the universal property of the relative normalization of $x\colon U \to \mathcal{X}$. By Lemma 2.4, we know $\mathcal{C}''$ is normal. We also know that $\overline{x}''$ is representable, integral and of finite type, hence finite by [66, Tag 01WJ]. Then $\mathcal{C}''$ has finite diagonal, so it has a coarse space. Since π' is an isomorphism over U, we see $\mathcal{C}'' \to C$ is a coarse space which is an isomorphism over U; this follows from the same argument used to establish this fact for $\mathcal{C} \to C$. So, $\mathcal{C}''$ is a tuning stack for x. Finally, Lemma 2.3 (3) shows that g is an isomorphism, and so $\mathcal{C}' \longrightarrow \mathcal{C}'' \xrightarrow{g^{-1}} \mathcal{C}$ is our desired map of tuning stacks. □

Corollary 2.6. *Let $(\mathcal{C}', \overline{x}', \pi')$ be a tuning stack. Then $(\mathcal{C}', \overline{x}', \pi')$ is a universal tuning stack if and only if $\overline{x}'$ is representable.*

Proof. Let $(\mathcal{C}, \overline{x}, \pi)$ be the universal tuning stack constructed in Proposition 2.5. By construction, $\overline{x}$ is representable. Now, if $(\mathcal{C}', \overline{x}', \pi')$ is a universal tuning stack, by definition of universality, there is an isomorphism $f\colon \mathcal{C}' \to \mathcal{C}$ of tuning stacks. Then $\overline{x}' = \overline{x}\circ f$ shows that $\overline{x}'$ is representable.

Conversely, if $(\mathcal{C}', \overline{x}', \pi')$ is a tuning stack, then by universality of $\mathcal{C}$, we have a morphism $f\colon \mathcal{C}' \to \mathcal{C}$ of tuning stacks. The result then follows from Lemma 2.3 (3). □

Remark 2.7. We note that the universal tuning stack $\mathcal{C}$ inherits many properties of $\mathcal{X}$. For instance, if $\mathcal{X}$ is Deligne–Mumford, then so is $\mathcal{C}$ (since the map $\mathcal{C} \to \mathcal{X}$ is representable); similarly, $\mathcal{C}$ is separated.

Example 2.8 (Root stacks). Cadman [17, Section 2] introduced the notion of a root stack, which we will use repeatedly both in examples and in proofs. Given an algebraic stack Y and an effective Cartier divisor E on Y, the root stack $\widetilde{Y} \to Y$ of order r is obtained by formally adjoining an rth root $\widetilde{E}$ of E; in other words, for a scheme T and a map $f\colon T \to Y$, a lift of f to $\widetilde{Y}$ corresponds to an effective Cartier divisor E' on T and an equivalence $rE' \sim f^*E$.

Remark 2.9. Not every tuning stack is universal. For example, given any tuning stack $(\mathcal{C}, \overline{x}, \pi)$ and a smooth nonstacky closed point P of $\mathcal{C}$, let $f\colon \mathcal{C}' \to \mathcal{C}$ be a root stack along P; then f is an isomorphism

away from P and the composite $\overline{x} \circ f \colon \mathcal{C}' \to \mathcal{X}$ is not representable. So Corollary 2.6 shows that $(\mathcal{C}', \overline{x} \circ f, \pi \circ f)$ is a tuning stack which is not universal.

Occasionally, we will need to work with the universal tuning stack itself, for example, in Section 4 where we define the essential deformation dimension. However, we prove in Proposition 2.13 that our notion of height is independent of the choice of tuning stack. In practice, it is frequently more convenient to construct a tuning stack via a more direct procedure than relative normalization, such as taking a quotient stack, or as a root stack; see Section 3 for examples.

Definition 2.10. Let $\mathcal{V}$ be a vector bundle on $\mathcal{X}$. If $x \in \mathcal{X}(K)$ and $(\mathcal{C}, \overline{x}, \pi)$ is a choice of tuning stack, then we refer to $\pi_* \overline{x}^* \mathcal{V}^\vee$ (which is a vector bundle by Corollary B.4) as the tuning sheaf associated to x, $\mathcal{V}$, and $\mathcal{C}$.

2.2. Heights

We are now ready to give the definition of the height of a rational point on a stack (with respect to a given vector bundle). We define the height to be the degree of any associated tuning sheaf. The tuning sheaf is, in general, a vector bundle, so by degree we mean the degree of the top wedge power, which is now a line bundle (metrized in the arithmetic case) on C. We show that this is well defined in Proposition 2.13.

Definition 2.11. Let $\mathcal{X}$ be a stack over C, and let $K = K(C)$. Let $\mathcal{V}$ be a vector bundle on $\mathcal{X}$ and $x \in \mathcal{X}(K)$ be a rational section. If $\mathcal{C}$ is any tuning stack for x and $\mathcal{T}_{x,\mathcal{V},\mathcal{C}}$ is the associated tuning sheaf, we let $\mathrm{ht}_{\mathcal{V}}(x) = -\deg(\mathcal{T}_{x,\mathcal{V},\mathcal{C}})$. In other words, the height of the rational point $x \in \mathcal{X}(K)$ with respect to $\mathcal{V}$ is

$$\mathrm{ht}_{\mathcal{V}}(x) := -\deg(\pi_* \overline{x}^* \mathcal{V}^\vee),$$

where $(\mathcal{C}, \overline{x}, \pi)$ is any choice of tuning stack for x.

If L is a finite extension of K, we can define the height of a point of $\mathcal{X}(L)$ by letting C' be $\operatorname{Spec} \mathcal{O}_L$ (if K is a number field) or the smooth projective curve with function field L (if K is a function field), and consider $\mathcal{X}' = \mathcal{X} \times_C C'$, which carries a vector bundle obtained by pulling back $\mathcal{V}$. Then we define the height of a point of $\mathcal{X}(L)$ to be the height of the corresponding point of $\mathcal{X}'(L)$.

At this point, we need to comment on a piece of notation. When C is a curve over a finite field k, the degree of a divisor $D = P_1 + \cdots + P_r$ on C is understood to be $\sum_i \log |k_{P_i}|$, where k_{P_i} is the residue field of the closed point P_i. In particular, $\deg D$ does not lie in $\mathbf{Z}$ but in $(\log q)\mathbf{Z}$, where $q = |k|$. This choice of notation is most natural in a context, as here, where we want to write down theorem statements and arguments which treat the case of number fields and function fields at once. The reader who wants to work in the context where C is a curve over a fixed finite field k and avoid the number field case is free to take heights to be integers, which just modifies everything in this paper by a multiplicative factor of $\log q$.

The reader may wonder why the height is defined as the negative of the degree of a bundle obtained from $\mathcal{V}^\vee$, rather than as the degree of a bundle obtained from $\mathcal{V}$ itself. The answer is that, in cases arising naturally, the heights as defined here will typically be bounded below (Northcott property) while a height defined to be $\deg(\pi_* \overline{x}^* \mathcal{V})$ will often take values unbounded both above and below, or only bounded above (Southcott property).

Another natural question: Why do we not define the height of x as $\deg_{\mathcal{C}} \overline{x}^* \mathcal{V}$ (where degree is defined in Definition B.5), which would be more similar to the usual definition? The main reason is that, as we shall see, $\deg_{\mathcal{C}} \overline{x}^* \mathcal{V}$ is identically zero for many choices of $\mathcal{X}$ and nontrivial $\mathcal{V}$ (e.g., for any line bundle on BG). Nonetheless, this function will play a key role for us (it will differ from $\mathrm{ht}_{\mathcal{V}}(x)$ by local terms supported on the stacky locus of $\mathcal{C}$, as we will see in §2.3), so we give it a name here.

Definition 2.12. Let $\mathcal{X}$, $\mathcal{V}$ and K be as in Definition 2.11. Then stable height $\mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}(x)$ is defined by

$$\mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}(x) = -\deg_{\mathcal{C}} \overline{x}^* \mathcal{V}^\vee$$

for any choice of tuning stack $\mathcal{C}$.

We justify the name 'stable height' in Proposition 2.14 below. When x is an *integral* point of $\mathcal{X}$, we may take C itself to be the tuning stack; in this case, π is the identity and $\mathrm{ht}(x)$ and $\mathrm{ht}^{\mathrm{st}}(x)$ agree.

Proposition 2.13 (Height and stable height are independent of tuning stack). *If $(\mathcal{C}_1, \overline{x}_1, \pi_1)$ and $(\mathcal{C}_2, \overline{x}_2, \pi_2)$ are two choices of tuning stacks for $x \in \mathcal{X}(K)$, then $-\deg(\pi_{1*}\overline{x}_1^*\mathcal{V}^\vee) = -\deg(\pi_{2*}\overline{x}_2^*\mathcal{V}^\vee)$ and $-\deg(\overline{x}_1^*\mathcal{V}^\vee) = -\deg(\overline{x}_2^*\mathcal{V}^\vee)$ for all vector bundles $\mathcal{V}$ on $\mathcal{X}$.*

In fact we show more: Not only the height, but the isomorphism class of the tuning sheaf is independent of the choice of tuning stack.

Proof. Let $(\mathcal{C}, \overline{x}, \pi)$ be the universal tuning stack for x whose existence we have shown in Proposition 2.5. By the universal property, there exist unique morphisms $f_i \colon \mathcal{C}_i \to \mathcal{C}$ of tuning stacks. Thus, we reduce immediately to the case where $\mathcal{C}_1$ is universal and $f : \mathcal{C}_2 \to \mathcal{C}_1$ is a map of tuning stacks. Now, let

$$\mathcal{C}_2 \to \underline{\mathrm{Spec}}\, f_*\mathcal{O}_{\mathcal{C}_2} \to \mathcal{C}_1$$

be the Stein factorization. Then $\mathrm{Spec}\, f_*\mathcal{O}_{\mathcal{C}_2} \to \mathcal{C}_1$ is a birational, finite, representable map with normal codomain and hence an isomorphism by Zariski's main theorem.

In particular f is Stein (i.e. the map $\mathcal{O}_{\mathcal{C}_1} \to f_*\mathcal{O}_{\mathcal{C}_2}$ is an isomorphism). Then for any vector bundle $\mathcal{W}$ on $\mathcal{C}_1$,

$$\mathcal{W} \simeq \mathcal{O}_{\mathcal{C}_1} \otimes_{\mathcal{O}_{\mathcal{C}_1}} \mathcal{W} \simeq f_*\mathcal{O}_{\mathcal{C}_2} \otimes_{\mathcal{O}_{\mathcal{C}_1}} \mathcal{W} \simeq f_*f^*\mathcal{W},$$

where the third isomorphism is the projection formula. Applying π_{1*} to the above isomorphism with $\mathcal{W} = \overline{x}_1^*\mathcal{V}^\vee$, we see $\pi_{1*}\overline{x}_1^*\mathcal{V}^\vee \simeq \pi_{2*}\overline{x}_2^*\mathcal{V}^\vee$ and so height is independent of the choice of tuning stack. (In the Arakelov case, we note that the tuning stacks are all birational so that the metric does not change.) Independence of the stable height follows from Lemma B.9 applied to f_i. □

The justification for the name 'stable height' is as follows. As we shall see, the height $\mathrm{ht}_{\mathcal{V}}(x)$ does not behave well under ramified base change. That is: If L/K is a finite extension, and x_L the point of $\mathcal{X}(L)$ obtained by composing $x \colon \mathrm{Spec}\, K \to \mathcal{X}$ with the structure map $p \colon \mathrm{Spec}\, L \to \mathrm{Spec}\, K$, the relationship between $\mathrm{ht}_{\mathcal{V}}(x)$ and $\mathrm{ht}_{\mathcal{V}}(x_L)$ is not in general very transparent. For example, if $\mathcal{X} = BG$ and $x \in \mathcal{X}(K)$ corresponds to a Galois extension L/K with Galois group G, then $\mathrm{ht}_{\mathcal{V}}(x_L) = 0$, but $\mathrm{ht}_{\mathcal{V}}(x) \neq 0$ in general. For stable height, by contrast, the situation is much as we are used to from heights on schemes.

Proposition 2.14 (Stable height is stable under base change). *With $\mathcal{X}$, $\mathcal{V}$, x and x_L as above, and L/K is a separable extension, then*

$$\mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}(x_L) = [L : K]\, \mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}(x).$$

Proof. If L is a number field, then let $C' = \mathrm{Spec}\, \mathcal{O}_L$; if L is a function field, then let C' be the projective normal curve with function field L. Let $\mathcal{C}$ be a tuning stack for x_K. Then the normalization $\mathcal{C}'$ of the fiber product $\mathcal{C} \times_C C'$ is a tuning stack for x_L, and we compute that

$$\mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}(x_L) := \deg \overline{x}_L^*\mathcal{V} = \deg g \cdot \deg \overline{x}^*\mathcal{V} = [L : K]\, \mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}(x),$$

where g is the projection $\mathcal{C}' \to \mathcal{C}$ and the middle inequality is Lemma B.9. □

When $\mathcal{X}$ is a scheme, we can take $\mathcal{C} = C$ and so stable height and height are the same. More generally, height agrees with stable height whenever the vector bundle $\mathcal{V}$ is pulled back from a vector bundle on a scheme.

Proposition 2.15. *Suppose $f \colon \mathcal{X} \to Y$ is a morphism over C, where Y is a scheme. Let V be a vector bundle on Y. Then, for all $x \in X(K)$,*

$$\mathrm{ht}_{f^*V}(x) = \mathrm{ht}^{\mathrm{st}}_{f^*V}(x).$$

Proof. Let $\mathcal{C}$ be a tuning stack for x, and let $\overline{x}\colon \mathcal{C}\to\mathcal{X}$ be an extension of x to $\mathcal{C}$. The map $f\circ\overline{x}\colon \mathcal{C}\to Y$ factors as $g\circ\pi$ for some $g\colon C\to Y$, by the universal property of the coarse space. So the vector bundle $\overline{x}^*f^*V$ can be written as π^*g^*V. Noting that duality commutes with pullback, we now have

$$\mathrm{ht}_{f^*V}(x) = -\deg_C \pi_*\pi^*g^*V^\vee$$

and

$$\mathrm{ht}^{\mathrm{st}}_{f^*V}(x) = -\deg_{\mathcal{C}} \pi^*g^*V^\vee = -\deg_C g^*V^\vee$$

(where the last equality follows from Lemma B.9 since $\deg\pi = 1$). The result now follows from the fact that for any bundle W on C,

$$W \cong \mathcal{O}_C\otimes_{\mathcal{O}_C} W \cong \pi_*\mathcal{O}_{\mathcal{C}}\otimes_{\mathcal{O}_C} W \cong \pi_*\pi^*W;$$

the last isomorphism is the projection formula, and the middle follows since the coarse map is Stein [61, Theorem 6.12]. □

Remark 2.16. Similarly, if $f\colon \mathcal{X}\to\mathcal{Y}$ is a morphism of stacks and $\mathcal{V}$ is a vector bundle on $\mathcal{Y}$, then for any $x\in\mathcal{X}(K)$,

$$\mathrm{ht}_{f^*\mathcal{V}}(x) = \mathrm{ht}_{\mathcal{V}}(f\circ x)$$

since a tuning stack for x is also a tuning stack for $f\circ x$.

Definition 2.17. We say that a vector bundle $\mathcal{V}$ on $\mathcal{X}/K$ satisfies the Northcott property if for every finite extension L/K and every integer B,

$$\{x\in\mathcal{X}(L)\colon \mathrm{ht}_{\mathcal{V}}(x)\le B\}$$

is finite.

This definition is slightly unsatisfactory because it will be too lenient for some choices of $\mathcal{X}$. For instance, if $\mathcal{X}$ is a curve of genus at least 2, it has finitely many points over every global field, so under this definition the Northcott property will be satisfied by *every* vector bundle. In the present paper, however, we will almost always be considering stacks $\mathcal{X}/K$ which have infinitely many K-rational points. Under such circumstances we expect $\mathcal{V}$ to satisfy the Northcott property if $\mathcal{V}$ is 'positive enough', which we demonstrate through several examples; see Section 3. (Be warned, however, that the Northcott vector bundles do not form a cone in any sense. For instance, it is possible that a line bundle $\mathcal{L}$ is Northcott but positive multiples $\mathcal{L}^{\otimes n}$ of it are not; the nontrivial line bundle on $B\mu_2$ has this property.) It is in order to ensure that natural examples exhibit the Northcott property that we use $\mathcal{V}^\vee$ rather than $\mathcal{V}$ when defining height.

Definition 2.18. Let $\mathcal{X}$, $\mathcal{V}$ and K be as in Definition 2.11, with $\mathcal{V}$ Northcott. We define the counting function associated to $\mathcal{V}$ and K to be

$$N_{\mathcal{V},K}(B) := \#\{x\in\mathcal{X}(K)\colon \mathrm{Ht}_{\mathcal{V}}(x)\le B\}.$$

Remark 2.19. In case $\mathcal{V}$ is a vector bundle of rank greater than 1, it would probably be better still to consider a definition of height which associates to x the tuning sheaf $\mathcal{T}_{x,\mathcal{V},\mathcal{C}}$ *itself*, rather than its degree. One might call such a height a "lattice height." For instance, the lattice height of a $\mathbf{Q}$-point on $\mathcal{X}$ would be a lattice Λ in $\mathbb{R}^{\mathrm{rank}\,\mathcal{V}}$, rather than a real number; the height we study in the present paper would be the covolume of Λ. This point of view is interesting even when $\mathcal{X}$ is a scheme; see for instance the notion of *slopes of a rational point* introduced by Peyre in [57, §4.2] and [58], and the related work of Browning and Sawin in the Hardy–Littlewood regime [16]. On the other hand, when $\mathcal{X}$ is BG and $\mathcal{V}$

is a permutation representation $G \hookrightarrow S_n$, the lattice height of a rational point of $\mathcal{X}$ corresponding to a degree-n number field L/Q is the ring of integers of $\mathcal{O}_L$ considered as a lattice in $L \otimes_{\mathbf{Q}} \mathbf{R}$; the covolume of this lattice is the absolute value of the discriminant of the number field, which is indeed the height in the sense considered in this paper. This lattice is often called the 'shape' of the number field, and the problem of counting number fields subject to constraints on shape is already an area of substantial activity; see, for instance, [39, 38].

2.3. *Computing heights: local discrepancies*

We now turn to the problem of practical computation of heights of points on stacks.

As above, let C be the spectrum of the ring of integers of a number field or a smooth curve over a finite field, let K be the fraction field of C and let $\mathcal{X}$ a normal proper Artin stack over C with finite diagonal. Let $\mathcal{V}$ be a vector bundle on $\mathcal{X}$, where we recall once again that, if K is a number field, $\mathcal{V}$ is a metrized vector bundle, as defined in §A.4.

Let $x\colon \operatorname{Spec} K \to \mathcal{X}$ be a rational point of $\mathcal{X}$, let $\mathcal{C}$ be a tuning stack, $\pi\colon \mathcal{C} \to C$ the coarse moduli map and $\overline{x}\colon \mathcal{C} \to \mathcal{X}$ an integral extension of x.

By Definition 2.11, the height of x is

$$\mathrm{ht}_{\mathcal{V}}(x) = -\deg \pi_* \overline{x}^* \mathcal{V}^\vee,$$

and by Definition 2.12 we have

$$\mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}(x) = -\deg \overline{x}^* \mathcal{V}^\vee.$$

Our goal in this section is to study the *difference* between height and stable height. To this end, we recall the natural map of vector bundles on $\mathcal{C}$

$$\pi^* \pi_* \overline{x}^* \mathcal{V}^\vee \to \overline{x}^* \mathcal{V}^\vee \tag{2.20}$$

whose cokernel is a sheaf $M(\overline{x}^* \mathcal{V}^\vee)$ on $\mathcal{C}$ with trivial generic fiber. This map is the counit of adjunction and we claim that it is injective. Indeed, we can check injectivity locally and assume that C is affine, in which case $\pi_* \overline{x}^* \mathcal{V}^\vee = \Gamma(\overline{x}^* \mathcal{V}^\vee)$, and the map (2.20) is thus the inclusion

$$\Gamma(\overline{x}^* \mathcal{V}^\vee) \otimes_{\mathcal{O}_C} \mathcal{O}_{\mathcal{C}} \to \overline{x}^* \mathcal{V}^\vee$$

of global sections.

Let C' be a smooth proper curve (or in the arithmetic case, $\operatorname{Spec} \mathcal{O}_L$ for some étale algebra L/K) endowed with a finite flat surjection $p\colon C' \to \mathcal{C}$ whose degree we denote by m; such a C' exists by Proposition B.3. The sheaf $p^* M(\overline{x}^* \mathcal{V}^\vee)$ is now a generically trivial and finitely generated sheaf on C', which is to say it is a finite abelian group with the structure of an $\mathcal{O}_{C'}$-module. It follows from Proposition B.10 and exactness of p^* that

$$\begin{aligned} \log |p^* M(\overline{x}^* \mathcal{V}^\vee)| &= \deg p^* \overline{x}^* \mathcal{V}^\vee - \deg p^* \pi^* \pi_* \overline{x}^* \mathcal{V}^\vee \\ &= m(\deg \overline{x}^* \mathcal{V}^\vee - \deg \pi^* \pi_* \overline{x}^* \mathcal{V}^\vee) \\ &= m(\mathrm{ht}_{\mathcal{V}}(x) - \mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}(x)). \end{aligned}$$

Now, $p^* M(\overline{x}^* \mathcal{V}^\vee)$ is a finite $\mathcal{O}_{C'}$-module and as such has a canonical decomposition as a finite direct sum $\oplus_v p^* M(\overline{x}^* \mathcal{V}^\vee)_v$, where v varies over non-Archimedean places of C'.

Definition 2.21. With all notation as above, the local discrepancy $\delta_{\mathcal{V};v}$ is defined as

$$\delta_{\mathcal{V};v}(x) = \frac{1}{m} \log |p^* M(\overline{x}^* \mathcal{V}^\vee)_v|.$$

We thus arrive at the formula

$$\mathrm{ht}_{\mathcal{V}}(x) = \mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}(x) + \sum_{v} \delta_{\mathcal{V};v}(x). \tag{2.22}$$

One can think of the structural information imparted by equation (2.22) as follows. The height $\mathrm{ht}_{\mathcal{V}}$ is a nonadditive function which changes under field extensions and lacks a canonical decomposition into local terms. However, it canonically decomposes into two pieces; one of which, $\mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}$, is additive and stable under field extensions, while the other, $\sum_v \delta_{\mathcal{V};v}$, canonically decomposes into local terms. These good features of the summands often make it manageable to compute them individually.

Concretely, we may think of local discrepancy as follows. Write K_v for the completion of K at v. Define $L_v = K_v \otimes_C C'$ so that L_v is an étale algebra over K_v. We can thus write $C'_v = \mathrm{Spec}\, \mathcal{O}_{L_v}$. Choose an identification of $\overline{x}^*\mathcal{V}^\vee|_{\mathrm{Spec}\, K_v}$ with K_v^r. Then the generic stalk of $p^*\overline{x}^*\mathcal{V}^\vee$ is identified with L_v^r, and we can think of $p^*\overline{x}^*\mathcal{V}^\vee$ as a C'_v-lattice Λ in the vector space L_v^r. Then the $\mathcal{O}_{K_v}$-module $\pi_*\overline{x}^*\mathcal{V}^\vee$ is $\Lambda \cap K_v^r$, and so $p^*\pi^*\pi_*\overline{x}^*\mathcal{V}^\vee$ is

$$(\Lambda \cap K_v^r) \otimes_{\mathcal{O}_{K_v}} \mathcal{O}_{L_v} \subset \Lambda$$

and

$$\delta_{\mathcal{V};v}(x) = \frac{1}{m} \log \left| \frac{\Lambda}{(\Lambda \cap K_v^r) \otimes_{\mathcal{O}_{K_v}} \mathcal{O}_{L_v}} \right|.$$

Remark 2.23. One particularly illustrative example is when L_v is a degree d Galois extension of K_v with Galois group $G \subset S_d$, and $\overline{x}^*\mathcal{V}^\vee$ is the G-representation obtained from the permutation representation of S_d. In this case, Λ is the $\mathcal{O}_{L_v}$-module $\mathcal{O}_{L_v}^{\oplus d}$ and $\sigma \in G$ acts on the i-th basis vector e_i by $\sigma(e_i) = e_{\sigma(i)}$. Since Λ is G-linearized, it follows that $\sigma(\alpha e_i) = \sigma(\alpha) e_{\sigma(i)}$ for any $\alpha \in L_v$. Said another way, Λ is the G-linearized $\mathcal{O}_{L_v}$-module given by the skew group ring $G * \mathcal{O}_{L_v}$. If we label the elements of G by $\sigma_1, \ldots, \sigma_d \colon L_v \to L_v$, then we see $\Lambda \cap K_v^d = \Lambda^G$ is the set of sums of the form $\sum_i \sigma_i(\alpha) e_i$ with $\alpha \in L_v$. From this description, it is clear that the permutation representation is related to the discriminant. This relation will be further expanded upon in §3.1.

Proposition 2.24. *Let E_v be an unramified extension of K_v of degree d, let x be a point of $\mathcal{X}(K_v)$ and let x_E be the corresponding point of $\mathcal{X}(E_v)$. Then*

$$\delta_{\mathcal{V};v}(x_E) = d\delta_{\mathcal{V};v}(x).$$

Proof. (This proof is essentially the same as that of the 'geometric' part of [74, Lemma 3.4].)

Write Λ_E for $(\Lambda \otimes_{\mathcal{O}_{K_v}} \mathcal{O}_{E_v})$. Observe first that

$$\Lambda_E \cap E_v^r = (\Lambda \cap K_v^r) \otimes_{\mathcal{O}_{K_v}} \mathcal{O}_{E_v}$$

since the condition of being in K^r is cut out by K-linear conditions on L^r considered as a K-module; the same linear conditions applied to $(L \otimes_K E)^r$ cut out E^r. We then get an equality

$$d\delta_{\mathcal{V};v}(x) = d\frac{1}{m} \log \left| \frac{\Lambda}{(\Lambda \cap K_v^r) \otimes_{\mathcal{O}_{K_v}} \mathcal{O}_{L_v}} \right| = \frac{1}{m} \log \left| \frac{\Lambda_E}{(\Lambda \cap K_v^r) \otimes_{\mathcal{O}_{K_v}} \mathcal{O}_{L_v} \otimes_{\mathcal{O}_{K_v}} \mathcal{O}_{E_v}} \right|.$$

On the other hand, writing F_v for the etale algebra $E_v \otimes_{K_v} L_v$, we have

$$\delta_{\mathcal{V};v}(x_E) = \frac{1}{m} \log \left| \frac{\Lambda_E}{(\Lambda_E \cap E_v^r) \otimes_{\mathcal{O}_{E_v}} \mathcal{O}_{F_v}} \right| = \frac{1}{m} \log \left| \frac{\Lambda_E}{(\Lambda \cap K_v^r) \otimes_{\mathcal{O}_{K_v}} \mathcal{O}_{F_v}} \right|.$$

The desired equality now follows from the fact that, since E_v/K_v is unramified, we have $\mathcal{O}_{F_v} = \mathcal{O}_{L_v} \otimes_{\mathcal{O}_{K_v}} \mathcal{O}_{E_v}$. □

There will be some cases where our life is simpler if we can ignore a specified finite set of primes. The following proposition is useful when we need to show this negligence does not perturb our height functions by very much.

Proposition 2.25. *Suppose K is a number field. There is a constant $C(\mathcal{X}, \mathcal{V}, v)$ such that*

$$\delta_{\mathcal{V};v}(x) \leq C(\mathcal{X}, \mathcal{V}, v)$$

for all x in $\mathcal{X}(K_v)$.

Proof. (The following proof is adapted from a nice proof of Hilbert 90 that we learned from [8, Lemma 3.3].)

There is some constant B such that every point $x \in \mathcal{X}(K_v)$ extends to an integral point of $\mathcal{X}(L_v)$ for some finite Galois extension L of K of degree at most B; this follows from the fact that $\mathcal{X}$ has a finite cover by a scheme; see [62, Theorem B]. Since K is a number field, there are only finitely many isomorphism classes of extensions of K_v of degree at most B. We may thus prove the required bound for a single choice of L_v.

Write G for $\mathrm{Gal}(L/K)$. Write $\alpha_1, \ldots, \alpha_m$ for a subset of $\mathcal{O}_{L_v}$ which freely spans $\mathcal{O}_{L_v}$ as an $\mathcal{O}_{K_v}$-module. Let λ be an element of Λ, and for each i in $1, \ldots, m$ define

$$\lambda_i = \sum_{g \in G} (\alpha_i \lambda)^g.$$

The action of G permutes the summands above, so λ_i is fixed by G and thus lies in $\Lambda \cap K_v^r$.

We can also write

$$\lambda_i = \sum_{g \in G} (\alpha_i^g)(\lambda^g). \tag{2.26}$$

Write A for the $m \times m$ matrix in with rows indexed by $\alpha_1, \ldots, \alpha_m$ and columns by the elements of G; by Dedekind's lemma this matrix lies in $\mathrm{GL}_m(L_v)$. Write $\overrightarrow{\lambda}$ for the vector $\lambda_1, \ldots, \lambda_m \in L_v^m$ and $\overrightarrow{\mu}$ for the vector whose entries are $\{\lambda^g\}_{g \in G}$. With this notation, equation (2.26) becomes

$$\overrightarrow{\lambda} = A\overrightarrow{\mu}$$

which we can rewrite as

$$\overrightarrow{\mu} = A^{-1}\overrightarrow{\lambda}.$$

In particular, we can write

$$\lambda = \sum a_i \lambda_i, \tag{2.27}$$

where a_i are entries of A^{-1}. But note that A depends only on the choice of α_i; in particular, there is some constant C such that the entries of A^{-1} lie in $C^{-1}\mathcal{O}_{L_v}$. Thus, equation (2.27) expresses an arbitrary $\lambda \in \Lambda$ as a linear combination of the λ_i, which lie in $\Lambda \cap K_v^r$, with coefficients in $C^{-1}\mathcal{O}_{L_v}$. We conclude that

$$\Lambda \subset C^{-1}[(\Lambda \cap K_v^r) \otimes_{\mathcal{O}_{K_v}} \mathcal{O}_{L_v}]$$

which provides a bound for

$$\left| \frac{\Lambda}{(\Lambda \cap K_v^r) \otimes_{\mathcal{O}_{K_v}} \mathcal{O}_{L_v}} \right|$$

depending only on L_v, as required. □

We note that Proposition 2.25 does *not* hold in general when K has characteristic p. For instance, we will see that the local discrepancy at v for a point of BG, with $\mathcal{V}$ the regular representation of G, computes the discriminant of the local extension: but we know that the discriminant of a $\mathbf{Z}/p\mathbf{Z}$-extension of $\mathbb{F}_p((t))$ can be arbitrarily large, by contrast with the discriminant of a $\mathbf{Z}/p\mathbf{Z}$-extension of $\mathbf{Q}_p$.

2.4. Computing heights: line bundles on $\mathcal{X}$ with globally generated powers

In this section, we consider the special case where $\mathcal{V}$ is a line bundle $\mathcal{L}$. It turns out that, speaking loosely, if some tensor power $\mathcal{L}^m$ has "enough sections," we can use these sections to compute heights of rational points on $\mathcal{X}$ with little explicit reference to stacks. (Whether this is a virtue depends on the reader's taste.)

Suppose $\mathcal{X}$ is a stack over C, $\mathcal{L}$ is a metrized line bundle on $\mathcal{X}$, and $s_1, \dots, s_k$ are sections of $\mathcal{L}$. We say $\mathcal{L}$ is generically globally generated by $s_1, \dots, s_k$ if the cokernel $\mathcal{F}$ of the corresponding morphism of sheaves

$$\mathcal{O}_{\mathcal{X}}^{\oplus k} \to \mathcal{L}$$

vanishes over the generic point of C. In particular, this implies that $\mathcal{F}$ is supported at finitely many places v of C. More specifically: For each non-Archimedean v with uniformizer $\pi_v \in \mathcal{O}_{C_v}$, there is an integer m_v such that the restriction of $\mathcal{F}$ to $\mathcal{X} \times_C \mathcal{O}_{C_v}$ is killed by $\pi_v^{m_v}$ (since $\mathcal{X}$ is finite type, it suffices to check this on a finite flat cover). In the case where C has no Archimedean places, we say $\mathcal{L}$ is globally generated by $s_1, \dots, s_k$ when the map from $\mathcal{O}_{\mathcal{X}}^{\oplus k}$ to $\mathcal{L}$ is surjective. We write q_v for the order of the residue field at v, if v is a non-Archimedean place; when v is Archimedean, we can take $q_v = e$.

Proposition 2.28. *Suppose $\mathcal{X}$ is a stack, and suppose $\mathcal{L}$ is a metrized line bundle on $\mathcal{X}$ such that some power $\mathcal{L}^n$ is generically globally generated by sections $s_1, \dots, s_k$. Let K be a global field, and let $x\colon \operatorname{Spec} K \to \mathcal{X}$ be a point of $\mathcal{X}(K)$. Choose an identification of $x^*\mathcal{L}$ (whence also $x^*\mathcal{L}^n$) with K, and write $x_1, \dots, x_k$ for the pullbacks of $s_1, \dots, s_k$ by x. Then*

$$\mathrm{ht}_{\mathcal{L}}(x) = \sum_v \left\lceil (1/n) \log_{q_v} \max(|x_1|_v, \dots, |x_k|_v) \right\rceil \log q_v + E(x),$$

where $E(x)$ is a function bounded above and below on $\mathcal{X}(K)$. When C has no Archimedean places and $\mathcal{L}^n$ is globally generated by $s_1, \dots, s_k$ we have

$$\mathrm{ht}_{\mathcal{L}}(x) = \sum_v \left\lceil (1/n) \log_{q_v} \max(|x_1|_v, \dots, |x_k|_v) \right\rceil \log q_v$$

exactly.

From now on, we denote a bounded function on $\mathcal{X}(K)$ by $O_{\mathcal{X}(K)}(1)$. Note that, when $K = \mathbf{Q}$, we may take $x_1, \dots, x_k$ to be integers, with the property that, for every p, there is some x_i which is not a multiple of p^n. We say such a tuple $(x_1, \dots, x_k) \in \mathbf{Z}^k$ is in minimal form. Suppose $(x_1, \dots, x_k)$ corresponds to a point x of $\mathcal{X}(\mathbf{Q})$ as in Proposition 2.28. The hypothesis of minimal form implies that the non-Archimedean contributions all vanish, and we are left with

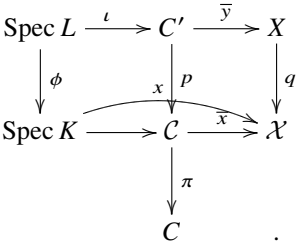
$$\mathrm{ht}_{\mathcal{L}}(x) = (1/n) \log \max_i |x_i|_{\mathbf{R}} + O_{\mathcal{X}(K)}(1) \tag{2.29}$$

up to a function bounded on $\mathcal{X}(\mathbf{Q})$. (The ceiling function can now be neglected since, having restricted to a single summand, the difference between a number and its floor is bounded and can be absorbed into the error term.)

We now prove Proposition 2.28.

Proof. We note, first of all, that we have not specified the choice of metric on $\mathcal{L}$ at Archimedean places, but this choice can be absorbed in the error term; if $\mathcal{L}$ and $\mathcal{L}'$ are line bundles which differ only with respect to the Archimedean metric, it is easy to see from the proof below that $\mathrm{ht}_{\mathcal{L}} - \mathrm{ht}_{\mathcal{L}'} = O_{\mathcal{X}(K)}(1)$. (At the moment when we say 'Fubini–Study metric on $\mathcal{O}(1)$ on complex projective space', just insert your own favorite metric, which differs from Fubini–Study by a bounded function.)

We begin by computing the degree of $\overline{x}^*\mathcal{L}^n$ on $\mathcal{C}$. Let L/K be a finite extension of some degree m such that the pullback of x to $\operatorname{Spec} L$ extends to a morphism $y\colon C' \to \mathcal{X}$, where C' is the curve (or Dedekind domain) with fraction field L. We then have a commutative diagram:



Now, $\deg_{\mathcal{C}} \overline{x}^*\mathcal{L}^n = (1/m)\deg_{C'} p^*\overline{x}^*\mathcal{L}^n$. The latter is a metrized line bundle on $\mathcal{O}_L$ whose degree we can compute by means of a section. For ease of notation, write Λ for $p^*\overline{x}^*\mathcal{L}^n$.

$$\deg_{C'} p^*\overline{x}^*\mathcal{L}^n = \log|\Lambda/s_1\mathcal{O}_L| - \sum_{\sigma\colon L\to\mathbf{C}} |\sigma^* s_1|_\sigma.$$

Write Λ' for the submodule of Λ spanned by $s_1, \dots, s_k$. By hypothesis, there is a bound independent of x for the size of Λ/Λ'. Thus, we may replace Λ with Λ' and get

$$\deg_{C'} p^*\overline{x}^*\mathcal{L}^n = \log|\Lambda'/s_1\mathcal{O}_L| - \sum_{\sigma\colon L\to\mathbf{C}} |\sigma^* s_1|_\sigma + O_{\mathcal{X}(K)}(1).$$

Now, the torsion $\mathcal{O}_L$-module $\Lambda'/s_1\mathcal{O}_L$ can be broken up into v-adic components T_v, one for each non-Archimedean place v of K, and by the explicit description of Λ' we have

$$\log|T_v| = m(\log\max_i |x_i|_v - \log|x_1|_v).$$

Thus, we have

$$\log|\Lambda'/s_1\mathcal{O}_L| = \sum_{v\nmid\infty} m(\log\max_i |x_i|_v - \log|x_1|_v).$$

We now turn to the Archimedean places, which requires us to specify the metric on $\mathcal{L}^n$. The sections $s_1, \dots, s_n$ provide a map of complex manifolds $f\colon \mathcal{X}(\mathbf{C}) \to \mathbb{P}^{k-1}(\mathbf{C})$, and $\mathcal{L}^n|\mathcal{X}(\mathbf{C})$ is pulled back from $\mathcal{O}(1)$ under f. So we may choose for our metric on $\mathcal{L}^n|\mathcal{X}(\mathbf{C})$ the pullback of the Fubini–Study metric on $\mathcal{O}(1)$. Having done so, we have

$$\sum_{\sigma\colon L\to\mathbf{C}} |\sigma^* s_1|_\sigma = \sum_{v|\infty} m(\log|x_1|_v - \log\max_i |x_i|_v) + O_{\mathcal{X}(K)}(1).$$

To sum up, we have computed that

$$\log|\Lambda'/s_1\mathcal{O}_L| - \sum_{\sigma\colon L\to C} |\sigma^* s_1|_\sigma = -\sum_v m|x_i|_v + \sum_v m\log\max_i |x_i|_v = \sum_v \log\max_i |x_i|_v + O_{\mathcal{X}(K)}(1),$$

whence

$$\deg_{C'} p^*\overline{x}^*\mathcal{L}^n = (\sum_v m \log \max_i |x_i|_v) + O_{\mathcal{X}(K)}(1),$$

whence

$$\deg_C \overline{x}^*\mathcal{L}^\vee = -(1/n)(\sum_v \log \max_i |x_i|_v) + O_{\mathcal{X}(K)}(1).$$

We note that, in the case where K is a function field and $s_1, \dots, s_k$ globally generate $\mathcal{L}^n$, the expression $(\sum_v m \log \max_i |x_i|_v)$ is just the usual expression for the degree of a line bundle pulled back from $\mathcal{O}(1)$ on $\mathbb{P}^{k-1}$ by a morphism with coordinates $(x_1 : \dots : x_k)$.

Having computed this degree, which is the negative of the stable height $\mathrm{ht}^{\mathrm{st}}_{\mathcal{L}}(x)$, we can compute $\mathrm{ht}_{\mathcal{L}}(x)$ by adding local discrepancies as in the previous section. First of all, if v is one of the finitely many non-Archimedean places where $\mathcal{L}^n$ is not generated by $s_1, \dots, s_k$, we observe that $\delta_{\mathcal{L};v}(x)$ is $O_{\mathcal{X}(K)}(1)$ by Proposition 2.25, and since the number of such places is bounded independently of x, we can absorb the contribution of those local discrepancies $\delta_{\mathcal{L};v}(x)$ into the error term.

So let v be a non-Archimedean place where $\mathcal{L}^n$ is generated by $s_1, \dots, s_k$. Then, given our choice of identification of $x^*\mathcal{L}^n$ with K, and writing L_v for the etale algebra $L \otimes_K K_v$, we can write $\overline{x}^*\mathcal{L}^n$ as the Galois-stable lattice I in L_v spanned as an $\mathcal{O}_{L_v}$-module by $x_1, \dots, x_k$. Then $\overline{x}^*\mathcal{L}^\vee$ is the submodule $I^{-1/n}$ of L_v consisting of all $\alpha \in L_v$ such that $\alpha^n I \subset \mathcal{O}L_v$. The pushforward $\pi_*\overline{x}^*\mathcal{L}^\vee$ is then the submodule $I^{-1/n} \cap K_v$ of K_v consisting of all $\beta \in K_v$ with $\beta^n x_i \subset \mathcal{O}_{K_v}$ for all i, which is to say it is the fractional ideal $m_v^{c_v}$, where

$$c_v = \lceil -(1/n) \min_i \mathrm{ord}_v\, x_i \rceil = \lceil (1/n) \log_{q_v} \max_i |x_i|_v \rceil.$$

So

$$\delta_{\mathcal{L};v}(x) = (1/m) \log |I^{-1/n}/I^{-1/n} \cap K| = (\log q_v) \lceil (1/n) \log_{q_v} \max_i |x_i|_v \rceil - (1/n) \log \max_i |x_i|_v.$$

Recalling from above that

$$\mathrm{ht}^{\mathrm{st}}_{\mathcal{L}}(x) = (1/n) \sum_v \log \max_i |x_i|_v + O_{\mathcal{X}(K)}(1),$$

we conclude that

$$\mathrm{ht}_{\mathcal{L}}(x) = \mathrm{ht}^{\mathrm{st}}(x) + \sum_v \delta_{\mathcal{L};v}(x) = \sum_v \lceil (1/n) \log_{q_v} \max_i |x_i|_v \rceil \log q_v + O_{\mathcal{X}(K)}(1)$$

which was the desired result. □

3. Examples

In this section, we show how to compute heights of points on various stacks that often arise in practice, emphasizing the fact that in these cases the output of our definition often recovers an invariant which was already widely used to measure the 'size' of the objects parametrized by rational points on those stacks.

3.1. Heights on BG

Let G be a constant finite group scheme over C, let $\mathcal{X}$ be the classifying stack BG/C, and let $q\colon C \to BG$ be the universal G-cover. Let $x\colon \operatorname{Spec} K \to \mathcal{X}$ be a rational point, and let $\overline{x}\colon \mathcal{C} \to BG$ be the extension of x to a tuning stack. Then we have a commutative diagram

$$\begin{array}{ccc} C' & \xrightarrow{x_{C'}} & C \\ \downarrow{\scriptstyle p} & & \downarrow{\scriptstyle q} \\ \mathcal{C} & \xrightarrow{\overline{x}} & BG, \end{array}$$

where C' is a smooth proper curve (not necessarily irreducible) whose fiber over $\operatorname{Spec} K$ is an étale G-algebra L/K.

Let $\mathcal{V}$ be a vector bundle of rank r on BG; in other words, $\mathcal{V}$ is an r-dimensional representation V of G over C. Then, by equation (2.22), we have

$$\operatorname{ht}_{\mathcal{V}}(x) = \operatorname{ht}^{\mathrm{st}}_{\mathcal{V}}(x) + \sum_v \delta_{\mathcal{V};v}(x).$$

First of all, note that $p^*\overline{x}^*\mathcal{V}^\vee = x_{C'}^* q^* \mathcal{V}$ is a vector bundle on C' pulled back from the trivial bundle on C, and thus has degree 0. So

$$\operatorname{ht}^{\mathrm{st}}_{\mathcal{V}}(x) = -\deg \overline{x}^*\mathcal{V}^\vee = -(\deg p)^{-1} \deg p^*\overline{x}^*\mathcal{V}^\vee = 0.$$

We have thus reduced ourselves to the local problem of computing $\delta_{\mathcal{V};v}(x)$ at the finite set of non-Archimedean places v of K where L/K is ramified. Let v be such a place.

The pullback of $\mathcal{V}^\vee$ along $x_{C'}^*$ from C to C' is $\mathcal{O}_{C'} \otimes_{\mathcal{O}_C} V^\vee$.

Thus, locally, the G-stable lattice $\Lambda_v \subset L_v^r$ we use to compute the local discrepancy can be written as

$$\mathcal{O}_{L_v} \otimes_{\mathcal{O}_{K_v}} V^\vee.$$

We note that this is precisely the G-module studied by Yasuda and Wood in section 3 of [74]. (The free rank $r\mathcal{O}_{L_v}$-module we call Λ_v is identified with $\mathcal{O}_{L_v}^r$ in their notation.) In particular, the free rank $r\mathcal{O}_{K_v}$-module Λ^G is precisely the *tuning submodule* in [74, Def 3.1], and the local discrepancy $\delta_{\mathcal{V};v}(x)$ is exactly the quantity denoted $\mathbf{v}_\tau(\rho)$ in [74, Def 3.3]. Thus, we can make use of their results to compute the local discrepancies explicitly.

The case where $\mathcal{V}$ is a permutation representation is an important example; in this case, we find that the discriminant of a field extension can be understood as a height on BG in the sense of this paper. In particular: When $\mathcal{V}$ is a degree-n permutation representation of G, and x is a point of $BG(K)$, we can associate to x a map

$$\rho_x\colon \operatorname{Gal}(K) \to G \to S_n$$

which in turn specifies a degree-n étale algebra L/K.

Proposition 3.1. *Let $\mathcal{V}$ be a vector bundle on BG corresponding to a degree-n permutation representation ρ of G, let x be a point in $BG(K)$ and let L/K be the algebra corresponding to x as described above. Then*

$$\operatorname{ht}_{\mathcal{V}}(x) = (1/2)\log|\Delta_{L/K}|. \tag{3.2}$$

Proof. It follows from [74, Theorem 4.8] that

$$\delta_{\mathcal{V};v}(x) = (1/2)a_v(\rho_x),$$

where a_v is the Artin conductor of $\rho_x|_{K_v}$, which is precisely the local component at v of $\Delta_{L/K}$. Thus,

$$\mathrm{ht}_{\mathcal{V}}(x) = \sum_v \delta_{\mathcal{V};v}(x) = (1/2)\log|\Delta_{L/K}|, \tag{3.3}$$

where by $|\Delta_{L/K}|$ we mean the absolute norm of the discriminant, that is, the order of the finite group $\mathcal{O}_C/\Delta_{L/K}$. □

In other words, the general definition of height introduced here, when applied to a G-extension (thought of as a point of BS_n), recovers the discriminant. Of course, a point of BS_n can be thought of as a G-extension in different ways; one might have in mind a degree-n extension, the Galois S_n-extension obtained by applying Galois closure or some other number field with the same Galois closure. Each such field corresponds to a permutation representation of S_n (in the first and second case above, the standard representation and the regular representation) and the discriminant of the field is computed by the height with respect to the vector bundle $\mathcal{V}$ specified by the corresponding permutation representation.

The case $\mathcal{X} = BG$ demonstrates the necessity of computing heights with respect to vector bundles of arbitrary rank, not only line bundles. Line bundles on BG correspond to one-dimensional representations of G. If, for example, G is a finite group with trivial abelianization, there are no nontrivial line bundles at all. In order to have a theory of heights rich enough to capture the invariants of G-extensions, we have no alternative than to consider vector bundles of higher rank on BG.

The work of Yasuda and Wood is not limited to permutation representations. For example, Wood and Yasuda work out in [74, Example 4.10] the example where $G = \mathbf{Z}/p\mathbf{Z}$, K is a function field of characteristic p and $\mathcal{V}$ is the two-dimensional nonsemisimple representation of $\mathbf{Z}/p\mathbf{Z}$ over K. A rational point of BG corresponds to a $\mathbf{Z}/p\mathbf{Z}$-extension L/K. If v is a place of K, we denote by j_v the largest integer i such that the higher ramification group G_i at v surjects onto $\mathbf{Z}/p\mathbf{Z}$. Then Yasuda and Wood's computation shows

$$\mathrm{ht}_v(x) = 1 + \left\lfloor \frac{j_v}{p} \right\rfloor. \tag{3.4}$$

When $K = \mathbb{F}_q(t)$ with q a power of p, the points of $B(\mathbf{Z}/p\mathbf{Z})(K)$ correspond to Artin–Schreier curves, and the height of an Artin–Schreier curve with respect to this $\mathcal{V}$ is the sum of the local terms (3.4) over all places v of $\mathbb{F}_q(t)$ which are ramified in the Artin–Schreier cover. We do not know if this notion of height of an Artin–Schreier curve corresponds to anything that has appeared in previous literature, but we note that the expression above is closely related to that appearing in the computation of dimensions of irreducible components of moduli space for Artin–Schreier curves of specified p-rank in the work of Pries and Zhu [60, Theorem 1.1].

This example also illustrates the important point that the height function $\mathrm{ht}_{\mathcal{V}}$ is *not* determined by the class of $\mathcal{V}$ in K_0 of the category of vector bundles; the vector bundle above is an extension of the trivial line bundle by the trivial line bundle, but its associated height function is not zero.[2]

3.2. *Heights on* $B\mu_n$

Suppose $\mathcal{X} = B\mu_n$, and $\mathcal{L}$ is the line bundle on $B\mu_n$ corresponding to the standard one-dimensional representation $\mu_n \to \mathbb{G}_m$. In that case, $\mathcal{L}^n$ is the trivial bundle on $\mathcal{X}$ and thus admits a generating section s. On the other hand, if x is a K-point of $B\mu_n$, the pullback $x^*\mathcal{L}$ is isomorphic to K. The obstruction to $x^*s \in \Gamma(\mathrm{Spec}\, K, x^*\mathcal{L}^n)$ being an nth power of an nonzero section of $x^*\mathcal{L}$ now yields a class in $K^*/(K^*)^n$. Put another way: Choosing an identification of $x^*\mathcal{L}$ with K induces an identification of $x^*\mathcal{L}^n$ with K, under which x^*s is identified with an element $x_0 \in K^*$, which represents the class in $K^*/(K^*)^n$ corresponding to x. Note that a change in the choice of s will apply a translation to the identification

[2]This is specifically due to the fact that $B(\mathbf{Z}/p\mathbf{Z})$ is not a tame stack over $\mathbb{F}_q(t)$, so π_* is not exact. Although $\overline{x}^*\mathcal{V}^\vee$ is an extension of $\mathcal{O}_C$ by itself, $\pi_*\overline{x}^*\mathcal{V}^\vee$ is no longer the extension of $\mathcal{O}_C$ by itself.

$B\mu_n(K) \cong K^*/(K^*)^n$, but such a change will modify heights by a bounded quantity, and if K is a function field over a finite field k and we require s to globally generate $\mathcal{L}^n$, the ambiguity in s imposes translation by k^*, which will not change the heights we compute at all. (If we want to remove this ambiguity entirely, we can fix for all time a choice of universal μ_n-torsor $q\colon \operatorname{Spec} K \to B\mu_n/K$ and an identification of $q^*\mathcal{L}$ with K; having done so, we can require that s pull back under q to an element of $(K^*)^n$.)

We note that the above setup applies even when $\operatorname{char} K$ divides n.

In particular: Proposition 2.28 yields

$$\mathrm{ht}_{\mathcal{L}}(x) = \sum_v \left\lceil \frac{1}{n} \log_{q_v} |x_0|_v \right\rceil \log q_v.$$

We note that our formula for $\mathrm{ht}_{\mathcal{L}}(x)$ is unchanged, as it must be, when x_0 is modified by an element of $(K^*)^n$.

By the computation above, when $K = \mathbf{Q}$ we see that the height of a point x of $B\mu_n(\mathbf{Q}) = \mathbf{Q}^\times/(\mathbf{Q}^\times)^n$ is obtained as follows: The class of $\mathbf{Q}^\times/(\mathbf{Q}^\times)^n$ corresponding to x is represented uniquely by an integer N with no nth power divisor, and as in equation (2.29) we have

$$\mathrm{ht}_{\mathcal{L}}(x) = \log |N|^{1/n}.$$

(In the examples, we will often suppress the $O_{\mathcal{X}(K)}(1)$ error term when no confusion is likely.)

Once again, the height recovers the measure of complexity most frequently used in practice; when enumerating the elements of $\mathbf{Q}^*/(\mathbf{Q}^*)^n$, one typically identifies the elements of the group with nth power-free integers and lists in order of absolute value.

Of course, this choice $\mathcal{L}$ is not the only option. Suppose, for instance, $K = \mathbf{Q}$ and $n = 3$; then there are two equally good choices of nontrivial line bundle on $\mathcal{X}$, namely $\mathcal{L}$ and $\mathcal{L}^2$. Suppose $x \in B\mu_3(\mathbf{Q})$ corresponds to $NM^2 \in \mathbf{Q}^\times/(\mathbf{Q}^\times)^3$, with N and M coprime and squarefree. Then, as we have already observed above,

$$\mathrm{ht}_{\mathcal{L}}(x) = \log |NM^2|^{1/3} = (1/3)\log N + (2/3)\log M.$$

On the other hand, consider $\mathcal{L}' = \mathcal{L}^2$. Then, having chosen s as above, s^2 is a generating section of $(\mathcal{L}')^3$, so we can take x_1 to be x^*s^2, which corresponds to $N^2M^4 \in \mathbf{Q}^\times$. Putting this integer in minimal form modifies it to N^2M, and another application of equation (2.29) shows that

$$\mathrm{ht}_{\mathcal{L}'}(x) = (2/3)\log N + (1/3)\log M.$$

As a final illustration, we can see how the above two computations combine to yield Proposition 3.1 for $B\mu_3$. Let $\mathcal{V}$ be the vector bundle $\mathcal{L} \oplus \mathcal{L}^2 \oplus \mathcal{O}_{\mathcal{X}}$. Then

$$\mathrm{ht}_{\mathcal{V}}(x) = \log N + \log M$$

which we note is also $(1/2)\Delta_{L/\mathbf{Q}}$, where $L = \mathbf{Q}((NM^2)^{1/3}) = \mathbf{Q}((N^2M)^{1/3})$ is the cubic extension of $\mathbf{Q}$ arising from x. This is as it must be, as we now explain. First, note that $\mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}(x) = 0$ for all x just as in the case $\mathcal{X} = BG$, because $\mathcal{V}$ pulls back to a trivial bundle on a finite cover of $B\mu_3$. So

$$\mathrm{ht}_{\mathcal{V}}(x) = \sum_v \delta_{\mathcal{V};v}(x).$$

Now, the size of $\delta_{\mathcal{V};3}(x)$ is bounded by Proposition 2.25, so at the expense of a bounded error term we can write

$$\mathrm{ht}_{\mathcal{V}}(x) = \sum_{v \neq 3} \delta_{\mathcal{V};v}(x).$$

Let x' be the point of $B(\mu_3)(\mathbf{Q}(\zeta_3))$ obtained by base change from x. Since every prime other than 3 is unramified in $\mathbf{Q}(\zeta_3)/\mathbf{Q}$, Proposition 2.24 tells us that

$$\mathrm{ht}_{\mathcal{V}}(x') = 2\sum_{v\neq 3}\delta_{\mathcal{V};v}(x) = 2\,\mathrm{ht}_{\mathcal{V}}(x).$$

On the other hand, over $\mathbf{Q}(\zeta_3)$, there is an isomorphism between $B(\mu_3)$ and $B(\mathbf{Z}/3\mathbf{Z})$, which carries $\mathcal{V}$ to the reduced permutation representation of $\mathbf{Z}/3\mathbf{Z}$, which we denote by $\mathcal{W}$. In fact, this isomorphism extends to $\mathbf{Z}[\zeta_3][1/3]$. Let y be the point of $B(\mathbf{Z}/3\mathbf{Z})(\mathbf{Q}(\zeta_3))$ corresponding to x' under this isomorphism, which we can also think of as the point associated to the Galois $\mathbf{Z}/3\mathbf{Z}$-extension $L(\zeta_3)/\mathbf{Q}(\zeta_3)$. Then

$$\delta_{\mathcal{W};v}(y) = \delta_{\mathcal{V}';v}(x')$$

for all places v of $\mathbf{Q}(\zeta_3)$ not dividing 3. We conclude that (as always, up to bounded error)

$$\mathrm{ht}_{\mathcal{W}}(y) = \sum_{v\neq 3}\delta_{\mathcal{W};v}(y) = \sum_{v\neq 3}\delta_{\mathcal{V}';v}(x') = 2\,\mathrm{ht}_{\mathcal{V}}(x).$$

On the other hand, by equation (3.3) we have

$$\mathrm{ht}_{\mathcal{W}}(y) = (1/2)\log|\Delta_{L(\zeta_3)/\mathbf{Q}(\zeta_3)}| = \log\Delta_{L/\mathbf{Q}}$$

which shows that $\mathrm{ht}_{\mathcal{V}}(x) = (1/2)\log|\Delta_{L/\mathbf{Q}}|$.

3.3. Heights on weighted projective space and weighted projective stacks

In this section, we consider rational points on the *weighted projective space* $\mathcal{X} = \mathbb{P}(a_0,\dots,a_k)$. This stack is, by definition, the quotient $[\mathbb{A}^{k+1}\setminus 0/\mathbb{G}_m]$ where $\mathbb{G}_m$ acts by the rule

$$\lambda\cdot(X_0,\dots,X_k) = (\lambda^{a_0}X_0,\dots,\lambda^{a_k}X_k).$$

Then $\mathbb{P}(a_0,\dots,a_k)$ is a smooth proper stack, and $\mathbb{A}^{k+1}\setminus 0$ is the total space of a line bundle on $\mathcal{X}$, whose dual is the *tautological bundle* $\mathcal{O}_{\mathbb{P}(a_0,\dots,a_k)}(1)$; for simplicity of notation, we denote the tautological bundle by $\mathcal{L}$ for the rest of this section. The coordinate function X_i is a section of $\mathcal{L}^{a_i}$. Writing A for the least common multiple of the a_i, the $k+1$ sections X_i^{A/a_i} of $\mathcal{L}^A$ generate $\mathcal{L}^A$. So we can compute heights of points in $\mathbb{P}(a_0,\dots,a_k)(K)$ by applying Proposition 2.28, as we now explain.

Let x be a point of $\mathbb{P}(a_0,\dots,a_k)(K)$. As in Proposition 2.28, we choose an identification of $x^*\mathcal{L}$ with K; this assigns a value in K to each of the $k+1$ coordinates, which values we denote $x_0,\dots,x_k$. Changing the identification of $x^*\mathcal{L}$ with K modifies this tuple by elementwise multiplication by tuples of the form $\lambda^{a_0},\dots,\lambda^{a_k}$, and we say that two tuples $x_0,\dots,x_k$ are equivalent if they differ by such a transformation. Then Proposition 2.28 tells us that

$$\mathrm{ht}_{\mathcal{L}}(x) = \sum_v \lceil \log_{q_v} \max_i |x_i|_v^{1/a_i}\rceil \log q_v. \tag{3.5}$$

In particular, when $K=\mathbf{Q}$, a rational point x of $\mathbb{P}^1(a_0,\dots,a_k)(\mathbf{Q})$ can be identified with a tuple of integers $(M_0:\dots:M_k)$ such that there is no prime p with $p^{a_i}|M_i$ for all i. Given a tuple which is in minimal form in this sense, the non-Archimedean primes contribute nothing to equation (3.5), and we get

$$\mathrm{ht}_{\mathcal{L}}(x) = \log\max_i |M_i|^{1/a_i}. \tag{3.6}$$

We note that this definition recovers the notion called 'naive height' for points of weighted projective space in [9].

Here is another means by which it is often practical to compute heights on weighted projective space when K is a global function field. Let F be a section of $\mathcal{L}^A$—for instance, it might be X_i^{A/a_i} for some i—and let y be the pullback of F along x to $x^*\mathcal{L}^A$, which we have identified with K. We define the *minimal valuation* of F at a place v of K as follows. Let $\pi_v \in K^*$ be an element which is a uniformizer at v, and define

$$c_v = \min\lfloor (1/a_i) \operatorname{ord}_v x_i \rfloor.$$

Note that $c_v = 0$ if and only if all the x_i are integral at v and there is some i such that $\operatorname{ord}_v x_i < a_i$. In this case, we say that $(x_0, \dots, x_k)$ is in minimal form. If $(x_0, \dots, x_k)$ is not in minimal form, we find an equivalent tuple in minimal form by modifying each x_i by $\pi_v^{-a_i c_v}$; the effect of this transformation on y is multiplication by $\pi_v^{-Ac_v}$. We therefore define the minimal valuation of F to be

$$\operatorname{ord}_v^{\min} F = \operatorname{ord}_v y - Ac_v = \operatorname{ord}_v y - A \min\lfloor (1/a_i) \operatorname{ord}_v x_i \rfloor.$$

We note that this quantity does not depend on the identification of $x^*\mathcal{L}$ with K, but only on F and v. Furthermore, we have

$$\sum_v \operatorname{ord}_v^{\min} F = \sum_v \operatorname{ord}_v y - \sum_v A \min\lfloor (1/a_i) \operatorname{ord}_v x_i . \rfloor = A \sum_v \max\lceil (1/a_i) \log_{q_v} \max |x_i|_v \rceil \log q_v$$

and, by Proposition 2.28, this last quantity, taking X_i^{A/a_i} to be the sections generating $\mathcal{L}^A$, is exactly $A \operatorname{ht}_{\mathcal{L}} x$. We conclude that

$$\operatorname{ht}_{\mathcal{L}} x = (1/A) \sum_v \operatorname{ord}_v^{\min} F \log q_v . \tag{3.7}$$

The classical theory of Weil heights is often set up by defining heights on projective spaces, and then defining a height $\operatorname{ht}_{\mathcal{O}(1)}$ on $X(K)$ for other projective schemes X by restriction. In a similar manner, one can define height functions on weighted projective stacks $\mathbb{P}(a_0, \dots, a_n)$ and obtain a height function $\operatorname{ht}_{\mathcal{L}}$ on $\mathcal{X}(K)$ whenever $\mathcal{L}$ is a generically globally generated power as in Section 2.4. However, we stress that this naive approach does *not* apply to all stacks of interest. Indeed, if $\mathcal{X}$ is any stack with a nonabelian stabilizer group, it does not embed into a weighted projective stack, hence the necessity of our construction of heights given in Section 2.2.

One example of weighted projective stacks which is of great interest is the moduli stack of elliptic curves $\overline{\mathcal{M}}_{1,1}$. If K is a field of characteristic not equal to 2 or 3, this stack is isomorphic over K to the weighted projective line $\mathbb{P}(4,6)$: concretely, given an elliptic curve E/K, we can write it in Weierstrass form $y^2 = x^3 + Ax + B$ with A, B in K. This Weierstrass form is unique up to transformations $(A, B) \to (\lambda^4 A, \lambda^6 B)$. So $(A : B)$ is a well-defined point on $\mathbb{P}(4,6)$. Moreover, the isomorphism takes the line bundle $\mathcal{O}(1)$ on $\mathbb{P}(4,6)$ to the Hodge bundle $\mathcal{L}$ on $\overline{\mathcal{M}}_{1,1}$ (the bundle whose kth powers have weight $2k$ modular forms as sections). We conclude that, if E/K is an elliptic curve over a global field of characteristic at least 5, with Weierstrass equation $y^2 = x^3 + Ax + B$, thought of as a K-point of $\overline{\mathcal{M}}_{1,1}$, we have

$$\operatorname{ht}_{\mathcal{L}} E = \log \max(|A|^{1/4}, |B|^{1/6}).$$

In other words, the familiar 'naive height' of an elliptic curve is indeed a height in the sense of this paper.

When K is a number field, the identification of $\overline{\mathcal{M}}_{1,1}/\mathbf{Q}$ with $\mathbb{P}(4,6)/\mathbf{Q}$ does not extend to $\operatorname{Spec} \mathbf{Z}$ but only to $\operatorname{Spec} \mathbf{Z}[1/6]$. However, this is enough to ensure that $\mathcal{L}^{12}$ is still *generically* globally generated by A^3 and B^2 in the sense of Proposition 2.28, and so equation (3.3) still holds up to a bounded error term.

When K is a global function field of characteristic at least 5, we can also apply equation (3.7); here, $A = \mathrm{lcm}(4, 6) = 12$ and the discriminant Δ is a natural section of $\mathcal{L}^{12}$ to use. So we find

$$\mathrm{ht}_{\mathcal{L}}\, E = \frac{1}{12} \sum_v \mathrm{ord}_v^{\min} \Delta, \tag{3.8}$$

where $\mathrm{ord}_v^{\min} \Delta$ is the valuation of the discriminant of a Weierstrass equation for E which is minimal at v.

We will return to the interesting case where K is a global function field of characteristic 2 or 3 in Section 3.4.

More generally, the moduli space of hyperelliptic curves over K with a marked Weierstrass point can be thought of as a weighted projective space as long as the characteristic of K is large enough: If $Y \to \mathbb{P}^1$ is the hyperelliptic map, we can move the image of the marked Weierstrass point to ∞ and (assuming the characteristic of K is not 2) complete the square in y so that the curve has affine equation

$$y^2 = x^{2g+1} + a_1 x^{2g} + \cdots + a_{2g+1}$$

then (again throwing out a finite set of characteristics for K) modify by the automorphism $x \to x + \frac{a_1}{2g+1}$ of $\mathbb{P}^1$ in order to make $a_1 = 0$. We now have an equation for Y of the form

$$y^2 = x^{2g+1} + a_2 x^{2g-1} + \cdots + a_{2g+1} \tag{3.9}$$

which is unique up to the operation of multiplying a_i by λ^{2i} for $\lambda \in K^*$. In other words, the moduli stack of hyperelliptic curves with marked Weierstrass point is isomorphic over K to the weighted projective $(2g-1)$-space $\mathbb{P}(4, 6, 8, \ldots, 4g+2)$. So a hyperelliptic curve over K can be thought of as a point x on $\mathbb{P}(4, 6, 8, \ldots, 4g+2)$, whose height with respect to $\mathcal{O}(1)$ we have computed above. In particular, if Y is a hyperelliptic curve over $\mathbf{Q}$ with Weierstrass equation (3.9), where the a_i are chosen to be integers so that there is no prime p with $p^{2i} | a_i$, the height of Y is $\log \max |a_i|^{1/2i}$, which again is equivalent to the notion of height typically used for hyperelliptic curves with a specified Weierstrass point as in, for example, the work of Bhargava and Gross [11].

Question 3.10. A weighted projective space is an example of a toric stack, as in [32]. What is the height of a rational point on a more general toric stack?

3.4. Heights of abelian varieties

We have established above in equation (3.3) that, when K is a global field of characteristic at least 5, the height of an elliptic curve with respect to the Hodge bundle on $\overline{\mathcal{M}}_{1,1}$ is the same as the customary naive height. There is another natural height on an elliptic curve over a global field: the *Faltings height* $\mathrm{ht}_{\mathrm{Fal}}(E)$. In this section, we study the extent to which Faltings height can be seen as a height in the sense of the present paper.

We note first that Faltings height satisfies some of the same formal properties as the heights defined in this paper do. For example: If L/K is a field extension, it is not necessarily the case that $\mathrm{ht}_{\mathrm{Fal}}(E/L)$ is $[L : K]\, \mathrm{ht}_{\mathrm{Fal}}(E/K)$; however, this equality *does* hold if E/K has everywhere semistable reduction, so we can define a stable Faltings height $\mathrm{ht}_s(E/K)$ to be $[L : K]^{-1}\, \mathrm{ht}_{\mathrm{Fal}}(E/L)$ for any L/K such that E/L has everywhere semistable reduction. The height $\mathrm{ht}_{\mathcal{V}}$ for any vector bundle on $\overline{\mathcal{M}}_{1,1}$ has the same properties, since an elliptic curve over $L = K(C')$ with everywhere semistable reduction is an integral point of $\overline{\mathcal{M}}_{1,1}$, that is, a morphism from C' to $\overline{\mathcal{M}}_{1,1}$. Lastly, $\mathrm{ht}_{\mathrm{Fal}}(E/K) - \mathrm{ht}_s(E/K)$ has a canonical local decomposition, just as does $\mathrm{ht}_{\mathcal{V}}(E/K) - \mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}(E/K)$; see (2.22).

It is thus natural to ask whether Faltings height is $\mathrm{ht}_{\mathcal{V}}$ for some vector bundle $\mathcal{V}$ or at least whether the two heights differ by a bounded function. One can even guess which vector bundle one might use; for everywhere semistable E/K, or in other words morphisms $f\colon C \to \overline{\mathcal{M}}_{1,1}$, we have

$$\mathrm{ht}_{\mathrm{Fal}}(E) = \deg f^*\mathcal{L},$$

where $\mathcal{L}$ is the Hodge bundle $\Omega^1_{\mathcal{E}/\overline{\mathcal{M}}_{1,1}}$ and $\mathcal{E}$ the universal semielliptic curve over the moduli stack.

So does $\mathrm{ht}_{\mathrm{Fal}}$ differ from $\mathrm{ht}_{\mathcal{L}}$ by a bounded function? Unfortunately, the answer is in general no—remember, in the number field case, $\mathrm{ht}_{\mathcal{L}}$ is naive height, and the difference between the naive height and the Faltings height of an elliptic curve over a number field K is *not* bounded, as one can see, for instance, in the proof of [55, Lemma 3.2].

The reason for this is the following. When K is a number field, the specification of the degree above requires a choice of metrization on $\mathcal{L}$ at the Archimedean places; for Faltings height, the appropriate Hermitian norm actually has a singularity at the cusp of moduli space, and in the present paper we have not considered metrized line bundles in this level of generality; rather, we have assumed that our choice of metrization on $\mathcal{L}$ is defined on all of $\overline{\mathcal{M}}_{1,1}(\mathbf{C})$, including the cusp.

However, when K is a global function field, this Archimedean issue is absent, and we find the following.

Proposition 3.11. *Let K be a global function field of characteristic at least 5, and let $\mathcal{L}$ be the Hodge bundle on $\overline{\mathcal{M}}_{1,1}$ as above. Then*

$$\mathrm{ht}_{\mathrm{Fal}}(E) = \mathrm{ht}_{\mathcal{L}}(E)$$

for all elliptic curves E/K.

Proof. For global function fields of characteristic larger than 3, the Faltings height of E/K is $(1/12)$ times the sum over all places of the valuation of a minimal discriminant; see for example, [6, Def 2.2]. We have already seen in equation (3.8) that $\mathrm{ht}_{\mathcal{L}}(E)$ is given by the same expression. □

The case of small characteristic is a different story. Let K be the function field of a curve C in characteristic p. Then the Faltings height of an elliptic curve over K is still the valuation of a minimal discriminant divisor on C, even if the characteristic of K is 2 or 3, and the Faltings height has the Northcott property.[3]

On the other hand, $\mathrm{ht}_{\mathcal{L}}$ is *not* Northcott in this setting, Note for instance that $\overline{\mathcal{M}}_{1,1}/\mathbf{F}_3$ contains as a closed substack a copy of BG lying over the coarse point $j = 0 = 1728$, where G is the automorphism group scheme of an elliptic curve with j-invariant 0 in characteristic 3. The group scheme G has order 12 and sits in an exact sequence

$$1 \to A \to G \to \mu_4 \to 1,$$

where $A \cong \mathbf{Z}/3\mathbf{Z}$ (see, for instance, [65, Exercise A.1]) and $\lambda \in \mu_4$ acts on A by multiplication by λ^2. The pullback of $\mathcal{L}$ to BG is a line bundle on BG, which is necessarily trivial on the commutator subgroup A. So $\mathcal{L}$ pulls back to the zero bundle under the composition $BA \to BG \to \overline{\mathcal{M}}_{1,1}$, which means that any point x in the image of $BA(K) \to \overline{\mathcal{M}}_{1,1}(K)$ has $\mathrm{ht}_{\mathcal{L}}(x) = 0$. There are infinitely many such points, corresponding to the $\mathbf{Z}/3\mathbf{Z}$-extensions of K. Concretely, elliptic curves given by Weierstrass equations of the form

$$y^2 = x^3 - x - f(t) \tag{3.12}$$

all have height 0 with respect to $\mathcal{L}$. Another way to see this is to observe that the space of sections of $\mathcal{L}^{12}$—that is, of weight-12 modular forms of level 1 in characteristic 3—is two-dimensional and is

[3] We do not know a citation for this fact in the published literature but learned it via personal communication from Xinyi Yuan.

spanned by Δ and b_2, where b_2 is the Hasse invariant. [22, Prop 6.2]. Any Weierstrass equation of type (3.12) has $b_2(E) = 0$ and $\Delta(E) = 1$ ([65, Appendix A, Prop 1.1. b)]). So by Proposition 2.28, using the fact that Δ is constant, we see again that $\mathrm{ht}_{\mathcal{L}}(E) = 0$ for any such E.

This does not mean, however, that Faltings height is a different kind of height from those discussed in this paper; it only means it does not agree with the height arising from the Hodge bundle or any of its powers. But, as explained in a paper of Meier [51], there are other vector bundles! When K is a field of characteristic greater than 3, every vector bundle on $\overline{\mathcal{M}}_{1,1}$ is isomorphic to a direct sum of line bundles, which can only be powers of the Hodge bundle [51, Cor 3.6]), essentially because $\overline{\mathcal{M}}_{1,1}$ is a weighted projective line in this case. But in characteristic 2 and 3, Meier constructs indecomposable higher-rank vector bundles on $\overline{\mathcal{M}}_{1,1}/K$.[4] Thus, the following question still makes sense.

Question 3.13 (A. Landesman). When K is a global field of characteristic 2 or 3, is there a vector bundle $\mathcal{V}$ on $\overline{\mathcal{M}}_{1,1}/K$ such that $\mathrm{ht}_{\mathcal{V}} = c\,\mathrm{ht}_{\mathrm{Fal}}$ for some $c \in \mathbf{Z}$?

We originally asked this question with $c = 1$; that is, is the Faltings height itself a height in our sense? Landesman showed in his thesis [45] that this is too much to hope for; in characteristic 3, there is no vector bundle $\mathcal{V}$ on $\overline{\mathcal{M}}_{1,1}/\mathbf{F}_3$ with $\mathrm{ht}_{\mathcal{V}} = \mathrm{ht}_{\mathrm{Fal}}$. However, Landesman then raises the question stated above, which remains open: Is there a vector bundle which computes some integer multiple of the Faltings height in characteristic 3? For that matter, is there even a vector bundle whose associated height is Northcott?

Furthermore, one may ask the same question about abelian varieties of higher dimension. The Faltings height is usually thought of as being related to the Hodge bundle on the moduli stack $\bar{\mathcal{A}}_g$. But the stacky height associated to this line bundle, or *any* line bundle, will not be Northcott on $\bar{\mathcal{A}}_g$, for the same reason it failed to be Northcott for $\overline{\mathcal{M}}_{1,1}$ in low characteristic; there are abelian varieties of dimension d with nonabelian automorphism group, which give rise to maps $BG \hookrightarrow \bar{\mathcal{A}}_g$ for nonabelian G, and no line bundle on BG can be Northcott. This problem can be avoided by computing heights on $\bar{\mathcal{A}}_g$ with respect to the rank-g vector bundle $\mathcal{V} = e^*\Omega^1_{A/\bar{\mathcal{A}}_g}$, where A is the universal principally polarized abelian variety over the moduli stack, rather than with respect to its determinant, the Hodge bundle. There will still be problems in low characteristic, as we have seen from the case of elliptic curves. One way of understanding the difficulty with curves of the form (3.12) is that a wildly ramified extension of K is necessary in order to arrive at a curve with semistable reduction; this cannot be the case for elliptic curves over fields of characteristic 5 or greater. The following question thus seems reasonable.

Question 3.14. When K is a global function field, $\mathcal{V}$ is the vector bundle $e^*\Omega^1_{A/\overline{\mathcal{A}}_g}$ on $\overline{\mathcal{A}}_g$, and A/K is an abelian variety that becomes semistable over an everywhere tamely ramified extension of K, is it the case that

$$\mathrm{ht}_{\mathcal{V}}(A) = \mathrm{ht}_{\mathrm{Fal}}(A)?$$

If Questions 3.13 and 3.14 both have a positive answer, one might well ask the common descendant of both questions: are there "exotic" vector bundles on $\mathcal{A}_g$ in small (relative to g) characteristic which compute the Faltings height of abelian varieties that require a wild extension to become semistable?

Finally, we return for a moment to the number field case. Because of the singularity at the boundary of $\overline{\mathcal{A}}_g$ of the Faltings metric, we cannot expect $\mathrm{ht}_{\mathcal{V}}$ to match $\mathrm{ht}_{\mathrm{Fal}}$ exactly. But there is a way to ask whether the two heights agree 'apart from the Archimedean place.' Namely, we can ask the following.

Question 3.15. Let K be a global field, let v be a non-Archimedean place of K, and let A/K be an abelian variety which becomes semistable over a tamely ramified extension of K_v. Is the component at v of $\mathrm{ht}_{\mathrm{Fal}}(A) - \mathrm{ht}_s(A)$ equal to $\delta_{\mathcal{V};v}(A)$?

This is a purely local question which has to do with the behavior of the tangent space to the Néron model of A under ramified base change. A positive answer to Question 3.15 would imply a positive

[4] Meier only describes these bundles on $\mathcal{M}_{1,1}$. but it is not hard to show they extend to the compactification.

answer to Question 3.14, as follows. The stable Faltings height $\mathrm{ht}_s(A)$ agrees with $\mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}$ in this setting, because both are given by the degree of the pullback of the Hodge bundle to an integral point $C' \to \overline{\mathcal{A}}_g$, where C' is a cover of C. And since there are no Archimedean places, the positive answer to Question 3.15 shows that

$$\mathrm{ht}^{\mathrm{Fal}}(A) - \mathrm{ht}_s(A) = \sum_v \delta_{\mathcal{V};v}(A) = \mathrm{ht}_{\mathcal{V}}(A) - \mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}(A).$$

3.5. *Heights on footballs*

A football[5] $\mathcal{F}(a,b)$ is a $\mathbb{P}^1$ rooted at 0 and ∞, with residual gerbes μ_a and μ_b, respectively. Let K be a global field; we emphasize that K is allowed to have any characteristic, including characteristics dividing a or b. (When K has one of these characteristics, $\mathcal{F}(a,b)$ is a tame Artin stack but not a Deligne–Mumford stack.) As an illustration of the (moderate) subtlety of the Northcott condition in the stacky case, we will work out which line bundles on $\mathcal{F}(a,b)$ are Northcott.

There are three kinds of K-points of $\mathcal{F}(a,b)$, which may be treated separately.

- The points supported at 0; these are naturally identified with K-points of $B(\mu_a)$, which are in turn identified with the set $K^*/(K^*)^a$;
- The points supported at ∞; these are naturally identified with K-points of $B(\mu_b)$, which are in turn identified with the set $K^*/(K^*)^b$;
- The rest of the points, which are naturally identified with the points on $\mathbb{P}^1(K)$ other than 0 and ∞; that is, these points are in bijection with K^*.

Any divisor on $\mathcal{F}(a,b)$ is linearly equivalent to one of the form $d[P] + n[0] + m[\infty]$, where P is some point on $\mathbb{G}_m$; such a divisor has degree $d + n/a + m/b$. This expression is not unique but is subject to the relations $a[0] \sim b[\infty] \sim [P]$. Take $\mathcal{L}$ to be the line bundle on $\mathcal{F}(a,b)$ corresponding to $d[P] + n[0] + m[\infty]$. We now explain how to compute $\mathrm{ht}_{\mathcal{L}}(x)$ for $x \in \mathcal{F}(a,b)(K)$.

For the first two types of points, this computation of height has already been carried out in Section 2.4. For a point x of the first type, d and m are irrelevant. The class in $K^*/(K^*)^a$ associated to x is represented by a function $f \in K^*$, and the height of x is a sum over places of K:

$$\mathrm{ht}_{\mathcal{L}}(x) = \sum_v \left\lceil \frac{n}{a} \mathrm{ord}_v(f) \right\rceil.$$

Similarly, for a point of the second type, represented by the class of g in $K^*/(K^*)^b$ the height is

$$\mathrm{ht}_{\mathcal{L}}(x) = \sum_v \left\lceil \frac{m}{b} \mathrm{ord}_v(g) \right\rceil.$$

We now treat points of the third, or generic type. For simplicity of description, take K to be the function field of a smooth proper curve $C/\mathbb{F}_q$. Then x affords a rational map ϕ from C to $\mathcal{F}(a,b)$. Write $\phi_c : C \to \mathbb{P}^1$ for the composition of ϕ with the coarse moduli map, denote $\deg \phi_c = \deg \phi$ by e, and write $\sum e_i P_i$ for the divisor $\phi_c^*[0]$ and $\sum e_i' Q_i$ for the divisor $\phi_c^*[\infty]$. Then $\sum e_i \deg P_i = \sum e_i' \deg Q_i = e$.

We may take $\mathcal{C}$ to be a root stack with residual gerbe μ_a at the P_i and μ_b at the Q_i. Then $\overline{x}^* \mathcal{L}^\vee$ is the divisor

$$-\left(d\phi^{-1}(P) + \sum_i \frac{e_i n}{a} P_i + \sum_i \frac{e_i' m}{b} Q_i \right)$$

whose degree, as it must be, is $-e \deg \mathcal{L}$.

[5]The 'football' here is understood to be an American football, which has two singular points. In the professional sporting context, the residual gerbes at these points are not specified.

We then have

$$\mathrm{ht}_{\mathcal{L}}(x) = -\deg \pi_* \overline{x}^* \mathcal{L} = -\left(-ed + \sum_i \left\lfloor -\frac{e_i n}{a} \right\rfloor \deg P_i + \sum_i \left\lfloor -\frac{e_i' m}{b} \right\rfloor \deg Q_i\right) \log q \tag{3.16}$$

$$= \left(ed + \sum_i \left\lceil \frac{e_i n}{a} \right\rceil \deg P_i + \sum_i \left\lceil \frac{e_i' m}{b} \right\rceil \deg Q_i\right) \log q. \tag{3.17}$$

In particular, we note that $\mathrm{ht}_{\mathcal{L}}(x) \geq e \log q \deg \mathcal{L}$, with equality holding just when every e_i is a multiple of a and every e_i' is a multiple of b, which is to say, when x actually extends to an integral point of $\mathcal{F}(a, b)$.

This description suffices to tell us which line bundles have the Northcott property. We already see that the set of Northcott line bundles does not form a cone because it is not closed under addition. (Indeed, we could have already seen that from the case $B(\mathbf{Z}/2\mathbf{Z})$, where the nontrivial line bundle $\mathcal{L}$ is Northcott and $\mathcal{L}^{\otimes 2}$, which is trivial, is not Northcott.)

Proposition 3.18. *Choose a, b coprime integers, and let K be the function field of a curve C. A divisor $L = d + n[0] + m[\infty]$ on $\mathcal{F}(a, b)$ is Northcott if and only if* $\deg L > 0$ *and* $(n, a) = (m, b) = 1$.

Proof. Suppose $(n, a) = r > 1$. Then any point of $\mathbb{P}(a, b)$ of the first type which corresponds to $f \in (K^*)^{a/r}/(K^*)^a \subset K^*/(K^*)^a$ has height 0 with respect to L, which contradicts Northcott. The argument is just the same if $(m, b) > 1$.

We observe that there are infinitely many maps $\mathbb{P}^1 \to \mathcal{F}(a, b)$; namely, those whose coarse map $\mathbb{P}^1 \to \mathbb{P}^1$ is of the form $[B(s, t)^b : A(s, t)^a]$. Any such map, pulled back to C via a map $C \to \mathbb{P}^1$, gives an integral point $C \to \mathcal{F}(a, b)$ of some coarse degree e, whose height is $e \deg L$; we can make e as large as we want, which shows that L cannot be Northcott if $\deg L \leq 0$.

Suppose, on the other hand, that all three conditions are met. We have already shown that points x of the third type have $\mathrm{ht}_L(x) \geq e \log q \deg L$; since $\deg L$ is positive, $\mathrm{ht}_L(x)$ gives an upper bound for e, which makes the set of possible x finite. For points of the first type represented by $f \in (K^*)/(K^*)^a$, we observe that

$$\mathrm{ht}_{\mathcal{L}}(x) = \sum_v \left\lceil \frac{n}{a} \operatorname{ord}_v(f) \right\rceil = \sum_v \left\{ \frac{n}{a} \operatorname{ord}_v(f) \right\},$$

the latter equality following from $\sum_v \operatorname{ord}_v(f) = 0$. So a bound on the height of x yields a bound on the number of places, where $\frac{n}{a} \operatorname{ord}_v(f)$ is not an integer; since $(n, a) = 1$, this bounds the number of places where (more precisely: the degree of the divisor where) $\operatorname{ord}_v(f)$ is not a multiple of a. Bounding this quantity places f within a finite set of cosets of $(K^*)^a$, so we are done. The case of points of the second type is exactly the same. □

3.5.1. Consistency check: footballs and weighted projective lines

When a and b are relatively prime, the football $\mathcal{F}(a, b)$ is isomorphic to the weighted projective line $\mathbb{P}(a, b)$; on K-points, the isomorphism ψ from $\mathbb{P}(a, b)$ to $\mathcal{F}(a, b)$ sends $(s : t)$ to the point t^a/s^b when $st \neq 0$. Let m, n be integers such that $ma + nb = 1$; then the line bundle $\mathcal{L} = n[0] + m[\infty]$ on $\mathbb{F}(a, b)$ has degree $1/ab$, and its pullback to $\mathbb{P}(a, b)$ is the tautological bundle $\mathcal{O}_{\mathbb{P}(a,b)}(1)$. If x is a point of $\mathbb{P}(a, b)(K)$, we have

$$\mathrm{ht}_{\mathcal{O}_{\mathbb{P}(a,b)}(1)}(x) = \mathrm{ht}_{\mathcal{L}}(\psi(x)).$$

This provides an opportunity to check consistency between the formulas we have given for the height of a point on weighted projective space and the height of a point on a football. Let $x = (s : t)$ be a point of $\mathbb{P}(a, b)$. Then by equation (3.5), we have

$$\mathrm{ht}_{\mathcal{O}_{\mathbb{P}(a,b)}(1)}(x) = \sum_v \lceil \log_{q_v} \max(|s|_v^{1/a}, |t|_v^{1/b}) \rceil \log q_v. \tag{3.19}$$

We now compute $\mathrm{ht}_{\mathcal{L}}(\psi(x))$. Recall that $\psi(x)$ is the point on $\mathbb{F}(a,b)$ corresponding to the point t^a/s^b of $\mathbb{P}^1(K)$. In the notation of the above section, the points P_i correspond to those places v of K, where $a\,\mathrm{ord}_v\, t - b\,\mathrm{ord}_v\, s > 0$, and the points Q_i to those places where $a\,\mathrm{ord}_v\, t - b\,\mathrm{ord}_v\, s < 0$. When v is a prime with $a\,\mathrm{ord}_v\, t - b\,\mathrm{ord}_v\, s > 0$, we have, again maintaining the notation of equation (3.17),

$$e_i = a\,\mathrm{ord}_v\, t - b\,\mathrm{ord}_v\, s$$

and

$$\deg P_i = \log q_v / \log q.$$

So the contribution of v to equation (3.17) is

$$\begin{aligned}\left\lceil \frac{(a\,\mathrm{ord}_v\, t - b\,\mathrm{ord}_v\, s)n}{a} \right\rceil \log q_v &= \left(n\,\mathrm{ord}_v\, t - \left\lceil \frac{nb\,\mathrm{ord}_v\, s}{a} \right\rceil \right) \log q_v \\ &= \left(n\,\mathrm{ord}_v\, t + m\,\mathrm{ord}_v\, s - \left\lceil \frac{\mathrm{ord}_v\, s}{a} \right\rceil \right) \log q_v .\end{aligned}$$

By a similar argument, one shows that when $a\,\mathrm{ord}_v\, t - b\,\mathrm{ord}_v\, s < 0$ one gets a contribution of

$$\left(n\,\mathrm{ord}_v\, t + m\,\mathrm{ord}_v\, s - \left\lceil \frac{\mathrm{ord}_v\, t}{b} \right\rceil \right) \log q_v .$$

Since the first case obtains exactly when $\mathrm{ord}_v\, s/a < \mathrm{ord}_v\, t/b$, we can express the contribution of v uniformly as

$$\left(n\,\mathrm{ord}_v\, t + m\,\mathrm{ord}_v\, s - \left\lceil \min\left(\frac{\mathrm{ord}_v\, s}{a}, \frac{\mathrm{ord}_v\, t}{b} \right) \right\rceil \right) \log q_v .$$

Summing over v, the first two terms vanish by the product formula, and we are left with

$$\mathrm{ht}_{\mathcal{L}}(\psi(x)) = -\sum_v \min\left(\left\lceil \frac{\mathrm{ord}_v\, s}{a} \right\rceil, \left\lceil \frac{\mathrm{ord}_v\, t}{b} \right\rceil \right) \log q_v$$

which is just equation (3.19) in another form.

3.6. Heights on symmetric powers of varieties

There is a substantial literature about points on varieties of bounded algebraic degree. We explain how these questions look through the lens of heights on stacks. Let X be a smooth proper scheme of dimension n over K. A point on X of algebraic degree m over K can be thought of as a K-point on the stack $\mathrm{Sym}^m X = [X^m/S_m]$. In this section, we explain how to compute the height of such a point. Slightly more generally, let G be a subgroup of S_m, and let $\mathcal{X}$ be the quotient $[X^m/G]$; when $G = S_m$, our stack $\mathcal{X}$ is $\mathrm{Sym}^m X$.

In order to talk about height, we need to choose a vector bundle $\mathcal{V}$ on $\mathcal{X}$; this is the same thing as an G-equivariant vector bundle on X^m. The choice we make is as follows: Let V_0 be some vector bundle of rank r on X, and let $\pi_1, \ldots, \pi_m \colon X^m \to X$ be the m projections. Then $\widetilde{V} = \oplus_i \pi_i^* V_0$ is an G-equivariant vector bundle of rank mr, which descends to a vector bundle $\mathcal{V}$ of rank mr on $\mathcal{X}$.

Let x be a point of $\mathcal{X}(K)$. We begin by computing the stable height $\mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}(x)$. The Cartesian square

$$\begin{array}{ccc} \mathrm{Spec}\, L & \xrightarrow{x_L} & X^m \\ \downarrow & & \downarrow \\ \mathrm{Spec}\, K & \xrightarrow{x} & \mathcal{X} \end{array}$$

provides an étale algebra L over K which carries an S_m-action and a rational point x_L which extends to an integral point $C \to X^m$. By Proposition 2.14,

$$\mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}(x) = [L : K]^{-1} \, \mathrm{ht}^{\mathrm{st}}_{\widetilde{V}}(x_L).$$

(Should L be an étale algebra which is not a field but rather a direct sum $\oplus_i F_i$, our convention is that the height of a point of $X^m(L)$ is $\sum_i \mathrm{ht}(P_i)$, where $P_i \in X^m(F_i)$ are the points corresponding to the restriction of x_L : $\mathrm{Spec}\, L \to X^m$ to connected components of $\mathrm{Spec}\, L$.)

Since X^m is a scheme, we have

$$\mathrm{ht}^{\mathrm{st}}_{\widetilde{V}}(x_L) = \mathrm{ht}_{\widetilde{V}}(x_L).$$

The latter quantity is a very natural one, what you might call the 'absolute height' of x. Suppose, for instance, that L/K is a field extension, necessarily Galois with Galois group G. Then x_L is a point of $X^m(L)$ on which $\mathrm{Gal}(L/K)$ acts by permutations; in other words, it is an element $(\alpha_1, \ldots, \alpha_m)$, where the α_i are conjugate and each α_i is contained in a degree-m extension L_i/K whose Galois closure is L. The (unordered) set $\alpha_1, \ldots, \alpha_m$ can be thought of as a K-rational Galois orbit of points on X, and the height of x_L is then given by the usual Weil height on X:

$$\mathrm{ht}_{\widetilde{V}}(x_L) = \sum_i \mathrm{ht}_{V_0;L} \, \alpha_i = m \, \mathrm{ht}_{V_0;L} \, \alpha_1,$$

where the subscript L is indicating that the height of α_i is understood to mean the height of α_i as a point of $X(L)$, not of $X(L_i)$; to sum up, this means that

$$\mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}(x) = |G|^{-1} \, \mathrm{ht}_{\widetilde{V}}(x_L) = m|G|^{-1} \, \mathrm{ht}_{V_0;L} \, \alpha_i = \mathrm{ht}_{V_0;L_i} \, \alpha_i$$

which is the same for every i. In fact, the reader will note that nothing we did actually used the hypothesis that L was a field, so the description of the stable height of x is valid also in the case where L is an étale algebra other than a field. For instance, if L splits completely as a product of copies of K, then L_i is isomorphic to K^m, and our point $x \in \mathcal{X}(K)$ may be thought of as an unordered m-tuple $\{Q_1, \ldots, Q_m\} \subset X(K)$; in that case,

$$\mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}(x) = \mathrm{ht}_{V_0;L_i}(Q_1, \ldots Q_m) = \sum_{i=1}^{m} \mathrm{ht}_{V_0;K} \, Q_i.$$

We now consider the discrepancy $\delta_{\mathcal{V}}(x) = \mathrm{ht}_{\mathcal{V}}(x) - \mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}(x)$.

Proposition 3.20. *The value of $\delta_{\mathcal{V}}(x)$ is the same for any V_0 and V_0' of the same rank r.*

Proof. We write $\widetilde{V}', \mathcal{V}'$ for the vector bundles on X^m and $\mathcal{X}$ respectively obtained from V_0' as $\widetilde{V}, \mathcal{V}$ were obtained from V_0.

The discrepancy is a sum of local terms $\delta_{v;\mathcal{V}}(x)$ where v ranges over a finite list of non-Archimedean places v of C where x does not extend to an $\mathcal{O}_{K_v}$-point; in particular, this list depends only on x, not on the choice of $\mathcal{V}$. Choose such a v; denoting by $\mathcal{C}_v$ the infinitesimal neighborhood of the tuning stack $\mathcal{C}$ over v, we have a commutative diagram

$$\begin{array}{ccc} \mathrm{Spec}\, \mathcal{O}_{L_v} & \xrightarrow{\overline{x}_{L_v}} & X^m \\ \downarrow & & \downarrow \\ \mathcal{C}_v & \xrightarrow{\overline{x}_{K_v}} & \mathcal{X}, \end{array}$$

where L_v denotes $L \otimes_K K_v$, so $\mathcal{O}_{L_v}$ is a disjoint union of dvrs. Composing $\overline{x}_{L_v}$ with the projection maps $p_1, \dots, p_m$ yields maps $q_1, \dots, q_m \colon \operatorname{Spec} \mathcal{O}_{L_v} \to X$ which are permuted by composition with the action of G on $\operatorname{Spec} \mathcal{O}_{L_v}$. We may take $U \subset X$ to be an open subscheme containing the image of the q_i on which V_0 and V_0' become isomorphic (and indeed we may choose U to make both isomorphic to $\mathcal{O}_U^r$).

Now, $\overline{x}_{L_v}^* \widetilde{V}$ can be described as $\oplus_i q_i^* V_0$, where the action of G permutes the factors; we note that this is *G-equivariantly* isomorphic to $\overline{x}_{L_v}^* \widetilde{V}' = \oplus_i q_i^* V_0'$. Thus, the vector bundle $\overline{x}_{K_v}^* \mathcal{V}$, which is the descent of $\overline{x}_{L_v}^* \widetilde{V}$, is isomorphic to $\overline{x}_{K_v}^* \mathcal{V}'$. Since $\delta_{v;\mathcal{V}}(x)$ depends only on $\overline{x}_{L_v}^* \widetilde{V}$, we conclude that

$$\delta_{v;\mathcal{V}}(x) = \delta_{v;\mathcal{V}'}(x)$$

as desired. □

Given Proposition 3.20, we are free to take $V_0 = \mathcal{O}_X^r$ when computing $\delta_{\mathcal{V}}(x)$. In this case, V is the direct sum of r copies of the vector bundle on $\mathcal{X}$ obtained by taking $V_0 = \mathcal{O}_X$; so we may simply take $V_0 = \mathcal{O}_X$ and multiply by r at the end.

In this case, we can describe $\mathcal{V}$ very concretely; in the diagram

$$\begin{array}{ccc} X^m & \longrightarrow & * \\ \downarrow & & \downarrow \\ \mathcal{X} & \xrightarrow{c} & BG \end{array}$$

we have that $\mathrm{ht}_{\mathcal{V}}(x) = \mathrm{ht}_{\rho}(c \circ x)$, where $c \circ x$ is the morphism from $\operatorname{Spec} K$ to BG corresponding to the etale G-extension L/K. It follows from Proposition 3.1 that $\mathrm{ht}_{\rho}(c \circ x) = (1/2) \log \Delta_{L_i/K}$ (which is the same for all i). The pullback of ρ to $*$ is trivial, so $\mathrm{ht}_{\rho}^{\mathrm{st}}$ is identically 0, whence the discrepancy $\delta_{\rho}(c \circ x)$ is also $(1/2) \log \Delta_{L_i/K}$. We can now conclude from the discussion above that, for any choice of V_0,

$$\delta_{\mathcal{V}}(x) = (r/2) \log \Delta_{L_i/K}.$$

Combining this with our computation of $\mathrm{ht}_{\mathcal{V}}^{\mathrm{st}}$, we finally arrive at a description of the height of a rational point on $\mathcal{X}$ with respect to $\mathcal{V}$. Recall that a point $x \in \mathcal{X}(K)$ provides us with a degree-m etale extension L_1/K and a point $\alpha_1 \in X(L_1)$. Denote by $\mathrm{ht}_{L_1}^W(\alpha_1)$ the usual Weil height of α_1 under the map $X(L_1) \to \mathbf{R}$ afforded by V_0. Then

$$\mathrm{ht}_{\mathcal{V}}(x) = \mathrm{ht}_{L_1}^W(\alpha_1) + (r/2) \log \Delta_{L_1/K}.$$

4. Counting rational points by height: a conjecture of Batyrev–Manin–Malle type

In this section, we formulate a conjecture of Batyrev–Manin–Malle type for rational points of bounded height on a stack $\mathcal{X}$. When $\mathcal{X}$ is a scheme, we recover the weak Batyrev–Manin conjecture about rational points on schemes; when we take $\mathcal{X} = BG$, we recover the weak Malle conjecture. We thus think of our conjecture as interpolating between the two conjectures, while at the same time generating many new cases of interest. As was the case for the original Batyrev–Manin, we develop our heuristics by consideration of the case $K = k(t)$ and the corresponding geometric problem of studying spaces of rational curves on $\mathcal{X}$.

By 'weak' in the above paragraph we mean that we propose conjectures that bound counting functions between X^a and $X^{a+\epsilon}$ for a specified exponent a. The 'strong' versions of Batyrev–Manin and Malle make a more precise conjecture, that counting functions are asymptotic to $X^a (\log X)^b$ for specified a, b. In work posted after the original version of this paper was released, Darda and Yasuda [21] have proposed a 'strong' conjecture about point-counting on stacks, with an explicit predicted power of $\log X$.

One could go further still and ask whether the counting functions discussed here are of the form $cX^a (\log X)^b + o(X^a (\log X)^b)$, with an explicit constant c; this has been quite an active area of

investigation in both the Batyrev–Manin and the Malle context. One remark in this regard: To get constants right, it is presumably important to remember that $\mathcal{X}(K)$ is naturally not a set but a groupoid, and counts of points should probably be weighted inversely to the size of the point's automorphism group. But issues of this kind will not be relevant for the coarser heuristics considered here.

4.1. Expected deformation dimension: stacky anticanonical height

In the Batyrev–Manin conjecture for a scheme X, when counting rational points with respect to a line bundle $\mathcal{L}$, the expected growth rate is given by $B^{a(\mathcal{L})}$, where the *Fujita invariant* $a(\mathcal{L})$ is the infimum of all a for which $a\mathcal{L} + K_X$ is effective. A technical hurdle we must overcome in defining $a(\mathcal{L})$ for stacks $\mathcal{X}$ is that for many stacks of interest, for example $\mathcal{X} = BG$, the canonical bundle $K_{\mathcal{X}}$ is trivial! Thus, the anticanonical height is not suitable for the purposes of obtaining the expected growth rate of point counts on stacks. Our solution is to introduce a new quantity, the expected deformation dimension (or edd), which is a suitable perturbation of the anticanonical height.

Before giving the definition of edd, we wish to sufficiently motivate it through geometric intuition. In the case of a proper scheme X over a function field $\mathbb{C}(t)$, a rational point $x\colon \operatorname{Spec}\mathbb{C}(t) \to X$, by the valuative criterion, extends to a map $\overline{x}\colon \mathbb{P}^1 \to X$. By Riemann–Roch, the anticanonical height $\mathrm{ht}_{-K_X}(x) = \deg(\overline{x}^* T_X)$ differs from $\chi(\overline{x}^* T_X)$ by a constant, and $\chi(\overline{x}^* T_X)$ is the expected dimension of the deformation space of $\overline{x}^*$.

The deformation theoretic point of view serves as our launching point for the definition of edd. Given a rational point $x\colon \operatorname{Spec} K \to \mathcal{X}$ of a stack, we can extend x to a *universal* tuning stack $\overline{x}\colon \mathcal{C} \to \mathcal{X}$; see Definition 2.1. The expected deformation dimension of $\overline{x}$ is then given by $\chi(L_{\overline{x}}^\vee[1])$, where $L_{\overline{x}}$ is the cotangent complex for the representable map $\overline{x}$. For the sake of motivational purposes, suppose both $\mathcal{X}$ and $\mathcal{C}$ are smooth tame Deligne–Mumford stacks, in which case the tangent complexes $L_{\mathcal{X}}^\vee$ and $L_{\mathcal{C}}^\vee$ are vector bundles, denoted by $T_{\mathcal{X}}$ and $T_{\mathcal{C}}$. Then

$$\chi(L_{\overline{x}}^\vee[1]) = \chi(\overline{x}^* T_{\mathcal{X}}) - \chi(T_{\mathcal{C}}),$$

which up to constants are the same as

$$\deg(\pi_* \overline{x}^* T_{\mathcal{X}}) - \deg(\pi_* T_{\mathcal{C}}). \tag{4.1}$$

Note that $\deg(\pi_* \overline{x}^* T_{\mathcal{X}}) = -\mathrm{ht}_{K_{\mathcal{X}}}(x)$. We next calculate $\deg(\pi_* T_{\mathcal{C}})$. Letting $\pi\colon \mathcal{C} \to C$ be the coarse space, we have

$$\Omega^1_{\mathcal{C}} = \pi^* \Omega^1_C \otimes \mathcal{O}_{\mathcal{C}}\Big(\sum (1 - e_p^{-1})p\Big) \tag{4.2}$$

by [71, Lemma 5.5.3 and Proposition 5.5.6]. So,

$$\pi_* T_{\mathcal{C}} = T_C \otimes \mathcal{O}_C\Big(\sum \lfloor e_p^{-1} - 1 \rfloor p\Big) = T_C(-R);$$

since the floors are equal to -1 if e_p is nontrivial and 0 otherwise, R is the divisor given by the ramified points taken without multiplicity. So, up to constants, $\deg(\pi_* T_{\mathcal{C}}) = -\deg(R)$.

In practice, however, we will want to consider stacks $\mathcal{X}_0$ over K for which we do not have in mind a particular model $\mathcal{X}/C$ which is normal and Deligne–Mumford or for which we do have in mind a model but it isn't Deligne–Mumford; for example, we don't want to exclude a stack like $B\mu_n/\operatorname{Spec}\mathbf{Z}$ which fails to be Deligne–Mumford in characteristics dividing n. Tuning stacks for rational points of such stacks are also generally not Deligne–Mumford. Presumably a more complicated definition involving the tangent *complex* would work, but in the interest of simplicity we have chosen for now to apply a technical workaround.

First, the universal tuning stack $\mathcal{C}$ of a rational point $x \in \mathcal{X}(K)$ is generically a scheme (and thus generically Deligne–Mumford). The coarse space map $\pi\colon \mathcal{C} \to C$ is birational and $\mathcal{C}$ is normal; if $\mathcal{C}$

is tame, then it is a root stack. To promote our working definition of edd (equation 4.1) to the general setting we are tempted to define

$$\Omega_{\mathcal{C}}^{1,\mathrm{fake}} = \pi^*\Omega_C^1 \otimes \mathcal{O}_{\mathcal{C}}\Big(\sum(1 - e_p^{-1})p\Big).$$

If p is a Deligne–Mumford point of $\mathcal{C}$ but is not tame, then one defines e_p via wild ramification [43, Proposition 7.1]. But if p is not a Deligne–Mumford point it is unclear how to define e_p. If p is tame, then the stabilizer of p is isomorphic to μ_m for some integer m, and it is tempting to define e_p to be $1/m$. This is ad hoc, but worse, not general enough: The stabilizer could be a group which is neither étale nor tame (such as $\mu_p \times \mathbf{Z}/p\mathbf{Z}$).

Our perspective is that the precise definition of e_p does not matter, as long as it is nontrivial at a stacky point. What we mean is: For the part of the definition of edd that relies on the universal tuning stack we only ever consider the quantity 'deg$(\pi_* T_{\mathcal{C}})$'. Since $T_{\mathcal{C}}$ is the dual of $\Omega_{\mathcal{C}}^{1,\mathrm{fake}}$,

$$\pi_* T_{\mathcal{C}} = T_C \otimes \mathcal{O}_C\Big(\sum\lfloor e_p^{-1} - 1\rfloor p\Big).$$

In particular, since we are taking floors the quantity $\lfloor e_p^{-1} - 1\rfloor$ is 0 is p is not stacky and is -1 otherwise. In equation 4.1, we thus abstain from defining $T_{\mathcal{C}}$ and instead replace $\deg(\pi_* T_{\mathcal{C}})$ with the following quantity.

Definition 4.3 (Reduced discriminant). Let $\pi\colon \mathcal{C} \to C$ be a tuning stack of a rational point $x \in \mathcal{X}(K)$. We define the reduced discriminant $\mathrm{rDisc}(x)$ of x to be the sum

$$\mathrm{rDisc}(x) = \sum \log q_v$$

over the stacky points v of $\mathcal{C}$, where q_v is the cardinality of the residue field of the point v.

To make sense more generally of the other term of equation 4.1, for the rest of this section, in addition to the assumptions of Subsection 1.1 and Section 2, we assume that the generic fiber $\mathcal{X}_K$ of our proper Artin stack $p\colon \mathcal{X} \to C$ is Deligne–Mumford so that it makes sense to talk about the canonical sheaf $K_{\mathcal{X}_K}$ of the generic fiber.

Definition 4.4. We say a line bundle on $\mathcal{X}$ is generically canonical if its restriction to $\mathcal{X}_K$ is $K_{\mathcal{X}_K}$.

We now define edd as follows, guided by the motivation above.

Definition 4.5 (Expected deformation dimension). Let K be a global field, and let C be either Spec O_K in the number field case or a smooth proper curve with function field K in the function field case. Let $\mathcal{X}$ be a proper Artin stack over C with finite diagonal such that $\mathcal{X}$ is a smooth proper Deligne–Mumford stack over K. Let $\widetilde{K}$ be a generically canonical line bundle on $\mathcal{X}$. Given $x \in \mathcal{X}(K)$, let $(\mathcal{C}, \overline{x}, \pi)$ be its universal tuning stack. The expected deformation dimension of x is

$$\mathrm{edd}(x) := -\mathrm{ht}_{\widetilde{K}}(x) + \mathrm{rDisc}(x).$$

Remark 4.6. Implicit in this definition is a conjecture: That the definition is independent of choices. More precisely, we expect that, given two different models of $\mathcal{X}_K$, and two different extensions of $K_{\mathcal{X}_K}$ to these models, the two functions $\mathrm{edd}(x)$ would differ by a function that is bounded as x ranges over $\mathcal{X}(K)$. In the examples that follow, we will simply choose a model $\mathcal{X}$ and choose a generically canonical line bundle on $\mathcal{X}$.

Remark 4.7. If $\mathcal{X} = X$ is a scheme, then the universal tuning stack is a curve, and edd agrees with the anticanonical height since $\mathrm{edd}(x) := -\mathrm{ht}_{K_X}(x) = \deg(\overline{x}^* T_X) = \mathrm{ht}_{-K_X}(x)$. On the other extreme, if $\mathcal{X} = BG$, then $K_{\mathcal{X}}$ is trivial, so $\mathrm{edd}(x)$ is the reduced discriminant of the field extension corresponding to x.

Example 4.8 (Extending a stacky curve and its canonical bundle). Let $\mathcal{X}_0$ be a smooth tame Deligne–Mumford stacky curve over K, and suppose that the coarse space map $\phi_0\colon \mathcal{X}_0 \to X_0$ is birational (equivalently, $\mathcal{X}_0$ has trivial generic inertia). By [33, Theorem 1 and Remark 4], such an $\mathcal{X}_0$ is isomorphic to a root stack over its coarse space. Let $p_1, \dots, p_k \in X_0$ be the ramification locus of ϕ_0; since $\mathcal{X}_0$ is a root stack, the stabilizer group over each p_i is isomorphic to μ_{e_i} for some integer $e_i \geq 2$, and $\mathcal{X}_0$ is the root stack of X_0 rooted along each p_i with order e_i.

The coarse space X_0 is a smooth proper curve over K and extends to a proper relative curve $X \to C$. Let D_i be the closure of p_i. After a possible normalization and sequence of blowups, we can assume that X is regular and that the D_i do not intersect each other or the singular points of the fibers of $X \to C$. Define $\phi\colon \mathcal{X} \to X$ to be the root stack of X rooted along each D_i with order e_i. The relative stacky curve $\mathcal{X}$ is a model of $\mathcal{X}_0$ and is tame. If there is some point v of C and some i such that the residue characteristic of v divides e_i, then $\mathcal{X}$ is an Artin stack which is not Deligne–Mumford; if $C = \operatorname{Spec} \mathcal{O}_K$ for some number field K, then there is always some such v and i.

As discussed above (see Equation 4.2) the canonical sheaf of $\mathcal{X}_0$ is

$$\Omega^1_{\mathcal{X}_0} = \phi_0^* \Omega^1_{X_0} \otimes \mathcal{O}_{\mathcal{X}_0}\Big(\sum (1 - e_i^{-1}) p_i\Big).$$

Define

$$\Omega^{1,\mathrm{fake}}_{\mathcal{X}} = \phi^* \omega_{X/C} \otimes \mathcal{O}_{\mathcal{X}}\Big(\sum (1 - e_i^{-1}) D_i\Big)$$

by the same 'formula'. Then $\Omega^{1,\mathrm{fake}}_{\mathcal{X}}$ is a generically canonical sheaf.

We have seen in Remark 4.7 that when $\mathcal{X}$ is a scheme, edd agrees with anticanonical height, that is, the height of the tangent bundle. It turns out that the same identity holds when $\mathcal{X}$ is a smooth, tame Deligne–Mumford stacky curve with no generic inertia, at least away from the accumulating subvarieties.

Proposition 4.9 (Curves with stacky points). *Let $\mathcal{X}_0$ be a smooth tame Deligne–Mumford stacky curve over K, and suppose that $\mathcal{X}_0$ is birational to its coarse space. Let $\mathcal{X}$ be the model of $\mathcal{X}_0$ given by extending the root data as in Example 4.8, and let $T_{\mathcal{X}}$ be the dual of the generically canonical bundle from Example 4.8. Let x be a point of $\mathcal{X}(K)$. Then*

$$\operatorname{edd}(x) = \operatorname{ht}_{T_{\mathcal{X}}}(x).$$

Proof. Let $\mathcal{C}$ be a tuning stack and $\overline{x}\colon \mathcal{C} \to \mathcal{X}$ the extension of x, as usual. The pullback $\overline{x}^* T_{\mathcal{X}}^\vee$ is a line bundle on $\mathcal{C}$. We first note that

$$\operatorname{ht}_{T_{\mathcal{X}}}(x) + \operatorname{ht}_{T_{\mathcal{X}}^\vee}(x) = \sum_v (\delta_{T_{\mathcal{X}};v}(x) + \delta_{T_{\mathcal{X}}^\vee;v}(x))$$

since

$$\operatorname{ht}^{\mathrm{st}}_{T_{\mathcal{X}}}(x) + \operatorname{ht}^{\mathrm{st}}_{T_{\mathcal{X}}^\vee}(x) = 0.$$

For each closed point v of C, the point x either reduces to a nonstacky point or reduces to a unique stacky point p with stabilizer group μ_m for some integer $m \geq 2$. Let k be the multiplicity of the reduction of x to p (i.e., the multiplicity of the intersection of the images of x and p in the coarse space X). If m divides k, then we can take the tuning stack $\mathcal{C}$ to be a scheme in a neighborhood of v, in which case the discrepancies are 0. Otherwise, $\mathcal{C}_v$ is a root stack which can be resolved by adjoining to K_v an mth

root of a uniformizer. Denote the resulting field extension by L_w. So as in Section 2.3, the restriction of $\overline{x}^* T_{\mathcal{X}}$ to $\mathcal{O}_{K_v}$ is identified with an ideal Λ in $\mathcal{O}_{L_w}$, and we have

$$\delta_{T_{\mathcal{X}};v}(x) = (1/m)\log\left|\frac{\Lambda}{(\Lambda \cap K_v)\otimes_{\mathcal{O}_{K_v}} \mathcal{O}_{L_w}}\right|.$$

Taking π_w to be a uniformizer of $\mathcal{O}_{L_w}$, we may write $\Lambda = \pi_w^{-k}\mathcal{O}_{L_w}$, and so

$$\delta_{T_{\mathcal{X}};v}(x) = ((-k/m) - \lfloor -k/m \rfloor)\log q_v.$$

The restriction $\overline{x}^* T_{\mathcal{X}}^\vee$, by the same argument, is identified with the ideal $\pi_w^k \mathcal{O}_{L_w}$. We conclude that

$$\delta_{T_{\mathcal{X}};v}(x) + \delta_{T_{\mathcal{X}}^\vee;v}(x) = ((-k/m) - \lfloor -k/m \rfloor + k/m - \lfloor k/m \rfloor)\log q_v$$

which is $\log q_v$ unless $m|k$, in which case it is zero. In other words,

$$\mathrm{ht}_{T_{\mathcal{X}}}(x) + \mathrm{ht}_{T_{\mathcal{X}}^\vee}(x) = \sum_v (\delta_{T_{\mathcal{X}};v}(x) + \delta_{T_{\mathcal{X}}^\vee;v}(x)) = \mathrm{rDisc}(x)$$

since $\mathrm{rDisc}(x)$ is precisely the sum of $\log q_v$ over the stacky points v of $\mathcal{C}$. We conclude that

$$\mathrm{edd}(x) = -\mathrm{ht}_{T_{\mathcal{X}}^\vee}(x) + \mathrm{rDisc}(x) = \mathrm{ht}_{T_{\mathcal{X}}}(x)$$

as claimed. □

Remark 4.10. If $\mathcal{X}'$ is a second model of the stacky curve $\mathcal{X}_0$ from Proposition 4.9 and if $\mathcal{X}'$ is tame, one can show that away from finitely many points of C, $\mathcal{X}'$ is a root stack and isomorphic to $\mathcal{X}$; shrinking C further the generically canonical sheaves agree. By Proposition 2.25, the value of $\delta_{T_{\mathcal{X}'};v}(x) + \delta_{T_{\mathcal{X}'}^\vee;v}(x)$ is bounded on $\mathcal{X}'(K)$, and thus the edd associated to the model $\mathcal{X}'$ will only differ by a constant which depends on $\mathcal{X}_0$ and K.

4.2. Weak form of the stacky Batyrev–Manin–Malle conjecture

Having now defined edd, we are ready to state a heuristic for counting rational points of bounded height on a stack. We then show that our heuristic recovers the weak form of the Batyrev–Manin when $\mathcal{X}$ is a scheme and recovers the weak form of the Malle conjecture when $\mathcal{X} = BG$.

Of course, we cannot expect to count points of bounded height unless the height function satisfies some kind of positivity property. In the Batyrev–Manin setting, this is achieved by restricting to heights corresponding to ample line bundles. One does not have as clear a geometric picture of vector bundles on stacks as one does in the setting of line bundles on schemes, so we use for the moment the following definition. We recall that stable height is well behaved under field extension (Proposition 2.14), so we can define an absolute $\mathrm{ht}_{\mathcal{V}}^{\mathrm{st;abs}}$ as a function on $\mathcal{X}(\overline{K})$ by the usual rule:

$$\mathrm{ht}_{\mathcal{V}}^{\mathrm{st;abs}}(x) = [L:K]^{-1}\,\mathrm{ht}_{\mathcal{V}}^{\mathrm{st}}(x)$$

for points of $\mathcal{X}(L)$.

Definition 4.11. We say a vector bundle $\mathcal{V}$ on a stack $\mathcal{X}$ is semipositive if the quantity $\mathrm{ht}_{\mathcal{V}}^{\mathrm{st;abs}}(x)$ is bounded below on $\mathcal{X}(\overline{K})$.

We note that the property of being semipositive is stable under field extensions by Remark 2.16.

Definition 4.12. Let f be a real-valued function on $\mathcal{X}(\overline{K})$. We say f is generically bounded below if there is a proper closed substack $\mathcal{Z}$ of $\mathcal{X}$ and a constant B such that the set of $x \in \mathcal{X}(\overline{K})$ such that $f(x) < [K(x):K]\cdot B$ is contained in $\mathcal{Z}(\overline{K})$, where $K(x)$ is the residue field of x.

Suppose $\mathcal{V}$ is a semipositive vector bundle on X. We consider the function

$$D_{a,\mathcal{V}}(x) = a\,\mathrm{ht}_{\mathcal{V}}(x) - \mathrm{edd}(x)$$

on $\mathcal{X}(\overline{K})$. We note that if $a' > a$, then

$$\begin{aligned} D_{a',\mathcal{V}}(x) &= D_{a,\mathcal{V}}(x) + (a'-a)\,\mathrm{ht}_{\mathcal{V}}(x) \geq D_{a,\mathcal{V}}(x) + (a'-a)\,\mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}(x) \\ &= D_{a,\mathcal{V}}(x) + (a'-a)[K(x):K]\,\mathrm{ht}^{\mathrm{st;abs}}_{\mathcal{V}}(x). \end{aligned}$$

Since $\mathcal{V}$ is semipositive, for fixed a' and a the quantity

$$(a'-a)[K(x):K]\,\mathrm{ht}^{\mathrm{st;abs}}_{\mathcal{V}}(x) > (a'-a)\,\mathrm{ht}^{\mathrm{st;abs}}_{\mathcal{V}}(x)$$

is bounded below on $\mathcal{X}(\overline{K})$. It follows that if $D_{a,\mathcal{V}}$ is generically bounded below, so is $D_{a',\mathcal{V}}$. So the set of a such that $D_{a,\mathcal{V}}(x)$ is generically bounded below is an interval, extending infinitely in the positive direction.

Definition 4.13. With notation as above, the Fujita invariant $a(\mathcal{V})$ of a semipositive $\mathcal{V}$ is the infimum of all positive real numbers a such that $D_{a,\mathcal{V}}$ is generically bounded below. If $D_{a,\mathcal{V}}$ is never generically bounded below we take $a = \infty$.

The main goal of this section is to propose a heuristic for counting points of bounded height on stacks. If $\mathcal{X}$ is a stack over C, $\mathcal{U}$ is an open dense substack of $\mathcal{X}$ and $\mathcal{V}$ is a Northcott vector bundle (as in Definition 2.17) on $\mathcal{X}$, define a counting function

$$N_{\mathcal{U},\mathcal{V},K}(B) = |\{x \in \mathcal{U}(K) : \mathrm{ht}_{\mathcal{V}}(x) \leq \log B\}|.$$

The Batyrev–Manin conjecture is customarily stated for *Fano* varieties, those with ample anticanonical bundle. As mentioned above, it is not clear what the right analogue of this condition is for stacks. For instance, we certainly do not want to exclude stacks like BG, on which all vector bundles have degree 0 and are thus in some sense not 'strictly positive,' but we *do* want to exclude stacks like abelian varieties, whose anticanonical bundle is trivial. To this end, we make the following defintion. Let $\mathcal{X}$ is a smooth proper Deligne–Mumford stack over a number field K, let $m > 0$ and B be real numbers, and let $d \geq 1$ an integer. We then define $S(\mathcal{X}, m, d, B)$ to be the set of pairs (L, P) with L a degree-d extension and $P \in \mathcal{X}(L)$, satisfying

$$\mathrm{edd}(P) + m\Delta_{L/K} < B.$$

We provisionally say $\mathcal{X}$ is *Fanoish* if $S(\mathcal{X}, m, d, B)$ is finite for all m, d and B.

We are now ready to state the heuristic that motivates this part of the paper.

Conjecture 4.14. *Let K be a number field, and let $C = \mathrm{Spec}\,\mathcal{O}_K$. Let $\mathcal{X}$ be a stack over C whose generic fiber $\mathcal{X}_K$ is a smooth proper Deligne–Mumford stack over K. Suppose further that $\mathcal{X}_K$ is Fanoish and that $\mathcal{X}(K)$ is Zariski dense in $\mathcal{X}_K$. If $\mathcal{V}$ is a semipositive vector bundle on $\mathcal{X}$, then there exists an open dense substack $\mathcal{U}$ of $\mathcal{X}$ such that, for every $\epsilon > 0$, there is a nonzero constant c_ϵ such that*

$$c_\epsilon^{-1} B^{a(\mathcal{V})} \leq N_{\mathcal{U},\mathcal{V},K}(B) \leq c_\epsilon B^{a(\mathcal{V})+\epsilon},$$

where $a(\mathcal{V})$ is the Fujita invariant defined in Definition 4.13.

Remark 4.15. Our point of view throughout has been to let K be a global field of any characteristic; however, in Conjecture 4.14 we restrict to the case where K has characteristic 0. The reason for this is that we aim to emulate the Batyrev–Manin conjecture, and the form that conjecture should take for global fields of characteristic p is not fully settled. Indeed, there are counterexamples to the most naive

formulations of Batyrev–Manin, even for the anticanonical height; see Starr–Tian–Zong [67, Lemma 5.1] and recent work of Beheshti, Lehmann, Riedl and Tanimoto [7].

Remark 4.16. The condition that $\mathcal{X}(K)$ is Zariski dense is present to handle cases where, for instance, $\mathcal{X}(K)$ is empty or supported on a closed subvariety due to a local obstruction.

Remark 4.17 (Accumulating loci can be zero-dimensional). One difference between this case and the traditional Batyrev–Manin conjecture is that the accumulating locus $\mathcal{X}\backslash\mathcal{U}$ can be zero-dimensional; indeed, on a stacky $\mathbb{P}^1$, the stacky points are accumulating subvarieties. An example of this phenomenon can be seen in the recent paper of Pizzo, Pomerance and Voight [59], which counts points on the moduli stack $X_0(3)$ with respect to (in our language) the height arising from the Hodge bundle. They find that the preponderance of points are those supported at the single (stacky) point over $j = 0$, and compute a lower-order asymptotic for points on the complement $\mathcal{U}$ of this point.

Remark 4.18. Conjecture 4.14 corresponds to the weak version of the Batyrev–Manin conjecture. An analogue of the strong version would be an assertion that $N_{\mathcal{U},\mathcal{V},K}(B)$ is asymptotic to a constant multiple of $B^{a(\mathcal{V})}(\log B)^{b(\mathcal{V},K)}$ for some explicit constant $b(\mathcal{V}, K)$. Getting the power of $\log B$ correct (not even to speak of the constant!) is very subtle even in the Batyrev–Manin setting where $\mathcal{X}$ is a scheme; we will not attempt to pin it down here, but it seems a rich problem for further investigation.

Remark 4.19. One could, in the same way, propose a heuristic for counting points on $\mathcal{X}$ of bounded *stable* height. Just above, one could define $D^{\mathrm{st}}_{a,\mathcal{V}}(x)$ to be $\mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}(x) - \mathrm{edd}(x)$ and define the stable Fujita invariant to be the infimum of those a such that $D^{\mathrm{st}}_{a,\mathcal{V}}$ is generically bounded below. This gives nothing new in the case where $\mathcal{X}$ is a scheme (where stable height and height are the same) or where $\mathcal{X} = BG$ (in which case stable height is 0) but is of interest in other cases: See Section 3.6 for an example. In the same vein, and in some sense analogously to the central case of Batyrev–Manin where we count by anticanonical height, one could count the number of points x of $\mathcal{X}(K)$ with $\mathrm{edd}(x) < \log B$, even though edd is not always a height in the sense of this paper. One could reasonably expect this count to be bounded between constant multiples of B and $B^{1+\epsilon}$. For example, when $\mathcal{X} = BS_3$ and $K = \mathbf{Q}$, this would amount to counting cubic fields $L/\mathbf{Q}$ such that the product of the primes ramified in L is at most $\log B$. This counting problem will be addressed in forthcoming work of Shankar and Thorne, where it is shown that the count is on order $B \log B$.

4.3. The case where $\mathcal{X}$ is a scheme: the Batyrev–Manin conjecture

Suppose $\mathcal{X}$ is a scheme X. Then, since $\mathrm{ht}_{\mathcal{V}} = \mathrm{ht}_{\wedge^r\mathcal{V}}$ for any rank r vector bundle $\mathcal{V}$ on $\mathcal{X}$, we may assume $\mathcal{V}$ is a line bundle $\mathcal{L}$. We have seen in Remark 4.7 that $\mathrm{edd}(x) = \mathrm{ht}_{-K_X}(x)$ for any $x \in X(\overline{K})$. So if X is Fano, it is Fanoish because the anticanonical height is an ample height and thus has the Northcott property. It is not immediately obvious that a Fanoish scheme is Fano, but it is also not unreasonable to hope so. To begin, $-K_X$ is nef: if there were a curve C on X with $-K_X|_C$ of negative degree, then for some d, there is a degree-d map $C \to \mathbb{P}^1$ which provides many degree-d algebraic points with more and more negative $-K_X$-height, not counteracted by $m\Delta_{L/K}$ if we make m small enough. We also note that a variety with trivial canonical sheaf may be expected *not* to be Fanoish; a K3 surface, for instance, is expected (though not in general known) to have a Zariski-dense set of points over some extension L of K, which implies that X is non-Fanoish since all these point have $-K_X$-height 0 and $\Delta_{L/K}$ fixed.

The question of which schemes 'should' satisfy the Batyrev–Manin conjecture is not wholly understood but is probably not limited to Fano schemes alone; if it turns out that 'Fanoish' delineates a class of schemes including some to which Batyrev–Manin does not apply, we will narrow the notion.

The condition that $\mathcal{L}$ is semipositive simply says that $\mathcal{L}$ is nef; a nef height is bounded below, and if $\mathcal{L}$ is not nef, there is a curve on which $\mathcal{L}$ has negative degree, whose $\overline{K}$-points thus have heights which are not bounded below.

Now,

$$D_{a,\mathcal{L}}(x) = a\,\mathrm{ht}_{\mathcal{L}}(x) - \mathrm{edd}(x) = \mathrm{ht}_{a\mathcal{L}+K_X}(x)$$

and $a(\mathcal{L})$ is the minimal a such that $\mathrm{ht}_{a\mathcal{L}+K_X}(x)$ is generically bounded below.

What does this say about the line bundle $a\mathcal{L}+K_X$? First of all, if M is a big line bundle on X, then the map $\phi_k : X \to \mathbb{P}^{N_k}$ induced by the global sections of $\mathcal{L}^k$ is a birational embedding for some sufficiently large k. It is then immediate that the absolute height $\mathrm{ht}_M(x)$ is bounded below on $X(\overline{K})$ away from the locus Z contracted by ϕ_k and that there are only finitely many points of $X(K)\backslash Z(K)$ with height below any given bound. So h_M is generically bounded below. On the other hand, the pseudoeffective cone is dual to the cone of *moving* curves by by a theorem of Boucksom, Demaily, Paŭn and Peternell [14, Th 0.2] (see [30, Th 2.22] for the case of characteristic p). So if M is not pseudoeffective, there is a moving curve Y on X on which M has negative degree; if Z is any closed locus, we can move Y to not be contained in Z, and then $Y(\overline{K})$ has points away from Z of arbitrarily negative height; in particular, h_M is not generically bounded below.

Since the pseudoeffective cone is the closure of the big cone, we conclude that the infimum of a such that $\mathrm{ht}_{a\mathcal{L}+K_X}(x)$ is generically bounded below is the same as the infimal a such that $a\mathcal{L}+K_X$ is pseudoeffective, which is the same as the infimal a such that $a\mathcal{L}+K_X$ is big. And this $a(\mathcal{L})$ is just the usual Fujita invariant appearing in the Batyrev–Manin conjecture for Fano varieties. So Conjecture 4.14 recovers the (weak form of the) Batyrev–Manin conjecture.

4.4. The case where $\mathcal{X}$ is BG: Malle's conjecture

Now, suppose $\mathcal{X} = BG$ over a number field K, and $\mathcal{V}$ is a vector bundle, that is, a representation of G. In particular, let us assume $\mathcal{V}$ is a faithful permutation representation corresponding to an embedding $G < S_n$. Each point x of BG corresponds to a G-extension of K (possibly an étale algebra), and, via the embedding of G into S_n, a degree-n extension L/K whose Galois closure is G. We have already computed that

$$\mathrm{ht}_{\mathcal{V}}(x) = (1/2)\log|\Delta_{L/K}| = \sum_{v\in R} e_v \log q_v,$$

where R is the set of non-Archimedean places of K ramified in L/K, and e_v is the local degree of the discriminant. If v is a place where L/K is tamely ramified, so that tame inertia acts on $\{1,\dots,n\}$ through a cyclic subgroup $\langle\pi\rangle < S_n$, the ramification e_v is just the *index* $\mathrm{ind}(\pi)$, the difference between n and the number of orbits of π.

First of all, note that $\mathcal{V}$ is semipositive since $\mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}$ is identically 0.

It follows from Remark 4.7 that for any extension E/K and any point $x \in BG(E)$ corresponding to a degree-n extension F/E, we have

$$\mathrm{edd}(x) = \sum_v \log q_v,$$

where the sum is over non-Archimedean places v of E which are ramified in F/E. Note in particular that, because this is positive, BG is Fanoish; the set of $(L, x \in BG(L))$ with $\mathrm{edd}(x) + m\Delta_{L/K} < B$ involves only the finite set of extensions L/K with discriminant at most B/m, and for each L, the set of $x \in BG(L)$ with $\mathrm{edd}(x) < B$ is finite since it consists of G-extensions of L with bounded discriminant.

Thus,

$$D_{a,\mathcal{V}}(x) = a\,\mathrm{ht}_{\mathcal{V}}(x) - \mathrm{edd}(x) = \sum_v ((1/2)ae_v - 1)\log q_v.$$

Suppose $a \geq 2\max_{\pi\in G}\mathrm{ind}(\pi)^{-1}$. Then $(1/2)ae_v - 1 \geq 0$ for all tame primes v. The contribution of nontame primes is bounded below by a constant depending only on $[E : K]$. Thus, the Fujita invariant of $\mathcal{V}$ is at most $2\max_{\pi\in G}\mathrm{ind}(\pi)^{-1}$.

Suppose, on the other hand, that a is strictly smaller than $2\,\mathrm{ind}(\pi)^{-1}$ for some $\pi \in G$. If E/K is an extension of K and L/E a G-extension such that every ramified prime is tame and has tame inertia acting via π, then the point x has

$$\mathrm{edd}(x) = \sum_v ((1/2)ae_v - 1)\log q_v$$

which is bounded above by a *negative* constant multiple of $\sum_v \log q_v$. Heuristically, it seems safe to suppose one can choose such (E, L) with $\sum \log q_v$ as large as one likes, which would mean that $D_{a,\mathcal{V}}$ was not generically bounded below. But this is perhaps not completely obvious: For instance, when $G = S_n$, one is saying that there are many field extensions with squarefree discriminant. One certainly expects this to be the case, but the fact, for example, that there are arbitrarily large squarefree integers which are discriminants of degree-n extensions of $\mathbf{Q}$ is a recent result of Kedlaya [40]. In fact, all we need is that for *some* extension K' of K there are extensions L/K' with larger and larger discriminants whose ramification is entirely or almost entirely drawn from the minimal-index conjugacy class in G. One can presumably construct such extensions using the method of regular extensions popular in work on the inverse Galois problem; using the Riemann existence theorem you write down a cover of curves $X \to \mathbb{P}^1_K$ with Galois group G and all ramification drawn from the minimal-index conjugacy class, then descend the picture to $X_0 \to \mathbb{P}^1_{K'}$ for some finite extension K'/K, then show that specialization to points of $\mathbb{P}^1(K')$ yields many extensions of K' with the desired properties. Since we are just formulating conjectures here, we will not push this argument through in detail.

An argument of the sort sketched in the above paragraph is necessary due to the fact that we defined the Fujita invariant in terms of heights of points over extension fields of K; presumably, a more conceptual geometric definition of the Fujita invariant of a vector bundle with zero stable height would automatically assign $\mathcal{V}$ the value $2\max_{\pi\in G}\mathrm{ind}(\pi)^{-1}$.

At any rate, if we grant the heuristic argument on the Fujita invariant above, we find that Conjecture 4.14 predicts that the number of degree-n extensions L/K with Galois group G and discriminant at most B—in other words, the number of points x on $BG(K)$ such that

$$\mathrm{ht}_{\mathcal{V}}(x) = (1/2)\log|\Delta_{L/K}| < (1/2)\log B$$

is bounded between $c_\epsilon^{-1}B^a$ and $c_\epsilon B^{a+\epsilon}$, where $a = \max_{\pi\in G}\mathrm{ind}(\pi)^{-1}$. This is exactly the weak Malle conjecture.

Remark 4.20. When $\mathcal{V}$ is a representation of G which is not a permutation representation, one still has some conjugacy-invariant function f from G to $\mathbf{R}_{>0}$ and an expression

$$\mathrm{ht}_{\mathcal{V}}(x) = \sum_{v\in R} c_v \log q_v,$$

where, for every tamely ramified prime v, the coefficient c_v is the value of f at an element of G generating the tame inertia group at v. In this case, Conjecture 4.14 asserts that the number of points $x \in BG(X)$ with $\mathrm{ht}_{\mathcal{V}}(x) < \log B$ should be on order B^a, where a is the reciprocal of the minimal value taken by $f(v)$ on nonidentity elements of G. Heuristics of this kind are well-known folk generalizations of Malle (see, e.g., [28, §4.2]) and have begun to be proved in some cases. For instance, the striking work of Altüg, Shankar, Varma and Wilson [5] can be thought of as proving Conjecture 4.14 in the case where $\mathcal{X} = BD_4$ and $\mathcal{V}$ corresponds to the two-dimensional action of D_4 by rigid motions of the square. (What they prove is much more refined than what Conjecture 4.14; they not only compute the power of B, but the power of $\log B$, and even the constant!)

The recent work of Alberts [4] on counting classes in $H^1(\mathrm{Gal}(\mathbf{Q}), A)$, where A is an abelian group with Galois action, can perhaps also be thought of in this way. Here, A corresponds to an étale but possibly nonconstant group scheme, so the stack BA is *geometrically* the classifying stack of the finite abelian group underlying A. In this case, the points of $BA(\mathbf{Q})$ are just the classes in $H^1(\mathrm{Gal}(\mathbf{Q}), A)$.

The "π-discriminant" of [4, Lemma 1.4] is the height attached to the vector bundle on BA descended from the regular representation of the finite group underlying A.

4.5. Symmetric powers of $\mathbb{P}^n$

Let $\mathcal{X}$ be the stack $\mathrm{Sym}^m \mathbb{P}^n = [(\mathbb{P}^n)^m/S_m]$, and let K be a global field of characteristic 0 or greater than m. For x a point of $\mathcal{X}(K)$, we have

$$\mathrm{edd}(x) = -\mathrm{ht}_{T_{\mathcal{X}}^\vee}(x) + \mathrm{rDisc}(x).$$

Note that we can associate to x a degree-m extension L_1 of K and a point y of $\mathbb{P}^n(L_1)$.

The cotangent bundle $T_{\mathcal{X}}^\vee$, considered as an S_m-equivariant bundle on $(\mathbb{P}^n)^m$, is the direct sum of the m pullbacks of the cotangent bundle from the m projections $\mathbb{P}^n$, and the height associated to the cotangent bundle on $\mathbb{P}^n$ is just the usual height associated to its determinant $\mathcal{O}(-n-1)$. So we are in the situation of Section 3.6, and we have

$$\mathrm{ht}_{T_{\mathcal{X}}^\vee}(x) = \mathrm{ht}_{\mathcal{O}(-n-1)}(y) + (n/2)\log \Delta_{L_1/K}.$$

Thus,

$$\mathrm{edd}(x) = \mathrm{ht}_{\mathcal{O}(n+1)}(y) + \sum_{v\in R}(1-(n/2)e_v)\log q_v,$$

where, as in §4.4, R is the set of tamely ramified places and e_v is the power of v in the discriminant of L_i/K; the contribution of the wildly ramified places, as in Section 4.4, is bounded by a constant (and if x varies over $\mathcal{X}(L)$ for some extension E/K, the wild contribution is bounded by a constant depending only on $[E:K]$).

We also have

$$\mathrm{ht}_{T_{\mathcal{X}}}(x) = \mathrm{ht}_{\mathcal{O}(n+1)}(y) + (n/2)\log \Delta_{L_1/K} = \mathrm{edd}(x) + \sum_{v\in R}(ne_v - 1)\log q_v.$$

In particular, $\mathrm{ht}_{T_{\mathcal{X}}}(x) - \mathrm{edd}(x)$ is always nonnegative, and $\mathrm{ht}_{T_{\mathcal{X}}}(x) = \mathrm{edd}(x)$ whenever x is a point of $\mathcal{X}$ in the image of the projection from $(\mathbb{P}^n)^m(K)$ to $\mathrm{Sym}^m \mathbb{P}^n(K)$. This shows that the Fujita invariant $a(T_{\mathcal{X}})$ is 1. Conjecture 4.14 thus suggests that, away from some proper closed substack, the number of rational points on $\mathrm{Sym}^m \mathbb{P}^n(X)$ with tangential height at most B is between $B^{1-\epsilon}$ and $B^{1+\epsilon}$.

There is a large existing literature about counting points on projective spaces of fixed algebraic degree and bounded height [63, 31, 49, 50, 73, 46, 36, 37]. Most typically, the question being asked is: How many points are there in $\mathbb{P}^n(\overline{K})$ which have absolute Weil height at most B and which are defined over a field $L_1/\mathcal{K}$ of degree m? As we have seen in §3.6, we can interpret this question as follows. Let $\mathcal{V}$ be the vector bundle on $\mathrm{Sym}^m \mathbb{P}^n$ obtained as in §3.6 taking V_0 as $\mathcal{O}_{\mathbb{P}^n}(1)$. If y is a point of $\mathbb{P}^n(L_1)$ and x the corresponding point of $\mathrm{Sym}^m \mathbb{P}^n$, we have

$$\mathrm{ht}^{\mathrm{abs}}_{\mathcal{O}(1)}(y) = m^{-1}\,\mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}(x).$$

So we are in the situation of Remark 4.19. In order to compute the stable Fujita invariant of $\mathcal{V}$, we need to study the function

$$D^{\mathrm{st}}_{a,\mathcal{V}}(x) = a\,\mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}(x) - \mathrm{edd}(x) = (a-n-1)\,\mathrm{ht}_{\mathcal{O}(1)}(y) - \sum_{v\in R}(1-(n/2)e_v)\log q_v.$$

When $n \geq 2$, we note that the local term $\sum_{v\in R}(1-(n/2)e_v)\log q_v$ is always nonpositive and is 0 when L_1 is K^m; in particular, the set of x in $\mathrm{Sym}^m \mathbb{P}^n(K)$ with $\mathrm{edd}(x) = (a-n-1)\,\mathrm{ht}^{\mathrm{st}}_{\mathcal{V}}(x)$ is Zariski dense for every K. Thus, $D^{\mathrm{st}}_{a,\mathcal{V}}$ will be generically bounded below for any $a \geq n+1$ but is not generically

bounded below for any smaller a. So the stable Fujita invariant is $n+1$. For each y in $\mathbb{P}^n(\overline{K})$ with $[K(y):K] = m$, we write x_y for the point of $\mathrm{Sym}^m \mathbb{P}^n$. Then Conjecture 4.14 suggests that for every $n \geq 2$ we should expect that, for some open dense $U \in \mathrm{Sym}^m \mathbb{P}^n$,

$$c_\epsilon^{-1} B^{m(n+1)} < \#\{y \in \mathbb{P}^n(\overline{K}) : [K(y):K] = m, , x_y \in U(K), \mathrm{ht}^{\mathrm{abs}}(y) < B\} < c_\epsilon B^{m(n+1)+\epsilon}.$$

When $n = 1$, the situation is more complicated. We now have

$$D^{\mathrm{st}}_{a,\mathcal{V}}(x) \geq (a-2)\,\mathrm{ht}_{\mathcal{O}(1)}(y) - \sum_{v \in R} (1/2) e_v \log q_v = (a-2)\,\mathrm{ht}_{\mathcal{O}(1)}(y) - (1/2)\log \Delta_{L_1/K}$$

with equality when L_1/K has squarefree discriminant. In order to understand how large a needs to be for $D^{\mathrm{st}}_{a,\mathcal{V}}(x)$ to be generically bounded below, we need to know how large $\log \Delta_{L_1/K}$ can be relative to $\mathrm{ht}_{\mathcal{O}(1)}(y)$. A point y of $\mathbb{P}^1(L_1)$ has a minimal binary m-ic form $F = a_0 X^m + \cdots + a_m Y^m$, where the height of the point $(a_0 : \ldots : a_m)$ in $\mathbb{P}^m(L)$ is on order $m\,\mathrm{ht}(y)$ since each coefficient is a monomial of degree m in the coordinates of y. The discriminant of L_1/K is at most the discriminant of F, with equality if $\mathrm{disc}\, F$ is squarefree. The discriminant of F is a product of $m(m-1)$ terms of the form $\alpha_i\beta_j - \alpha_j\beta_i$, where $(\alpha_i : \beta_i)$ and $(\alpha_j : \beta_j)$ are conjugates of y in $\mathbb{P}^1(\overline{K})$. So the log of $\mathrm{disc}\, F$, considered as an element of $\mathcal{O}_L$, is on order $2m(m-1)\,\mathrm{ht}_L(y)$ and the log of $\mathrm{disc}\, F$ considered as an element of $\mathcal{O}_K$ is thus $2(m-1)\,\mathrm{ht}(y)$. We conclude that

$$D^{\mathrm{st}}_{a,\mathcal{V}}(x) \geq (a-2)\,\mathrm{ht}_{\mathcal{O}(1)}(y) - (m-1)\,\mathrm{ht}_{\mathcal{O}(1)}(y) = (a-m-1)\,\mathrm{ht}_{\mathcal{O}(1)}(y).$$

So $D^{\mathrm{st}}_{a,\mathcal{V}}$ is generically bounded below when $a \geq m+1$, and as long as there is a Zariski-dense set of choices of y with $\mathrm{disc}\, F$ squarefree (perhaps this is obvious, but at any rate it follows from standard conjectures) $D_{a,\mathcal{V}}$ is *not* generically bounded below for any smaller a. So the stable Fujita invariant in this case is $m+1$ and Conjecture 4.14 asserts that, for some open dense U,

$$c_\epsilon^{-1} B^{m(m+1)} < \#\{y \in \mathbb{P}^1(\overline{K}) : [K(y):K] = m, , x_y \in U(K), \mathrm{ht}^{\mathrm{abs}}(y) < B\} < c_\epsilon B^{m(m+1)+\epsilon}. \quad (4.21)$$

In fact, equation (4.21) follows from a theorem of Masser and Vaaler [50], who prove a much more refined asymptotic, with U the whole of $\mathrm{Sym}^m \mathbb{P}^1$:

$$\#\{y \in \mathbb{P}^1(\overline{K}) : [K(y):K] = m, \mathrm{ht}^{\mathrm{abs}}(y) < B\} \sim A_{m,K} B^{m(m+1)}$$

with an explicit constant $A_{m,K}$. Of course to compute the constant in the case where K is a number field, one has to be careful about the metrization on $\mathcal{O}(1)$ in a way we are not attempting here. Le Rudulier [46] generalized the Masser–Vaaler result to the case of an arbitrary metrized line bundle on $\mathbb{P}^1$.

When $n \geq 2$, the asymptotics for points of bounded height on projective n-space with algebraic degree m is still the subject of active research. If n is large enough relative to m, the heuristic (4.5) is known to be correct; indeed, one has

$$\#\{y \in \mathbb{P}^n(\overline{K}) : [K(y):K] = m, \mathrm{ht}^{\mathrm{abs}}(y) < B\} \sim A_{m,n,K} B^{m(n+1)}$$

when K is a number field and $n > (5/2)m + O(1)$, by a result of Widmer [73] and when $n > m+1$ with m prime by a result of Guignard [37]. For the function field case, the result is proved by Thunder and Widmer [69] when $n > 2m+4$ (and generalized from $\mathbb{P}^n$ to smooth projective toric varieties by Bourqui in [15]). Schmidt in [64] showed that equation (4.5) holds in case $K = \mathbf{Q}$, $m = 2$ and $n = 2$; indeed, in that case, the growth rate is $B^6 \log B$, showing that the ϵ in the exponent is sometimes necessary. Mânzăteanu [48] extended Schmidt's result to function fields K of odd characteristic.

On the other hand, Schmidt in [63] gives a lower bound

$$\#\{y \in \mathbb{P}^n(\overline{K}) : [K(y):K] = m, \mathrm{ht}^{\mathrm{abs}}(y) < B\} > A_{m,n,K} B^{m(m+1)}$$

valid for all n and all sufficiently large B. When $m > n$, this is a larger exponent than that predicted in equation (4.5). But this does not contradict Conjecture 4.14. The source of Schmidt's lower bound is the simple observation that any choice of line in $\mathbb{P}^n$ yields an injection of $\mathrm{Sym}^m \mathbb{P}^1(K)$ into $\mathrm{Sym}^m \mathbb{P}^n(K)$, and the former already contains $B^{m(m+1)}$ points of height at most B. But any such point lies on the proper closed substack $Z \subset \mathrm{Sym}^m \mathbb{P}^n(K)$ lying under the locus in $(\mathbb{P}^n)^m$ parametrizing ordered m-tuples of *collinear* points. Thus, it remains possible that when some accumulating locus is removed, the asymptotic growth rate of the number of points is smaller. And indeed, Guignard [37, Theorem 1.2.3] shows exactly this in the case where K is a number field, $m = 3$ and $n = 2$. In this setting, Schmidt's lower bound shows that the number of cubic points on $\mathbb{P}^2$ with absolute height at most B is at least cB^{12}. Guignard shows that if you exclude those cubic points which lie on a K-rational line, the number of rational points that remain is bounded above by $c_\epsilon B^{9+\epsilon}$, precisely the exponent predicted by Conjecture 4.14.

We thus see that the present viewpoint is useful for understanding phenomena of accumulation in a uniform way. The algebraic points witnessing Schmidt's lower bound are clearly 'nongeneric' in some sense, but, considered as points of $\mathbb{P}^n(\overline{K})$, they are Zariski dense. Considering these points instead as points on $\mathrm{Sym}^m \mathbb{P}^n$ shows that the accumulation is a phenomenon that can be repaired by stripping out a proper closed subvariety, exactly as in the Batyrev–Manin setting. Of course, one does not need to invoke stacks to adopt this point of view—for instance, see §33.2 of Le Rudulier's thesis [46], where a degree-m algebraic point of $\mathbb{P}^2$ is thought of as a point on the coarse moduli scheme of $\mathrm{Sym}^m \mathbb{P}^2$ rather than the stack itself; since the two are birational, the observation that the collinear m-tuples lie on a subvariety on which rational points accumulate takes the same form for Le Rudulier as it does for us.

4.6. Footballs and multifootballs

Proposition 4.9 shows that edd agrees with tangential height $\mathrm{ht}_{T_\mathcal{X}}$ when $\mathcal{X}$ is a smooth proper one-dimensional stack over a number field K which is birational to a curve. In particular, Proposition 4.9 applies when $\mathcal{X}$ is a stacky curve birational to $\mathbb{P}^1$ which has r stacky points isomorphic to $B(\mu_{m_1}), \dots, B(\mu_{m_r})$. For short we will call such a curve an $(m_1, \dots, m_r)$-rooted $\mathbb{P}^1$. The football $\mathcal{F}(a, b)$ as in §3.5 is then an (a, b)-rooted curve.

Let $\mathcal{X}$ be an $(m_1, \dots, m_r)$-rooted $\mathbb{P}^1$. Now, Conjecture 4.14 predicts that, for some open dense $\mathcal{U}$ in $\mathcal{X}$, we have

$$c_\epsilon^{-1} B \le N_{\mathcal{U}, T_\mathcal{X}, K}(B) \le c_\epsilon B^{1+\epsilon}. \tag{4.22}$$

First of all, U is obtained by removing a finite set of points from $\mathcal{X}$, so we can interpret the above asymptotic as a heuristic for the number of points of $\mathcal{X}$ of bounded height which are not supported on the stacky locus.

The coarse map $\mathcal{X} \to \mathbb{P}^1$ is a birational isomorphism, and so without serious ambiguity we can denote a point x on $\mathcal{X}(K)$ not contained in stacky locus by its image $(a : b)$ in $\mathbb{P}^1(K)$. We will now compute tangential height explicitly. The tangent sheaf $T_\mathcal{X}$ is $2P + \sum_i (1/m_i - 1)P_i$, where P_i is the i'th stacky point and P is some other point on $\mathcal{X}$; the degree of $T_\mathcal{X}$ is thus $d = 2 - r + \sum_i (1/m_i)$. If N is an multiple of every m_i, then $NT_\mathcal{X}$ is linearly equivalent to Nd copies of P; in other words, it is pulled back from $\mathcal{O}(Nd)$ on the coarse space $\mathbb{P}^1$. We thus have

$$\mathrm{ht}^{\mathrm{st}}_{T_\mathcal{X}}(x) = (1/N)\,\mathrm{ht}^{\mathrm{st}}_{NT_\mathcal{X}}(x) = (1/N)\,\mathrm{ht}_{\mathcal{O}(Nd)}(a : b) = d\,\mathrm{ht}_{\mathcal{O}(1)}(a : b).$$

We note, in particular, that $T_\mathcal{X}$ is not semipositive unless $d \ge 0$, so we assume this from now on.

For expositional simplicity, we now restrict to the case $K = \mathbf{Q}$. So the stable height of x is $d \log \max(|a|, |b|)$, where a and b are now taken to be coprime integers. It remains to compute the local discrepancies. The local discrepancy $\delta_v(a : b)$ can be computed as follows. The tangent bundle $T_\mathcal{X}$ has local degree $1/m_i \in \mathbf{Q}/\mathbf{Z}$ at P_i, so the degree of $\overline{x}^* T_\mathcal{X}^\vee$ at the point of the tuning stack $\mathcal{C}$ over a

place v is $-k/m_i$, where $k = \operatorname{ord}_v L_i(a:b)$. Thus, the local degree of the pushforward $\pi_*\overline{x}^*T_{\mathcal{X}}$ on C is $\lfloor -k/m_i \rfloor = -\lceil k/m_i \rceil$, and so the local discrepancy is given by

$$\delta_v = (\lceil k/m_i \rceil - k/m_i) \log q_v.$$

Throw out the bounded contribution of any prime v where two distinct P_i intersect, and denote by L_i the linear form whose zero is at P_i. Then for each prime p, there is at most one $L_i(a,b)$ vanishing at p, and the local discrepancy is $(1/m_i)\log p^c$, where c is the least integer such that the p-adic valuation of $p^c L_i(a,b)$ is a multiple of m_i.

Definition 4.23. For integers m, N, define $\Phi_m(N)$ to be the unique m-th power free integer such that $N\Phi_m(N)$ is an mth power. Alternatively,

$$\Phi_m(N) = \prod_p p^{m\lceil \operatorname{ord}_p N/m \rceil - \operatorname{ord}_p N}.$$

When $m = 2$, we have that $\Phi_2(N)$ is the *squarefree part* of N, denoted $\operatorname{sqf}(N)$.

Putting this all together, we find

$$\operatorname{ht}_{T_{\mathcal{X}}}(a:b) = \sum_i (1/m_i)\log \Phi_{m_i}(L_i(a,b)) + (2 - r + \sum_i 1/m_i)\log\max(|a|,|b|).$$

When r is small, it is straightforward to see that equation (4.22) is satisfied. For example, consider a $\mathbb{P}^1$ rooted only at 0 with a copy of $B\mu_3$ (that is, $r = 1$ and $m_1 = 3$). Then (taking $\mathcal{U}$ to be the complement of the stacky locus) $N_{\mathcal{U},T_{\mathcal{X}},K}(B)$ is the number of pairs of coprime a, b such that

$$\Phi_3(a)^{1/3}\max(|a|,|b|)^{4/3} < B.$$

We can write a uniquely as $c^3 d_1 d_2^2$, where d_1, d_2 are coprime and squarefree, and clearly bounded above by a power of B. Then $\Phi_3(a) = d_1^2 d_2$ and we find that up to constants we are counting the positive c, d_1, d_2, b such that

$$d_1^{2/3} d_2^{1/3}\max(c^4 d_1^{4/3} d_2^{8/3}, b^{4/3}) = \max(c^4 d_1^2 d_2^3, b^{4/3} d_1^{2/3} d_2^{1/3}) < B.$$

For a given choice of coprime d_1, d_2, we see that the number of choices for c is $B^{1/4} d_1^{-1/2} d_2^{-3/4}$, while the number of choices for b is $B^{3/4} d_1^{-1/2} d_2^{-1/4}$, so the number of choices for the pair (c,b) is just $B d_1^{-1} d_2^{-1}$; summing this over all coprime pairs d_1, d_2 up to some power of B gives an asymptotic for $N_{\mathcal{U},T_{\mathcal{X}},K}(B)$ on order $B\log^2 B$, which agrees with the heuristic prediction (4.22).

John Yin has shown (personal communication) that equation (4.22) holds for a $(2,2)$-rooted $\mathbb{P}^1$; in fact, he addresses the more general case where the degree-2 stacky locus is irreducible over $\mathbf{Q}$ rather than being supported at two rational points, as in the cases discussed here.

Things get more difficult as r grows. Consider the case of a $(2,2,2)$-rooted $\mathbb{P}^1$ with the half-points located at $0, -1$, and ∞. Then

$$\operatorname{ht}_{T_{\mathcal{X}}}(a:b) = (1/2)\log(\operatorname{sqf}(a)\operatorname{sqf}(b)\operatorname{sqf}(a+b)\max(|a|,|b|))$$

so $N_{\mathcal{U},T_{\mathcal{X}},K}(B)$ is the number of pairs of coprime a, b such that

$$\operatorname{sqf}(a)\operatorname{sqf}(b)\operatorname{sqf}(a+b)\max(|a|,|b|) < B^2.$$

This set contains all pairs of coprime integers in $[0, \sqrt{B}]$, so it has size at least cB, as predicted.

In fact, in recent work, Pierre Le Boudec (in personal communication) and Nasserden–Xiao [53] have independently shown that $N_{\mathcal{U},T_{\mathcal{X}},K}(B)$ is bounded above and below by constant multiples of $B\log B^3$.

This seems a very interesting case to explore further; can one obtain an asmyptotic $N_{\mathcal{U},T_{\mathcal{X}},K}(B) \sim cB\log^3 B$, and if so, what is the constant?

We also note that some footballs are weighted projective lines; in recently announced work, Darda [20] proves counting results for weighted projective spaces.

4.7. When edd *is negative: A stacky Lang–Vojta conjecture*

Conjecture 4.14 is meant to apply to those 'Fanoish' stacks $\mathcal{X}$, where edd is positive in some appropriate sense. In this section, we consider the opposite scenario: where $\mathrm{edd}(x)$ is negative. When X is a scheme, this is the situation where the canonical bundle K_X is ample so that X is of general type; in this case, and assuming K is a number field, Lang's conjecture suggests that $X(K)$ should be supported on a proper closed subvariety of X. (When K is a global field of characteristic p, the situation is more subtle—the famous examples of Shioda show, for instance, that a variety can be of general type and also unirational! We thus restrict to the number field case for the remainder of the discussion.)

More precisely, conjectures of Vojta say that, for any X, any ample line bundle L, and any real $\delta > 0$, the set of rational points on $X(K)$ such that

$$\mathrm{ht}_{-K_X}(x) + \delta\,\mathrm{ht}_L(x) < 0$$

should be supported on a proper closed subvariety.

This suggests that one might tentatively propose a 'Vojta conjecture for stacks' as follows: Let $\mathcal{X}$ be a stack over a number field K, let L be a line bundle on $\mathcal{X}$ pulled back from an ample line bundle on the coarse space of $\mathcal{X}$ and let $\delta > 0$ a real number.

Conjecture 4.24. *The set of rational points of $\mathcal{X}(K)$ such that*

$$\mathrm{edd}(x) + \delta\,\mathrm{ht}_L(x) < 0$$

is supported on a proper closed substack of $\mathcal{X}$.

For example, if $\mathcal{X}$ is a $(4,4,4)$-rooted $\mathbb{P}^1$ with the $(1/4)$th-points at $0, 1, \infty$, then we have

$$\mathrm{edd}(a:b) = \log \Phi_4(a)^{1/4}\Phi_4(b)^{1/4}\Phi_4(a+b)^{1/4}\max(|a|,|b|)^{-1/4}$$

and the claim is then that the inequality

$$\Phi_4(a)\Phi_4(b)\Phi_4(a+b) < \max(|a|,|b|)^{1-\delta}$$

holds for only finitely many pairs of coprime integers a, b.

Another interesting case is that of a $(2,2,2,2,2)$-rooted $\mathbb{P}^1$ with the half-points at $0, 1, 2, 3, 4$. In this case, Conjecture 4.24 says there are only finitely many five-term arithmetic progressions $a_1, \ldots, a_5$ such that

$$\mathrm{sqf}(a_1a_2a_3a_4a_5) < \max(a_1, a_5)^{1-\delta}.$$

As Nasserden and Xiao explain in [54, Theorem 1.4], the assertion that Conjecture 4.24 holds for all stacky curves is equivalent to the abc conjecture, with a key ingredient being a result of Granville [34]; indeed, Granville's result shows immediately that the two examples above satisfy Conjecture 4.24 conditional on abc. What is the relation between Vojta's 'more general abc conjecture' from [72] applied to a divisor D on a scheme X, and Conjecture 4.24 for a stack obtained by rooting a scheme X at D?[6] One may hope that individual cases of Conjecture 4.24, like those described above, might not be as far out of reach as abc and its generalizations.

[6]We are grateful to Aaron Levin for useful discussions concerning this connection.

We note that a conjecture akin to Conjecture 4.24 also appears in the work of Abramovich and Várilly [1, Proposition 3.2]; they show their conjecture follows from the Vojta conjecture for schemes and derive from this a finiteness theorem, conditional on Vojta, for principally polarized abelian varieties with full m-level structure for large enough m. Their conjecture is expressed in terms of a height on $\mathcal{X}$ which, in the language of this paper, is $\mathrm{ht}^{\mathrm{st}}_{-K_X}$. And their conjecture, like Conjecture 4.24, can be expressed as an assertion that the set of points $x \in \mathcal{X}(K)$ with

$$\mathrm{ht}^{\mathrm{st}}_{-K_X}(x) + \delta\,\mathrm{ht}_L(x) + \sum_v \alpha_v(x) < 0$$

is not Zariski dense, for some local nonnegative contributions α_v supported at the points where x fails to extend to an integral point of $\mathcal{X}$. (In fact, their conjecture says more, making an assertion about all algebraic points of a fixed degree r.) The conjecture of Abramovich and Várilly is compatible with Conjecture 4.24 but is not identical to it. One interesting case where they differ is that of $\mathcal{X} = A/\pm 1$, with A an abelian variety of dimension g over a number field K. Let x be a point of $(A/\pm 1)(K)$, which is to say a quadratic extension F/K and a point of $A(F)$ with trace zero in $A(K)$. The stable height can be computed on the pullback to the étale cover A, where the canonical divisor on $\mathcal{X}$ is zero, so the Abramovich–Várilly conjecture bounds the set of $x \in \mathcal{X}(K)$ such that

$$\delta\,\mathrm{ht}_L(x) + \sum_v \alpha_v(x) < 0.$$

But the left-hand side is positive for all but finitely many x by the ampleness of L, so this is easy. On the other hand, Conjecture 4.24 says more. We have

$$\mathrm{edd}(x) = -\,\mathrm{ht}_{T^\vee_{\mathcal{X}}} + \mathrm{rDisc}(x).$$

Near a stacky point v, the tuning stack looks like $[(\operatorname{Spec}\mathcal{O}_{F,v})/\pm 1]$ and Λ, as in §2.3, is given by $\mathcal{O}_{F,v}^{\oplus g}$, where the ± 1-action sends the i-th basis vector e_i to $-e_i$; hence, if we let $\overline{\alpha}$ denote the quadratic conjugate of $\alpha \in F_v$, we see αe_i maps to $-\overline{\alpha} e_i$. It follows that if v is not of characteristic 2, then $\Lambda \cap L_v$ is the set of sums $\sum_i \alpha_i e_i$ with α_i of trace zero. An easy computation then shows the local discrepancy at v is $(1/2)g \log q_v$. We conclude that

$$\mathrm{ht}_{T^\vee_{\mathcal{X}}} = \mathrm{ht}^{\mathrm{st}}_{T^\vee_{\mathcal{X}}} + (g/2)\log \mathrm{disc}_{F/K} = (g/2)\log \mathrm{disc}_{F/K}.$$

Furthermore (still setting aside the bounded contribution of 2), the conductor $|\operatorname{Supp} R_\pi|$ is just equal to $\log \mathrm{disc}_{F/K}$. So Conjecture 4.24 says that the set of x with

$$(1 - g/2)\log \mathrm{disc}_{F/K} + \delta\,\mathrm{ht}_L(x) < 0$$

is supported on a closed subvariety, for any real $\delta > 0$. When $g \geq 2$ this is vacuous, but when $g \geq 3$ it has content. By changing δ we can absorb the constant on the right-hand side and say that the prediction is as follows: For any abelian variety A/K of dimension at least 3, and any real $\delta > 0$, there is a closed subvariety $Z_\delta \subset A$ such that, for any trace-zero quadratic point $P \in A(\overline{K})\backslash Z(\overline{K})$, the absolute logarithmic height of $P \in A(\overline{K})$ is at least $\delta^{-1} \log \mathrm{disc}_{F/K}$.

This formulation may seem a bit cumbersome, but it is necessary. Suppose, for example, that A is the Jacobian of a hyperelliptic curve X over K, and suppose X has a rational Weierstrass point so it embeds into A via an Abel–Jacobi map. Then X provides many quadratic points P on A whose heights are bounded above by $c \log \mathrm{disc}_{K(P)/K}$ for some real c. So if $\delta < c^{-1}$, the exceptional set Z_δ needs to include X. But if we take $\delta < (1/m)c^{-1}$, then every quadratic point on A lying on the curve $[m]X$ satisfies $\log \mathrm{disc}_{K(P)/K} + \delta\,\mathrm{ht}(x) < 0$, so we need to include not only X but $[2]X, [3]X, \ldots, [m]X$ in the exceptional locus Z_δ. On the other hand, no matter what δ is, there should be many quadratic points in $A\backslash Z_\delta$ because (at least under modest assumption on A) the functional equation of quadratic twists

of A will vary in sign with the twist, which means there will be many quadratic twists A_d of A which under Birch–Swinnerton-Dyer have positive rank. The heuristics here would suggest that the nontorsion points on such an A_d have very large height relative to d. Is this reasonable?

4.8. Further questions

There are many questions about the subject matter here which in the interests of length and time we have not addressed.

- How does one compute $\mathrm{edd}(x)$ explicitly when K is the function field of a curve in finite characteristic and $\mathcal{X}$ is not tame?
- Is Conjecture 4.14 geometrically consistent in the sense of Lehmann, Sengupta and Tanimoto [47]?
- How should one estimate the asymptotic growth of points on $\mathcal{X}$ which are integral with respect to a divisor D?
- As mentioned earlier in the paper, one might, rather than defining height in terms of the degree of $\pi_*\overline{x}^*\mathcal{V}$, simply keep track of the vector bundle $\pi_*\overline{x}^*\mathcal{V}$ itself. When $K = \mathbf{Q}$ this metrized vector bundle is a lattice of the same rank as $\mathcal{V}$. When $\mathcal{X}$ is a scheme, this point of view has been advanced by Peyre [58] as a more refined means of studying rational points on varieties. When $\mathcal{X} = BG$ and $\mathcal{V}$ is a permutation representation of G, this lattice is related to the shape of the integer lattice in the G-extension $L/\mathbf{Q}$ corresponding to x; the variation of these lattices as one ranges over G-extensions of bounded discriminant has been an object of much recent interest [12, 39, 13]. What can be said about intermediate cases, like $\mathrm{Sym}^m \mathbb{P}^n$?

Appendix A. Metrized vector bundles on stacks over number fields

A.1. Linear Algebra

An Hermitian pairing on a complex vector space V is a bilinear map $\langle\,,\rangle\colon V \to \mathbb{C}$ such that for all $v, w \in V$, $\langle w, v\rangle = \overline{\langle v, w\rangle}$ (whence $\langle v, v\rangle \in \mathbb{R}_{\geq 0}$). We define the associated Hermitian norm $\|\cdot\|\colon V \to \mathbb{R}$ via $\|v\| := \sqrt{\langle v, v\rangle}$. We call such a pair $\overline{V} := (V, \|\cdot\|_V)$ (or equivalently, $(V, \langle\,,\rangle_V)$) an Hermitian space. For $r \in \mathbb{R}_{\geq 0}$ we define the ball of radius r to be $B\left(\overline{V}, r\right) := \{v \in V \nmid \|v\| \leq r\}$ (and refer to $B\left(\overline{V}, 1\right)$ as the unit ball in $\overline{V}$). We define the standard Hermitian space to be $\overline{\mathbb{C}^n} := (\mathbb{C}^n, \langle\,,\rangle)$, where $\langle x, y\rangle := \sum x_i \overline{y_i}$.

A morphism $\phi \in \mathrm{Hom}\left(\overline{V}, \overline{W}\right)$ of Hermitian spaces is a linear map $\phi\colon V \to W$ such that $\|\phi(v)\|_W \leq \|v\|_V$ for all $v \in V$. The space $\mathrm{Hom}(V, W)$ admits a pairing

$$\langle \phi, \psi\rangle := \sup_{v \in B(\overline{V},1)} \langle \phi(v), \psi(v)\rangle_W.$$

The associated norm is $\|\phi\| = \sup_{v \in B(\overline{V},1)} \|\phi(v)\|_W$; we let $\underline{\mathrm{Hom}}\left(\overline{V}, \overline{W}\right)$ be the associated Hermitian space, whence $\mathrm{Hom}\left(\overline{V}, \overline{W}\right) := B\left(\underline{\mathrm{Hom}}\left(\overline{V}, \overline{W}\right), 1\right)$. We define the dual $\overline{V}^\vee$ of $\overline{V}$ to be $\underline{\mathrm{Hom}}\left(\overline{V}, \overline{\mathbb{C}}\right)$.

Let $\overline{V}$ be an Hermitian space, and let $0 \to V' \to V \xrightarrow{\pi} V'' \to 0$ be an exact sequence of complex vector spaces. Then the restriction of $\|\cdot\|_V$ to V' is an Hermitian norm $\|\cdot\|_{V'}$ on V'. The orthogonal complement $(V')^\perp$ of V' is naturally identified with V'', inducing a pairing $\langle\,,\rangle_{V''}$ on V'' via restriction of $\langle\,,\rangle_V$ and this identification; the induced quotient norm $\|\cdot\|_{V''}$ on V'' can thus be computed as $\|v\|_{V''} = \inf_{w \in \pi^{-1}(v)} \|w\|_V$.

Let $\overline{V}$ and $\overline{W}$ be Hermitian spaces. We define the direct sum $\overline{V} \oplus \overline{W} := (V \oplus W, \|\cdot\|_{V \oplus W})$ via the declaration $\langle v, w\rangle_{V \oplus W} = 0$ for $v \in V, w \in W$; one then computes that $\|v \oplus w\|_{V \oplus W} = \sqrt{\|v\|_V^2 + \|w\|_W^2}$. We define the tensor product $\overline{V} \otimes \overline{W} := (V \otimes W, \|\cdot\|_{V \otimes W})$ via the formula $\langle v_1 \otimes w_1, v_2 \otimes w_2\rangle_{V \oplus W} = \langle v_1, v_2\rangle_V \cdot \langle w_1, w_2\rangle_W$; one then computes that $\|v \otimes w\|_{V \otimes W} = \|v\|_V \cdot \|w\|_W$. We define the alternating

product $\bigwedge^n \overline{V}$ via $\langle v_1 \wedge \cdots \wedge v_n, w_1 \wedge \cdots \wedge w_n \rangle = \det(\langle v_i, w_j \rangle)$; this is not exactly equal to the quotient norm of $\|\cdot\|_{V^{\otimes n}}$ along the map $V^{\otimes n} \to \bigwedge^n V$, but rather is $\sqrt{n!}$ times the quotient norm.

A.2. Analytic spaces

Let X be a complex analytic space (as in [35]), and let $\mathcal{V}$ be a vector bundle on X. Let $\mathcal{C}_X$ denote the sheaf of continuous functions on X valued in $\mathbb{R}_{\geq 0}$. An Hermitian norm $|\cdot|$ on $\mathcal{V}$ is a morphism of sheaves

$$|\cdot| \colon \mathcal{V} \to \mathcal{C}_X$$

such that

1. $|s|(x) = 0$ if and only if $s(x) = 0$,
2. for all $f \in \mathcal{O}_X(U)$, we have $|fs| = |f||s|$, and
3. for every complex point $x \colon * \to X$, the restriction of $|\cdot|$ to $x^*\mathcal{V}$ is Hermitian (when viewed as a norm on $H^0(*, x^*\mathcal{V})$),

where, in condition (2), $|f|$ is the trivial norm on the line bundle $\mathcal{O}_X$ (i.e., $f \in \mathcal{O}_X(U)$ corresponds to a continuous function $f \colon U \to \mathbb{C}$, and we define $|f| \colon U \to \mathbb{R}_{\geq 0}$ by $|f|(x) = |f(x)|$). We call such a pair $\overline{\mathcal{V}} := (V, |\cdot|)$ a metrized vector bundle on the analytic space X.

We define direct sums, tensor products, alternating products and duals via the formulas from (A.1) (locally, and if necessary, we sheafify); for example, given metrized vector bundles $(\mathcal{V}_1, |\cdot|_1)$ and $(\mathcal{V}_2, |\cdot|_2)$, we define

$$|\cdot| \colon \mathcal{V}_1 \oplus \mathcal{V}_2 \to \mathcal{C}_X,$$

as

$$|v_1 \oplus v_2|(x) := \left((|v_1|_1(x))^2 + (|v_2|_2(x))^2 \right)^{1/2}.$$

Given a morphism $g \colon X \to Y$ of analytic spaces and a metrized vector bundle $\overline{\mathcal{V}} = (\mathcal{V}, |\cdot|)$ on Y, we define the pull back $g^*\overline{\mathcal{V}}$ to be the pair $((g^*\mathcal{V}), g^*|\cdot|)$, where $g^*|\cdot|$ is adjoint to the composition

$$\mathcal{V} \to \mathcal{C}_Y \to g_*\mathcal{C}_X,$$

and where the second map is given by composition of functions. If g is unramified and finite (in particular, $g_*\mathcal{V}$ is a vector bundle), we define the direct image $g_*\overline{\mathcal{V}}$ to be the pair $((g_*\mathcal{V}), g_*|\cdot|)$, where $g_*|\cdot|$ is defined via the composition

$$g_*\mathcal{V} \to g_*\mathcal{C}_X \to \mathcal{C}_Y,$$

and where $g_*\mathcal{C}_X \to \mathcal{C}_Y$ is defined by summation on fibers; in other words, for an open subset $U \subset Y$ and a function $h \in \mathcal{C}_X(g^{-1}(U))$, we define a map $U \to \mathbb{R}_{\geq 0}$ via the formula $y \mapsto \sqrt{\sum_{x \in g^{-1}(y)} h(x)^2}$. For a complex point $x \colon * \to X$ with image $y \colon * \to Y$, the natural map $(g^*\mathcal{V})_x \to \mathcal{V}_y$ is an isomorphism, and the norm is 'the same' on these fibers. In contrast, the fiber $(g_*\mathcal{V})_y$ of the direct image is naturally isomorphic to $\oplus_{x \in g^{-1}(y)} \mathcal{V}_x$, and the norm on this fiber is the direct sum norm defined in (A.1).

A.3. Schemes

By a variety over S, we mean a scheme of finite type over S. To a variety X over $\operatorname{Spec} \mathbb{C}$ and vector bundle $\mathcal{V}$ on X, associate the complex analytification $(X^{\mathrm{an}}, \mathcal{V}^{\mathrm{an}})$ (as in [35]). (We note that one can also associate an analytic space, functorially, to an algebraic space which is locally separated and locally of

finite type over $\mathbb{C}$ [42, Ch. I, 5.17], and that the setup here extends to that generality without any further modification.)

Let K be a number field, let X be a $\operatorname{Spec}\mathcal{O}_K$ variety and let $\mathcal{V}$ on X be a vector bundle on X. For an embedding $\sigma\colon K \to \mathbb{C}$ (i.e., a map $\sigma\colon \operatorname{Spec}\mathbb{C} \to \operatorname{Spec}K$), we let $X_\sigma := X \times_{K,\sigma} \mathbb{C}$ and let $\mathcal{V}_\sigma$ denote the pullback of $\mathcal{V}$ to X_σ. We define a metrized vector bundle on X to be a vector bundle $\mathcal{V}$ together with a choice of Hermitian norm $|\cdot|_\sigma$ on $\mathcal{V}_\sigma^{\mathrm{an}}$ for every embedding $\sigma\colon K \to \mathbb{C}$, with the following property: for every Zariski open $U \subset X$ and section $s \in \mathcal{V}(U)$, we have $|\sigma^* s|_\sigma(p) = |\overline{\sigma}^* s|_{\overline{\sigma}}(\overline{p})$.

We define direct sums, tensor products, alternating products, and duals via the formulas from (A.2). Given a morphism $g\colon X \to Y$ of $\operatorname{Spec}\mathcal{O}_K$ varieties and an embedding $\sigma\colon K \to \mathbb{C}$, the diagram

$$\begin{array}{ccc} X_\sigma & \xrightarrow{g_\sigma} & Y_\sigma \\ \downarrow & & \downarrow \\ X & \xrightarrow{g} & Y \end{array}$$

commutes. Given a metrized vector bundle $\overline{\mathcal{V}} = (\mathcal{V}, |\cdot|)$ on Y, it follows that $(g^*\mathcal{V})_\sigma$ is canonically isomorphic to $g_\sigma^*(\mathcal{V}_\sigma)$, and we define the pull back $g^*\overline{\mathcal{V}}$ to have underlying vector bundle $g^*\mathcal{V}$ and metrics $g_\sigma^*|\cdot|_\sigma$ defined via (A.2). Similarly, if g is finite, flat, and generically étale (and in particular locally free so that $g_*\mathcal{V}$ is a vector bundle), we define the direct image $g_*\overline{\mathcal{V}}$ to have underlying bundle $g_*\mathcal{V}$ and metrics $g_{\sigma,*}|\cdot|_\sigma$ defined via (A.2).

There is an alternative type of direct image, which highlights the choice of base in our definition. Let $K \subset L$ be an inclusion of number fields. Let $X \to \operatorname{Spec}\mathcal{O}_L$ be an $\mathcal{O}_L$ variety, and let $\overline{\mathcal{V}}$ be a metrized vector bundle on X. We define the restriction of scalars of $(X, \overline{\mathcal{V}})$ to be the pair $(\operatorname{Res}_{L/K} X, \operatorname{Res}_{L/K} \overline{\mathcal{V}})$, where $\operatorname{Res}_{L/K} X$ is the usual restriction of scalars (i.e., X itself, viewed as an $\mathcal{O}_K$ variety via the composition $X \to \operatorname{Spec}\mathcal{O}_L \to \operatorname{Spec}\mathcal{O}_K$) and where $\operatorname{Res}_{L/K}\overline{\mathcal{V}}$ has the same underlying vector bundle $\mathcal{V}$ and is endowed with a metric in the following way. Given an embedding $\sigma\colon K \hookrightarrow \mathbb{C}$, the space $\left(\operatorname{Res}_{L/K} X\right)_\sigma$ is isomorphic to $\coprod_{\sigma'|\sigma} X_{\sigma'}$, where the coproduct is taken over the set of $\sigma'\colon L \hookrightarrow \mathbb{C}$ extending σ; similarly, $\left(\operatorname{Res}_{L/K}\mathcal{V}\right)_\sigma$ is the vector bundle whose restriction to $X_{\sigma'}$ is $\mathcal{V}_\sigma$ (note that, by the sheaf axioms, $\Gamma(X_\sigma, \mathcal{V}_\sigma) = \bigoplus_{\sigma'|\sigma} \Gamma(X_{\sigma'}, \mathcal{V}_{\sigma'})$), and the norm

$$|\cdot|_\sigma\colon \left(\operatorname{Res}_{L/K}\mathcal{V}\right)_\sigma^{\mathrm{an}} \to \mathcal{C}_{\left(\operatorname{Res}_{L/K} X\right)_\sigma^{\mathrm{an}}}$$

is the one whose restriction to $X_{\sigma'} \subset \left(\operatorname{Res}_{L/K} X\right)_\sigma$ is $|\cdot|_{\sigma'}$.

Similarly, if $K \hookrightarrow L$ is an extension of number fields, X is an $\mathcal{O}_L$ variety, and $\overline{\mathcal{V}}$ is a metrized vector bundle on X *considered as an* $\mathcal{O}_K$ variety (equivalently, a metrized bundle on $\operatorname{Res}_{L/K} X$), we define base extension $\overline{\mathcal{V}}_L$ as follows. The underlying bundle is $\mathcal{V}$; for a place σ' of L with restriction $\sigma := \sigma'|_K$, the map $\phi\colon X_{\sigma'} \to \operatorname{Res} X_\sigma$ of $\mathbb{C}$ varieties is an isomorphism, and we define $|\cdot|_{\sigma'}$ to be the same as $|\cdot|_\sigma$ (under the identification ϕ).

The degree of a metrized line bundle $(\mathcal{V}, |\cdot|)$ on $\operatorname{Spec}\mathcal{O}_K$ (considered as an $\mathcal{O}_K$-variety) is defined to be

$$\deg(\mathcal{V}, |\cdot|) = \log|\Gamma(\mathcal{V})/\mathcal{O}_K \cdot s| - \sum_{\sigma\colon K \to \mathbb{C}} \log|\sigma^* s|_\sigma, \tag{A.1}$$

where $s \in \Gamma(\mathcal{V})$ is any nonzero section. Implicit here is that this definition is independent of the choice of s. When $(\mathcal{V}, |\cdot|)$ is a metrized vector bundle of rank $r > 1$, the degree of $(\mathcal{V}, |\cdot|)$ is by definition the degree of the metrized vector bundle $\wedge^n(\mathcal{V}, |\cdot|)$. If $K \hookrightarrow L$ is an extension of number fields and $(\mathcal{V}, |\cdot|)$ is a metrized line bundle on $\operatorname{Spec}\mathcal{O}_L$ *considered as an* $\mathcal{O}_K$-variety, then we define $\deg(\mathcal{V}, |\cdot|) := \deg(\mathcal{V}_L, |\cdot|)$, where $\mathcal{V}_L$ is the base extension of $\mathcal{V}$ to K.

If $K \subset L$ is a degree n extension of number fields, then the following direct computation shows that

$$\deg(\mathcal{V}_L, |\cdot|) = n \cdot \deg(\mathcal{V}, |\cdot|). \tag{A.2}$$

Indeed, pullbacks commute with top wedge power, so it suffices to check the equality when $\mathcal{V}$ is a line bundle, in which case

$$\sum_{\sigma' \colon L \to \mathbb{C}} \log |(\sigma')^* s|_{\sigma} = \sum_{\sigma \colon K \to \mathbb{C}} \left(\sum_{\sigma' | \sigma} \log |\sigma^* s|_{\sigma} \right) = \sum_{\sigma \colon K \to \mathbb{C}} n \cdot \log |\sigma^* s|_{\sigma}$$

and, since $\mathcal{O}_L$ is a flat $\mathcal{O}_K$-module,

$$|(\Gamma(\mathcal{V}) \otimes_{\mathcal{O}_K} \mathcal{O}_L)/\mathcal{O}_L \cdot s| = |(\Gamma(\mathcal{V})/\mathcal{O}_K \cdot s) \otimes_{\mathcal{O}_K} \mathcal{O}_L| = n \cdot |\Gamma(\mathcal{V})/\mathcal{O}_K \cdot s|.$$

A.4. Stacks

This generalizes to stacks in the following fairly formal way.

Let $\mathcal{X}$ be an algebraic stack, finite type over $\operatorname{Spec} \mathcal{O}_K$. We define a metrized vector bundle $\overline{\mathcal{V}}$ on $\mathcal{X}$ to be a vector bundle $\mathcal{V}$ on $\mathcal{X}$ together with, for every map $f \colon X \to \mathcal{X}$ from a variety X, a choice of metric on $f^*\mathcal{V}$ (in the sense of A.3) which we denote by $f^*|\cdot|$, and which is compatible with compositions in the following sense: For a map $g \colon X' \to X$ from an $\mathcal{O}_K$-variety X', there is a canonical isomorphism $g^*(f^*\mathcal{V}) \to (f \circ g)^*\mathcal{V}$, and we require that this isomorphism identifies $g^*(f^*|\cdot|)$ with $(f \circ g)^*|\cdot|$.

We again define direct sums, tensor products, alternating products and duals via the formulas from (A.1). Given a morphism $g \colon \mathcal{X} \to \mathcal{Y}$ and a metrized vector bundle $\overline{\mathcal{V}}$ on $\mathcal{Y}$, we define the pullback $g^*\overline{\mathcal{V}}$ to have underlying bundle $g^*\mathcal{V}$ and, for a map $f \colon X \to \mathcal{X}$ from an $\mathcal{O}_K$-variety X, define $f^*(g^*\overline{\mathcal{V}}) := (g \circ f)^*\overline{\mathcal{V}}$. For direct images, we restrict to the following special cases. Let $\overline{\mathcal{V}} = (\mathcal{V}, |\cdot|)$ be a metrized vector bundle on $\mathcal{X}$. If g is finite, flat and generically étale (and in particular representable), we define the direct image $g_*\overline{\mathcal{V}}$ to be the metrized vector bundle on $\mathcal{Y}$ which, for a map $f \colon Y \to \mathcal{Y}$ from a variety Y with corresponding fiber product

$$\begin{array}{ccc} X & \xrightarrow{f'} & \mathcal{X} \\ {\scriptstyle g'}\downarrow & & \downarrow{\scriptstyle g} \\ Y & \xrightarrow{f} & \mathcal{Y} \end{array}$$

pulls back to $f^*\left(g_*\overline{\mathcal{V}}\right) := g'_* f'^* \overline{\mathcal{V}}$. If instead g is proper, quasi-finite and birational, and $\mathcal{Y}$ is isomorphic to $\operatorname{Spec} \mathcal{O}_K$, then g is an isomorphism on a nonempty open subset $U \hookrightarrow \mathcal{X}$; we define $g_*\overline{\mathcal{V}}$ to have underlying bundle $g_*\mathcal{V}$ (which is a vector bundle by Proposition B.4) and the metric defined by $g^*|\cdot|$.

A.5. A detailed example

Let K be a number field, and let $X = \operatorname{Spec} \mathcal{O}_K$, considered as an $\mathcal{O}_K$ variety. We consider the trivial metrized vector bundle $(\mathcal{O}_X, |\cdot|)$ (where the trivial norms are defined in Subsection A.2). Explicitly, for an embedding $\sigma \colon K \hookrightarrow \mathbb{C}$, the scheme X_σ is simply $\operatorname{Spec} \mathbb{C}$, and the norm

$$|\cdot|_\sigma \colon \mathcal{O}^{\mathrm{an}}_{X,\sigma} \to \mathcal{C}_{X^{\mathrm{an}}_\sigma}$$

is the complex absolute value $\mathbb{C} \to \mathbb{R}_{\geq 0}$. Given a section $s \in \mathcal{O}_K$, $|\sigma^* s|_\sigma$ is equal to the complex absolute value $|\sigma(s)|$. Taking $s = 1$, we compute that the degree

$$\deg(\mathcal{V}, |\cdot|) = \log|\mathcal{O}_K/\mathcal{O}_K \cdot 1| - \sum_{\sigma\colon K \to \mathbb{C}} \log|\sigma^* 1|_\sigma = 0 - \sum_{\sigma\colon K \to \mathbb{C}} 0$$

is 0, as one would expect of a trivial bundle.

Next, let K be a number field, and again let $X = \operatorname{Spec} \mathcal{O}_K$, but now considered as a variety over $\mathbb{Z}$. We consider the 'trivial' metrized vector bundle $(\mathcal{O}_X, |\cdot|)$ (where the trivial norms are defined in Subsection A.2). This is the same as the pullback of the trivial bundle on $\operatorname{Spec} \mathbb{Z}$ along the map (of $\mathbb{Z}$ varieties) $\operatorname{Spec} \mathcal{O}_K \to \operatorname{Spec} \mathbb{Z}$. Explicitly, there is only one embedding $\sigma\colon \mathbb{Q} \hookrightarrow \mathbb{C}$, and the scheme X_σ is isomorphic to the disjoint union $\coprod_{\sigma'|\sigma} X_{\sigma'}$, where the coproduct is taken over the set of embeddings $\sigma'\colon K \hookrightarrow \mathbb{C}$ of K and where $X_{\sigma'} = X \times_{K,\sigma'} \mathbb{C}$ (i.e., considered as an $\mathcal{O}_K$ scheme); X_σ is thus a disjoint union of $[K : \mathbb{Q}]$ copies of $\operatorname{Spec} \mathbb{C}$. The norm

$$|\cdot|_\sigma\colon \mathcal{O}^{\mathrm{an}}_{X,\sigma} \to \mathcal{C}_{X^{\mathrm{an}}_\sigma}$$

is locally (on X_σ) again given by the complex absolute value. Label the embeddings $\sigma_1, \ldots, \sigma_n$, and let $s \in \mathcal{O}_K$. Then $\sigma^* s$ is equal to the tuple $(\sigma_1(s), \ldots, \sigma_n(s))$. Given our choice of base, it does not make sense to compute the degree. Note that this description is also the same as the restriction of scalars (as in Subsection A.3) of the trivial metrized bundle on $\operatorname{Spec} \mathcal{O}_K$ (as an $\mathcal{O}_K$ variety) from the previous paragraph.

Now, let $X = \operatorname{Spec} \mathcal{O}_K$ and $Y = \operatorname{Spec} \mathbb{Z}$, and let $\pi\colon X \to Y$ be the structure map. Consider the direct image $\pi_* \overline{\mathcal{O}_X} = (\pi_* \mathcal{O}_X, \pi_* |\cdot|)$, where we consider X as a variety over $\mathbb{Z}$ and where $|\cdot|$ is the trivial metric. Then $\pi_* \mathcal{O}_X \cong \overline{\mathcal{O}_K}$ and $\pi_* |\cdot|$ has the following description. Again, since our base is $\operatorname{Spec} \mathbb{Z}$, there is only one embedding $\sigma\colon \mathbb{Q} \hookrightarrow \mathbb{C}$; the scheme Y_σ is isomorphic to a single copy of $\operatorname{Spec} \mathbb{C}$, and the norm

$$(\pi_* |\cdot|)_\sigma\colon \left(\pi_* \mathcal{O}_{X,\sigma}\right)^{\mathrm{an}} \to \mathcal{C}_{Y^{\mathrm{an}}_\sigma}$$

is now a map of sheaves on a topological space which is a single point, and thus determined by the map of global sections

$$\mathcal{O}_K \otimes_{\mathbb{Z}} \mathbb{C} \cong \prod_{\sigma'|\sigma} \mathbb{C} \to \mathbb{R}_{\geq 0},$$

where the product is taken over the set of embeddings $\sigma'\colon K \hookrightarrow \mathbb{C}$ of K, which we label as $\sigma_1, \ldots, \sigma_n$. The map $\prod_{\sigma'|\sigma} \mathbb{C} \to \mathbb{R}_{\geq 0}$ is given by

$$(z_1, \ldots, z_n) \mapsto \sqrt{\sum |z_i|^2}$$

and the isomorphism $\mathcal{O}_K \otimes_{\mathbb{Z}} \mathbb{C} \cong \prod_{\sigma'|\sigma} \mathbb{C}$ is given by

$$\alpha \otimes 1 \mapsto (\sigma_1(\alpha), \ldots, \sigma_n(\alpha)).$$

We now compute the degree of $\pi_* \overline{\mathcal{O}_X}$. Let $\overline{\mathcal{V}} := \bigwedge^n \pi_* \overline{\mathcal{O}_X}$ be the top wedge power of $\pi_* \overline{\mathcal{O}_X}$, and choose a $\mathbb{Z}$ basis $\alpha_1, \ldots, \alpha_n$ of $\mathcal{O}_K$. Then $\bigwedge^n \mathcal{O}_K$ is a free $\mathbb{Z}$ module of rank 1 generated by the section $s = \alpha_1 \wedge \cdots \wedge \alpha_n$. We then compute that the degree is

$$\log|\Gamma(\mathcal{V})/\mathbb{Z} \cdot s| - \log|\sigma^* s|_\sigma = 0 - \log|\sigma^* s|_\sigma.$$

Next, we compute $\log|\sigma^* s|_\sigma$. The norm $\bigwedge^n \pi_*|\cdot|$ is given by the composition

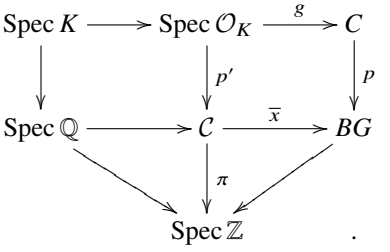
$$\left(\bigwedge^n \mathcal{O}_K\right) \otimes_{\mathbb{Z}} \mathbb{C} \cong \bigwedge^n (\mathcal{O}_K \otimes_{\mathbb{Z}} \mathbb{C}) \cong \bigwedge^n \prod_{\sigma'|\sigma} \mathbb{C} \cong \mathbb{C} \to \mathbb{R}_{\geq 0};$$

following s through these maps

$$\begin{aligned}(\alpha_1 \wedge \cdots \wedge \alpha_n) \otimes 1 &\mapsto (\alpha_1 \otimes 1) \wedge \cdots \wedge (\alpha_n \otimes 1)\\ &\mapsto (\sigma_1(\alpha_1), \ldots, \sigma_n(\alpha_1)) \wedge \cdots \wedge (\sigma_1(\alpha_n), \ldots, \sigma_n(\alpha_n))\\ &= \det(\sigma_j(\alpha_i)) \cdot (1 \wedge \cdots \wedge 1)\\ &\mapsto \left|\det(\sigma_j(\alpha_i))\right| = |\Delta_K|^{1/2}\end{aligned}$$

we conclude that $|\sigma^* s|_\sigma = |\Delta_K|^{1/2}$ and that the degree of $\pi_*\overline{\mathcal{O}_K}$ is $-\log|\Delta_K|^{1/2}$.

Finally: Let $C = \operatorname{Spec}\mathbb{Z}$, and let $BG = [C/G]$, with quotient map $p\colon C \to BG$. Let $\overline{\mathcal{V}} = \left(p_*\overline{\mathcal{O}_C}\right)^\vee$, where $\overline{\mathcal{O}_C}$ is the trivial metrized line bundle on C. (We dualize to facilitate the following quick global computation.) Let $x\colon \operatorname{Spec}\mathbb{Q} \to BG$ be a rational point corresponding to an extension $\mathbf{Q} \subset K$, and assume for this example that K is a number field (rather than just an étale algebra). We will now show that $\mathrm{ht}_{\overline{\mathcal{V}}}(x) = \log|\Delta_K|^{1/2}$. Let $\mathcal{C} = [\operatorname{Spec}\mathcal{O}_K/G]$. Then $\mathcal{C}$ is a tuning stack for x, summarized by the following diagram.



By definition, $\left(\overline{x}^*\overline{\mathcal{V}}\right)^\vee = p'_* g^* \overline{\mathcal{O}_C}$. Moreover, the tuning sheaf $\pi_* p'_* g^* \overline{\mathcal{O}_C}$ is isomorphic to $(p' \circ \pi)_* g^* \overline{\mathcal{O}_C}$, and $g^*\overline{\mathcal{O}_C}$ is the trivial metrized line bundle on $\operatorname{Spec}\mathcal{O}_K$ (as a $\mathbb{Z}$ variety). The height is then, by definition,

$$\mathrm{ht}_{\overline{\mathcal{V}}}(x) := -\deg\left((p' \circ \pi)_* g^* \overline{\mathcal{O}_C}\right);$$

we conclude that $\mathrm{ht}_{\overline{\mathcal{V}}}(x) = \log|\Delta_K|^{1/2}$.

Appendix B. One-dimensional Artin stacks with finite diagonal

In this appendix, we discuss a few technical aspects of the types of stacks that appear as the tuning stack of a rational point (Definition 2.1).

Fix a base scheme S. An Artin stack $\mathcal{C}$ (finite type over S) with finite diagonal admits a *coarse space* map $\pi\colon \mathcal{C} \to C$ [41, Corollary 1.3 (1)], which is (by definition) universal for maps to algebraic spaces and is a bijection on geometric points, and is moreover Stein (i.e., $\pi_*\mathcal{O}_\mathcal{C} \cong \mathcal{O}_C$) and a universal homeomorphism [61, Theorem 6.12]. If $S = \operatorname{Spec} k$ for some field k, then we say that $\mathcal{C}$ is geometric; if $S \to \operatorname{Spec}\mathbb{Z}$ is finite and flat, then we say that $\mathcal{C}$ is arithmetic.

Definition B.1. A stacky curve is a normal, one-dimensional Artin stack $\mathcal{C}$ with finite diagonal such that the coarse space map $\pi\colon \mathcal{C} \to C$ is birational and such that C/S is a proper curve if $\mathcal{C}$ is geometric and finite over S if $\mathcal{C}$ is arithmetic.

Normality of C follows from normality of $\mathcal{C}$, so C/k is a smooth proper curve in the geometric case and $C \cong \coprod \operatorname{Spec} \mathcal{O}_{K_i}$ for some number fields K_i in the arithmetic case. This is somewhat more general than the notion of stacky curve from [71, Chapter 5].

Our beginning lemma was pointed out to us by Sid Mathur.

Lemma B.2. *Let $\mathcal{C}$ be a stacky curve. Then $\mathcal{C}$ is regular.*

Proof. Since $\mathcal{C}$ is an Artin stack, it has a smooth cover $p\colon U \to \mathcal{C}$. Let $y \in \mathcal{C}(\Omega)$ be a geometric point. Then $\pi(y)$ is a geometric point of C. Since C has dimension at most 1, the point $\pi(y)$ has codimension at most 1 in C. Therefore, there exists a point $z \in U(\Omega)$ with $\pi \circ p(z) = \pi(y)$ such that z has codimension at most 1 in U. Since π is a coarse space map, $p(z) \simeq y$.

Since $\mathcal{C}$ is normal, U is as well and so z is a regular point of U. Therefore, there is an open neighborhood $V \subseteq U$ of z such that V is regular. Since the image of $p|_V : V \to \mathcal{C}$ contains $p(z) \simeq y$, we have found a smooth cover of a neighborhood of $y \in \mathcal{C}(\Omega)$ by a regular scheme. □

Proposition B.3. *There exists a finite flat surjection $p\colon C' \to \mathcal{C}$ with C' regular and with irreducible connected components. The composition $\pi \circ p\colon C' \to C$ is finite and flat.*

Proof. We may assume that $\mathcal{C}$ is connected. Since $\mathcal{C}$ has finite diagonal, we know from [25, Theorem 2.7] that there is a finite surjective map $p\colon C' \to \mathcal{C}$, where C' is a scheme. We can assume C' is normal by replacing it with its normalization. Since π is proper and quasi-finite, $q := \pi \circ p$ is proper and quasi-finite, hence finite. Since C is of dimension 1, so is C'. As C' is normal, it is regular. Since q is surjective, we can replace C' by one of its irreducible components which surjects onto C; note that this maintains surjectivity of p, as π is a bijection on geometric points. Since C and C' are regular, q is flat by [26, Corollary 18.17]. Similarly, since $\mathcal{C}$ is regular, letting $U \to \mathcal{C}$ be any smooth cover by a scheme, we see the pullback $p_U : C' \times_{\mathcal{C}} U \to U$ is a finite map between regular schemes. Again, [26, Corollary 18.17] tells us that p_U is flat and hence p is flat. □

Corollary B.4. *Let $\mathcal{E}$ be a vector bundle on $\mathcal{C}$. Then $\pi_*\mathcal{E}$ is a vector bundle.*

Proof. We can assume that $\mathcal{C}$ is connected. We claim that the canonical map $\mathcal{O}_{\mathcal{C}} \to p_*\mathcal{O}_{C'}$ is injective. It suffices to check this after passing to a smooth cover $\operatorname{Spec} A \to \mathcal{C}$. We see $C' \times_{\mathcal{C}} \operatorname{Spec} A \to \operatorname{Spec} A$ is finite, so the fiber product is of the form $\operatorname{Spec} B$. The induced map $\operatorname{Spec} B \to \operatorname{Spec} A$ is surjective, hence dominant, and $\operatorname{Spec} A$ is regular, hence reduced, so $A \to B$ is injective, proving our claim.

To finish the proof, tensor the injective map $\mathcal{O}_{\mathcal{C}} \to p_*\mathcal{O}_{C'}$ by with $\mathcal{E}$. This yields an injection $\mathcal{E} \to \mathcal{E} \otimes p_*\mathcal{O}_{C'} \cong p_*p^*\mathcal{E}$ (where the isomorphism is the projection formula) and hence an injection $\pi_*\mathcal{E} \to q_*p^*\mathcal{E}$. Since $p^*\mathcal{E}$ is a vector bundle and q is finite flat, we see $q_*p^*\mathcal{E}$ is a vector bundle, so $\pi_*\mathcal{E}$ is torsion-free and coherent. As C is regular of dimension 1, this implies $\pi_*\mathcal{E}$ is a vector bundle. □

We now address generalities about of the degree of a line bundle on an Artin stack. In the geometric case, if $\mathcal{C}$ is Deligne–Mumford, then Vistoli [70] developed a more general theory of intersection theory (see also [71, Chapter 5] for just the case of line bundles). In general, degrees of 0-cycles on stacks are not defined (see [24]), and in the Arakelov setting (as in A.1) some additional attention is needed even in the Deligne–Mumford case. However, we have shown in Proposition B.3 that every connected stacky curve $\mathcal{C}$ admits a finite flat surjection $C' \to \mathcal{C}$ with C' regular and irreducible, and by [25, Remark 2.8] this is all that one needs to develop intersection theory in our setting.

Definition B.5. Let $\mathcal{L}$ be a line bundle (resp. torsion sheaf) on $\mathcal{C}$, and let $p\colon C' \to \mathcal{C}$ be a finite and flat surjection from a regular scheme C'. We define the degree (resp. length) of $\mathcal{L}$ to be $\deg \mathcal{L} = \frac{1}{\deg(p)} \deg p^*\mathcal{L}$ (resp. $\operatorname{length} \mathcal{L} = \frac{1}{\deg(p)} \operatorname{length} p^*\mathcal{L}$).

Again, we emphasize the fact that in the arithmetic setting $\mathcal{L}$ is an Hermitian line bundle and we mean the Arakelov degree. For a torsion sheaf, the Archimedean contributions are 0 so there is no distinction.

Lemma B.6. *The degree (resp. length) of $\mathcal{L}$ is independent of the choice of p.*

Proof. Let $p_i\colon C_i \to \mathcal{C}$ be two such covers, and let C_3 be the normalization of some irreducible component of $C_1 \times_{\mathcal{C}} C_2$ such that the maps $q_i\colon C_3 \to C_i$ are both surjective (and thus finite and flat). We then have

$$\frac{\deg p_1^*\mathcal{L}}{\deg p_1} = \frac{\deg q_1^* p_1^*\mathcal{L}}{(\deg q_1)(\deg p_1)} = \frac{\deg q_2^* p_2^*\mathcal{L}}{(\deg q_2)(\deg p_2)} = \frac{\deg p_2^*\mathcal{L}}{\deg p_2}. \tag{B.7}$$

The proof for length is identical. □

Definition B.8. Let $f\colon \mathcal{C}' \to \mathcal{C}$ be a quasi-finite map of stacky curves. We define the degree of f to be the degree of the induced map $C' \to C$ of coarse spaces.

Lemma B.9. *Let $f\colon \mathcal{C}' \to \mathcal{C}$ be a quasi-finite map of stacky curves, and let $\mathcal{L}$ be a line bundle (resp. torsion sheaf) on $\mathcal{C}$. Then* $\deg f^*\mathcal{L} = \deg f \cdot \deg \mathcal{L}$ *(resp.* $\operatorname{length} f^*\mathcal{L} = \deg f \cdot \operatorname{length} \mathcal{L}$*).*

Proof. If $\mathcal{C}'$ is a scheme, then this follows from the definitions of degree. Let $p\colon C' \to \mathcal{C}'$ be a finite flat cover by a regular scheme C'. By [66, Tag 0CPT], f is proper; the composition $f \circ p$ is thus proper, quasi-finite and flat, and in particular finite. We then have

$$\deg f^*\mathcal{L} = \frac{\deg p^* f^*\mathcal{L}}{\deg p} = \deg f \frac{\deg p^* f^*\mathcal{L}}{(\deg p)(\deg f)} = \deg f \cdot \deg \mathcal{L}.$$

The proof for length is identical. □

Proposition B.10. *Let $0 \to \mathcal{V}' \to \mathcal{V} \to M \to 0$ be an exact sequence, where $\mathcal{V}' \to \mathcal{V}$ is a map of vector bundles (metrized, in the Arakelov case) and M is a finitely generated torsion sheaf on $\mathcal{C}$. Then*

$$\deg \mathcal{V} = \deg \mathcal{V}' + \operatorname{length} M.$$

Proof. In the geometric case, this is well known. In the Arakelov case, by Lemma B.9 we may assume that $\mathcal{C} = \operatorname{Spec} \mathcal{O}_K$ for some number field K. Since M is a torsion sheaf and thus has no Archimedean metric, the proof follows from the definition of degree (Equation A.1). □

Acknowledgement. It is a pleasure to thank Dan Abramovich, Jarod Alper, Eran Assaf, Frank Calegari, Antoine Chambert-Loir, Brian Conrad, Ratko Darda, Anton Geraschenko, Aaron Landesman, Pierre Le Boudec, Brian Lehmann, Aaron Levin, Siddharth Mathur, Lucia Mocz, Brett Nasserden, Martin Olsson, Fabien Pazuki, Rachel Pries, Tony Shaska, Jason Starr, Sho Tanimoto, Anthony Várilly-Alvarado, John Voight, Melanie Matchett Wood, Takehiko Yasuda and Xinyi Yuan for valuable conversations about the material in this paper.

The first author was supported by NSF grant DMS-1700884 and DMS-2001200, a Simons Foundation Fellowship and a Guggenheim Fellowship. The second author was supported by a Discovery Grant from the National Science and Engineering Board of Canada. The third author was partially supported by NSF grant DMS-1555048.

Conflict of Interest. The authors have no conflict of interest to declare.

References

[1] D. Abramovich and A. Várilly-Alvarado, 'Level structures on abelian varieties and Vojta's conjecture', *Compos. Math.* **153**(2) (2017), 373–394. With an appendix by Keerthi Madapusi Pera.

[2] D. Abramovich and A. Várilly-Alvarado, 'Campana points, Vojta's conjecture, and level structures on semistable abelian varieties', *J. Théor. Nombres Bordeaux* **30**(2) (2018), 525–532.

[3] D. Abramovich and A. Várilly-Alvarado, 'Level structures on Abelian varieties, Kodaira dimensions, and Lang's conjecture', *Adv. Math.* **329** (2018), 523–540.

[4] B. Alberts, 'Statistics of the first Galois cohomology group: A refinement of Malle's conjecture', *Algebra Number Theory* **15**(10) (2021), 2513–2569.

[5] S. A. Altuğ, A. Shankar, I. Varma and K. H. Wilson, 'The number of D_4-fields ordered by conductor', *J. Eur. Math. Soc. (JEMS)* **23**(8) (2021), 2733–2785.

[6] A. Bandini, I. Longhi and S. Vigni, 'Torsion points on elliptic curves over function fields and a theorem of Igusa', *Expositiones Mathematicae* **27**(3) (2009), 175–209.

[7] R. Beheshti, B. Lehmann, E. Riedl and S. Tanimoto, 'Rational curves on del Pezzo surfaces in positive characteristic', Preprint, 2021, arXiv:2110.00596. To appear, *Trans. Amer. Math. Soc. Ser. B.*

[8] G. Berhuy. *An Introduction to Galois Cohomology and Its Applications*, vol. **377** (Cambridge University Press, 2010).

[9] L. Beshaj, J. Gutierrez and T. Shaska, 'Weighted greatest common divisors and weighted heights', *J. Number Theory* **213** (2020), 319–346.

[10] M. Bhargava, 'Mass formulae for extensions of local fields, and conjectures on the density of number field discriminants', *Int. Math. Res. Not. IMRN* (17) (2007), Art. ID rnm052, 20.

[11] M. Bhargava and B. H. Gross, 'The average size of the 2-Selmer group of Jacobians of hyperelliptic curves having a rational Weierstrass point', in *Automorphic Representations and L-Functions*, Tata Inst. Fundam. Res. Stud. Math., vol. 22 (Tata Inst. Fund. Res., Mumbai, 2013), 23–91.

[12] M. Bhargava and P. Harron, 'The equidistribution of lattice shapes of rings of integers in cubic, quartic, and quintic number fields', *Compositio Mathematica* **152**(6) (2016), 1111–1120.

[13] W. Bolaños and G. Mantilla-Soler, 'The shape of cyclic number fields', Prepreint, 2022, arXiv:1912.07054. *To appear in Canadian Mathematical Bulletin.*

[14] S. Boucksom, J.-P. Demailly, M. Păun and T. Peternell, 'The pseudo-effective cone of a compact Kähler manifold and varieties of negative Kodaira dimension', *Journal of Algebraic Geometry* **22**(2) (2013), 201–248.

[15] D. Bourqui, 'Algebraic points, non-anticanonical heights and the severi problem on toric varieties', *Proceedings of the London Mathematical Society* **113**(4) (2016), 474–514.

[16] T. Browning and W. Sawin, 'Free rational curves on low degree hypersurfaces and the circle method', Preprint, 2018, arXiv:1810.06882.

[17] C. Cadman, 'Using stacks to impose tangency conditions on curves', *Amer. J. Math.* **129**(2) (2007), 405–427.

[18] B. Conrad, 'Keel–Mori theorem via stacks', Preprint, 2005.

[19] B. Conrad, 'Arithmetic moduli of generalized elliptic curves', *J. Inst. Math. Jussieu* **6**(2) (2007), 209–278.

[20] R. Darda, 'Rational points of bounded height on weighted projective stacks', Preprint, 2021, arXiv:2106.10120.

[21] R. Darda and T. Yasuda, 'The Batyrev–Manin conjecture for DM stacks', Preprint, 2022, arXiv:2207.03645.

[22] P. Deligne, 'Courbes elliptiques: formulaire d'après J. Tate', *in Modular Functions of One Variable, IV (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972)*, *Lecture Notes in Math.*, vol. 476, (1975), 53–73.

[23] A.-W. Deng, 'Rational points on weighted projective spaces', Preprint, 1998, arXiv:9812082.

[24] D. Edidin, A. Geraschenko and M. Satriano, 'There is no degree map for 0-cycles on Artin stacks', *Transform. Groups* **18**(2) (2013), 385–389.

[25] D. Edidin, B. Hassett, A. Kresch and A. Vistoli, 'Brauer groups and quotient stacks', *Amer. J. Math.*, **123**(4) (2001), 761–777.

[26] D. Eisenbud, 'Commutative Algebra', *Graduate Texts in Mathematics*, vol. 150 (Springer-Verlag, New York, 1995).

[27] N. D. Elkies, 'ABC implies Mordell', *International Mathematics Research Notices* **1991**(7) (1991), 99–109.

[28] J. S. Ellenberg and A. Venkatesh, 'Counting extensions of function fields with bounded discriminant and specified Galois group', in *Geometric Methods in Algebra and Number Theory* (Springer, 2005), 151–168.

[29] B. Fantechi, E. Mann and F. Nironi, 'Smooth toric Deligne–Mumford stacks', *J. Reine Angew. Math.* **648** (2010), 201–244.

[30] M. Fulger and B. Lehmann, 'Zariski decompositions of numerical cycle classes', *Journal of Algebraic Geometry* **26**(1) (2017), 43–106.

[31] X. Gao, 'On Northcott's theorem', PhD thesis, University of Colorado (1995).

[32] A. Geraschenko and M. Satriano, 'T'oric stacks I: The theory of stacky fans', *Trans. Amer. Math. Soc.* **367**(2) (2015), 1033–1071.

[33] A. Geraschenko and M. Satriano, 'A "bottom up" characterization of smooth Deligne–Mumford stacks', *Int. Math. Res. Not. IMRN* (**21**) (2017), 6469–6483.

[34] A. Granville, 'ABC allows us to count squarefrees', *Int. Math. Res. Not. IMRN* (**19**) (1998), 991–1009.

[35] H. Grauert and R. Remmert, *Coherent Analytic Sheaves*, Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 265 (Springer-Verlag, Berlin, 1984).

[36] R. Grizzard and J. Gunther, 'Slicing the stars: Counting algebraic numbers, integers, and units by degree and height', *Algebra & Number Theory* **11**(6) (2017), 1385–1436.

[37] Q. Guignard, 'Counting algebraic points of bounded height on projective spaces', *Journal of Number Theory* **170** (2017), 103–141.

[38] P. H. and R. Harron, 'The shapes of Galois quartic fields', *Transactions of the American Mathematical Society* **373**(10) (2020), 7109–7152.

[39] R. Harron, 'The shapes of pure cubic fields', *Proceedings of the American Mathematical Society*, **145**(2) (2017), 509–524.

[40] K. Kedlaya, 'A construction of polynomials with squarefree discriminants', *Proceedings of the American Mathematical Society* **140**(9) (2012), 3025–3033.

[41] S. Keel and S. Mori, 'Quotients by groupoids', *Ann. of Math. (2)*, **145**(1) (1997), 193–213.

[42] D. Knutson, 'Algebraic Spaces', Lecture Notes in Mathematics, vol. **203** (Springer-Verlag, Berlin-New York, 1971).

[43] A. Kobin, 'Artin–Schreier root stacks', *J. Algebra* **586** (2021), 1014–1052.

[44] A. Kresch and A. Vistoli, 'On coverings of Deligne–Mumford stacks and surjectivity of the Brauer map', *Bull. London Math. Soc.* **36**(2) (2004) 188–192.

[45] A. Landesman, 'A thesis of minimal degree: two', PhD thesis, Stanford University (2021).

[46] C. Le Rudulier, 'Points algébriques de hauteur bornée', PhD thesis, Rennes 1 (2014).
[47] B. Lehmann, A. K. Sengupta and S. Tanimoto, 'Geometric consistency of Manin's conjecture', Preprint, 2018, arXiv:1805.10580.
[48] A. Mânzăţeanu, 'Counting points on Hil$b^m\mathbf{P}^2$over function fields', Preprint, 2019, arXiv:1905.04772.
[49] D. Masser and J. D. Vaaler, 'Counting algebraic numbers with large height I, in *Diophantine Approximation* (Springer, 2008), 237–243.
[50] D. Masser and J. Vaaler, 'Counting algebraic numbers with large height II', *Transactions of the American Mathematical Society* **359**(1) (2007), 427–445.
[51] L. Meier, 'Vector bundles on the moduli stack of elliptic curves', *Journal of Algebra* **428** (2015), 425–456.
[52] L. Meier and V. Ozornova, 'Rings of modular forms and a splitting of $\mathrm{TMF}_0(7)$', *Selecta Mathematica* **26**(1) (2020), Paper No. 7.
[53] B. Nasserden and S. Y. Xiao, 'The density of rational points on $\mathbb{P}^1$with three stacky points', Preprint, 2020, arXiv:2011.06586.
[54] B. Nasserden and S. Y. Xiao, 'Heights and quantitative arithmetic on stacky curves', Preprint, 2021, arXiv:2108.04411.
[55] F. Pazuki, 'Modular invariants and isogenies', *Int. J. Number Theory* **15**(3) (2019), 569–584.
[56] E. Peyre, 'Hauteurs et mesures de Tamagawa sur les variétés de Fano', *Duke Math. J.* **79** (1) (1995), 101–218.
[57] E. Peyre, 'Liberté et accumulation', *Doc. Math.* **22** (2017), 1615–1659.
[58] E. Peyre, 'Beyond heights: slopes and distribution of rational points', in *Arakelov Geometry and Diophantine Applications*, Lecture Notes in Math., vol. 2276 (Springer, Cham, 2021), 215–279.
[59] M. Pizzo, C. Pomerance and J. Voight, 'Counting elliptic curves with an isogeny of degree three, *Proc. Amer. Math. Soc. Ser. B* **7** (2020), 28–42.
[60] R. Pries and H. J. Zhu, 'The p-rank stratification of Artin-Schreier curves', *Annales de l'Institut Fourier* **62** (2012), 707–726.
[61] D. Rydh, 'Existence and properties of geometric quotients', *J. Algebraic Geom.* **22**(4) (2013), 629–669.
[62] D. Rydh, 'Noetherian approximation of algebraic spaces and stacks', *J. Algebra* **422** (2015), 105–147.
[63] W. M. Schmidt, 'Northcott's theorem on heights I. A general estimate', *Monatshefte für Mathematik* **115**(1–2) (1993), 169–181.
[64] W. M. Schmidt, 'Northcott's theorem on heights II. The quadratic case', *Acta Arithmetica* **70**(4) (1995), 343–375.
[65] J. H. Silverman, 'The Arithmetic of Elliptic Curves', second edn., *Graduate Texts in Mathematics*, vol. 106 (Springer, Dordrecht, 2009).
[66] The Stacks Project Authors, Stacks Project, http://stacks.math.columbia.edu.
[67] J. Starr, Z. Tian and R. Zong, 'Weak approximation for Fano complete intersections in positive characteristic', Preprint, 2018, arXiv:1811.02466. To appear, *Ann. Inst. Fourier.*
[68] J. Starr and C. Xu, 'Rational points of rationally simply connected varieties over global function fields', Preprint, 2017, arXiv:1703.08334v1.
[69] J. L. Thunder and M. Widmer, 'Counting points of fixed degree and given height over function fields', *Bulletin of the London Mathematical Society* **45**(2) (2013), 283–300.
[70] A. Vistoli, 'Intersection theory on algebraic stacks and on their moduli spaces', *Invent. Math.* **97**(3) (1989) 613–670.
[71] J. Voight and D. Zureick-Brown, 'The canonical ring of a stacky curve', *Mem. Amer. Math. Soc.* **277**(1362) (2022), v+144.
[72] P. Vojta, 'A more general abcconjecture', *Internat. Math. Res. Notices* (**21**) (1998), 1103–1116.
[73] M. Widmer, 'Counting points of fixed degree and bounded height', *Acta Arithmetica* **140** (2009), 145–168.
[74] M. Machett Wood and T. Yasuda, 'Mass formulas for local Galois representations and quotient singularities. I: A comparison of counting functions', *International Mathematics Research Notices*, 2015.